Schistosoma mansoni: functional expression and characterisation of cathepsin L1 and cathepsin L2

Thesis presented for the degree of Ph.D.

by

Ciaran P. Brady, B.Sc.

under the supervision of

John P. Dalton, B.Sc., Ph.D.
School of Biotechnology
Dublin City University

December 1998

I hereby certify that this material, which I submit for assessment on the programme of study leading to the award of PhD is entirely my own work and has not been taken from the work of others save to the extent that such work has been cited and acknowledged within the text of my work.

Signed: Ciaran Brady. Date: 4/3/1998

Ciaran P. Brady

Acknowledegments

Firstly I would like to thank my supervisor, Dr. John Dalton, for providing me with the opportunity to pursue this project and for his guidance, enthusiasm and friendship throughout my time in DCU.

I would also like to thank all my fellow 'knackers', namely Andrew, Eamon, Sandra, Bernie, Jacinta, Sharon, Pepe, Leda and Paul, who also done their time in the molecular parasitology laboratory and provided many unforgettable moments, and many more forgettable ones!! To the above and all the postgrads in the Biology dept., both past and present, for the good times and never boring excursions! A special word of thanks to Andrew for his expert help with protein purifications, fermentations and enzyme kinetics and also for proof reading.

My grateful thanks to Prof. Paul Brindley for allowing me to work in his laboratory in Brisbane and looking after me so well during my time in Australia. I hope I wasn't too much of a burden! Also, many thanks to Lyn, Alec, Sharon, Bill, Sara, Joanna, Mishka, Christiana, Jason, Nayyar, Tom and Steve for their help and hospitality during my time in L16. Also, cheers to Luis, Joy, May-La, Andreas, Tonia and Stephanie; friday lunchtimes will never be the same again, keep up the good work in Toms' bar!! To all the other hard-drinking reprobates that I met Down Under, thanks for all the good times and friendship; love you all, love your work!!!

I am grateful to Cavan County Council, Forbairt and the Australian Vice-Chancellor's committee who have spent money on me throughout the course of this work.

Finally, to my family and friends (especially the other members of the magnificent seven!) for putting up with me and providing so much support over the years; yes, my studying days are at last drawing to a close, you will be glad to hear! A very special thanks to my parents, Tommy and Eileen, to whom I owe a dept that I will never be able to repay.

To my parents

Contents

Abstract

The cysteine proteinases, cathepsins L1 (SmCL1) and L2 (SmCL2) from the parasitic helminth *Schistosoma mansoni* were functionally expressed in *Saccharomyces cerevisiae*. By transforming the yeast with cDNAs encoding the complete pre-pro-enzymes, the recombinant proteinases were secreted into the culture medium from which they were purified by gel filtration and/or anion exchange chromatography. The enzymes were secreted by the yeast cells in their mature, active form and demonstrated characteristics typical of cathepsin L-like cysteine proteinases. Proteins in the purified fractions were reactive with anti-sera prepared against the recombinant enzymes which had been expressed in, and purified from extracts of *Escherichia coli*.

Yeast-expressed SmCL1 and SmCL2 displayed distinct differences in their specificities for synthetic peptide substrates; SmCL1 favours substrates containing aliphatic (P_3)-aliphatic (P_2)-positive or polar hydrophobic (P_1) residues, whereas SmCL2 showing a preference for aliphatic (P_3)-aromatic (P_2)-positive (P_1). SmCL1 cleaved a broader range of substrates tested and was also active over a wider pH range than SmCL2. SmCL1 showed a pH optimum of 6.5 for activity against Z-Phe-Arg-NHMec, while SmCL2 was optimally active at pH 5.5 against this substrate. RT-PCR and immunoblotting studies revealed expression profiles of these proteinases also differed in the various life cycle stages of *S. mansoni*; both proteinases are expressed in male and female adults but at different levels, and the pro-form of SmCL1 is also expressed in cercariae.

Immunolocalisation experiments demonstrated that SmCL1 was expressed in the gastrodermal cells lining the schistosome gut, as well as in the tegument of adult worms. Immunoblotting studies which detected the presence of SmCL1 in the excretory/secretory products of adult worms suggests that this enzyme is secreted by the parasite. Moreover, recombinant yeast-expressed SmCL1 had a marked preference for haemoglobin as substrate. Collectively, these results suggest that SmCL1 plays some role in the degradation of host haemoglobin in the schistosome gut.

SmCL2 has been localised to the reproductive organs of adult schistosomes by another laboratory and has been implicated to be involved in egg formation in female parasites. Therefore, SmCL1 and SmCL2 perform different functions in schistosomes. Both these enzymes now represent novel targets at which chemotherapy and/or immunoprophylaxis may be directed.

Abbreviations

ADH	Alcohol dehydrogenase
AEBSF	4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride
AFC	Aminofluoromethylcourmarin
BCIP	5-bromo-5-chloro-3-indolyl phosphate
Boc	t-butyloxycarbonyl
BSA	Bovine serum albumin
bSmCL1	Bacterially-expressed recombinant SmCL1
bSmCL2	Bacterially-expressed recombinant SmCL2
cDNA	Complementary DNA
CHN_2	diazomethylketone
DEAE	diethylaminoethyl
DFP	Diisopropylphosphofluoridate
DIG	Digoxygenin
DPP I	Dipeptidylpeptidase
DTT	dithiothreitol
EC	Enzyme classification
EDTA	Ethylenediaminetetraacetic acid
ES products	Excretory/secretory products
FITC	Fluorescein isothiocyanate
Hb	Haemoglobin
HIV	Human Immunodeficiency Virus
IL	Interleukin
IPTG	Isopropylthio-β-D-galactoside
k_{cat}	First order rate constant (turnover number)
K_m	Michelis-Menten constant
LAP	Leucine aminopeptidase
LB	Luria Bertani medium
M6P	Mannose-6-phosphate

NBT	Nitroblue tetrazolium
NHMec	Aminomethylcourmarin
NK	Natural killer
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
pI	Isoelectric point
QAE	Quaternary aminoethyl
RBC	Red blood cells
RPMI	Roswell Park Memorial Institute
SDS	Sodium dodecyl sulphate
SmCL1	*S. mansoni* cathepsin L1
SmCL2	*S. mansoni* cathepsin L2
Suc	Succinyl
TH	T helper cell
TIMP	Tissue inhibitor of metalloproteinases
Tos	Tosyl
TPI	Triose phosphate isomerase
Tris	Tris-(hydroxy-methyl)-methylamine(2-amino hydroxymetyl) propane-1,3-diol
WE	Adult schistosome worm extract
X-Gal	1,5-bromo-4-chloro-3-indoyl-β-D-galactosidae
ySmCL1	Yeast-expressed recombinant SmCL1
ySmCL2	Yeast-expressed recombinant SmCL2
Z	Benzoyl

Chapter 1

Introduction

1.1 : Helminth parasites

Parasitic infections can be generally divided into two groups; those caused by protozoa and those caused by helminths. Protozoa are unicellular, they multiply within the definitive host, and are frequently intracellular in habitat. Helminths, as defined by Anderson and May (1979) are, on the other hand, "macroparasites which tend to have much longer generation times than the protozoans. Direct multiplication within the host is either absent or occurs at a very low rate. The immune responses elicited by helminths generally depend on the number of parasites present in a given host, and tend to be of relatively short duration. These infections, therefore, tend to be of persistent nature with hosts being continually reinfected."

The study of helminths is regarded as the study of parasitic worms. Helminths typically parasitise vertebrates, although invertebrates act as intermediate hosts. The helminth diseases in man and domestic animals are caused by three groups of parasites belonging to the classes of trematoda (flatworms), nematoda (roundworms) and cestoda (tapeworms). Only a small proportion of individuals harbouring helminth infections, generally those individuals with heavy infection, are likely to develop overt disease. There is a low mortality/high morbidity rate associated with helminth infections. It is estimated that one quarter of the worlds population is infected with helminth parasites (Bundy, 1992), and about 150,000 die each year as a result of these infections (Bundy *et al.*, 1990).

1.2 : Schistosomiasis

Schistosomes, or blood flukes, are digenetic trematodes and are the causative agent of the parasitic disease schistosomiasis. Schistosomiasis, also known as Bilharzia, afflicts more than 250 million people in tropical regions of Central Africa, the Middle East, Eastern and South-East Asia, the Caribbean and in parts of South America, and more than 600 million are at risk of infection (World Health Organisation, 1996). It is estimated that 5-10% of heavily infected people will eventually die of their infection which accumulates to approximately 80,000 deaths per annum (Butterworth, 1989).

There are three common species which infect man, namely *Schistosoma mansoni*, *Schistosoma japonicum* and *Schistosoma haematobium*. The life cycle of *Schistosoma mansoni* is illustrated in Fig. 1.1. Eggs (Fig. 1.1, stage 1) that are passed in faeces of the primary host hatch into miracidiae in aggregations of water. The miracidiae (Fig. 1.1, stage 2) penetrate an intermediate host, the fresh water snail. In the snails, they develop two generations of sporocysts (Fig. 1.1, stages 3A and 3B), the second of which produces fork-tailed cercariae (Fig. 1.1, stage 4). These cercariae penetrate the skin (Fig. 1.1, stage 5) of the primary host when the host comes in contact with contaminated water. Once through the skin, the cercariae shed their tails and become schistosomulae which migrate through the tissues until they reach the portal venous system of the liver. Here, males and females (Fig. 1.1, stage 6) copulate before settling down in pairs in the mesenteric veins of the bowel or portal venous system of the liver. Adult females release numerous eggs each day and can do so for many years. The eggs move through the intestinal wall and are liberated into the lumen of the bowel, from which they reach the exterior.

Eggs are also carried with the circulation and are deposited in various body organs, particularly in the liver. The hosts inflammatory response to the eggs causes the tissue pathology associated with schistosomiasis (Mahmoud and Wahals, 1990). Eggs trapped in the liver elicit the development of an intense, cell-mediated granulomatous reaction which, with its associated fibrosis, is responsible for the clinical manifestations of portal hypertension and, in severe cases, rupture of the spleen and liver.

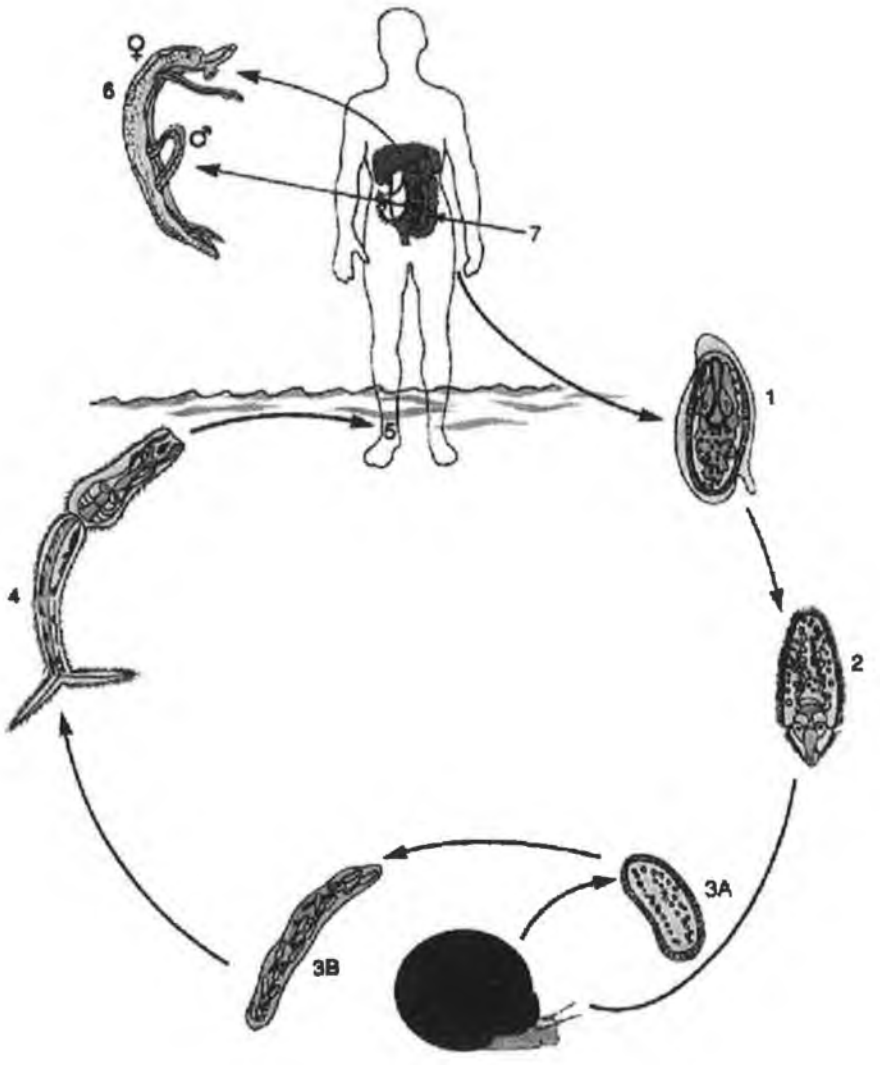


Fig 1.1 : Life cycle of *Schistosoma mansoni.* (Adapted from Peters and Gilles, 1995)

Traditionally, the control of schistosomiasis has concentrated on improvements in hygiene, removal of the intermediate host and chemotherapy using broad range drugs such as praziquantal. Praziquantal is, however, expensive and patients become reinfected in areas of high transmission and re-treatment must therefore be carried out frequently and for an indefinite period. Also, resistance to praziquantal has been reported (Brindley 1994a, 1994b) and it seems that resistant strains of the parasite are beginning to emerge. Resistance to praziquantel treatment has been induced in laboratory mice infected with *S. mansoni* (Brindley, 1994b).

Such drug resistance problems mean that the search for new strategies to control or eradicate schistosomiasis are essential. The successful eradication of schistosomiasis would involve the development of more effective and economically viable drugs, combined with an effective immunisation program designed to enhance host resistance to reinfection. Drugs with new modes of action and minimal toxicity to the host, as well as an effective vaccine, have still to be developed. The advent of recombinant DNA technology has opened up new opportunities in the identification of potential target molecules to which novel anti-schistosome chemotherapy and/or immunoprophylaxis could be directed.

1.3 : Proteinases

Proteinases are a complex group of enzymes that cleave peptide bonds. They display enormous variety in their physico-chemical and catalytic properties. Proteolytic enzymes are separated into two major groups - exopeptidases and endoproteinases. Exopeptidases remove one or more terminal amino acids from proteins and are further subdivided

according to whether they act at the amino (aminopeptidases, dipeptidylpeptidases, tripeptidylpeptidases, etc.) or the carboxyl terminus (carboxypeptidases). Endoproteinases cleave internal peptide bonds in polypeptide substrates and are classified into four major groups on the basis of their active sites, mechanism of peptide bond cleavage, and three-dimensional structure. These are serine-, cysteine-, aspartic- and metallo-proteinases.

1.3.1 : Endoproteinase groups

The serine proteinase group (EC 3.4.21) consists of two distinct families : the mammalian serine proteinases (e.g. chymotrypsin (EC 3.4.21.1), trypsin (EC 3.4.21.4) and elastase (EC 3.4.21.11)) and the bacterial serine proteinases (e.g. subtilisin (EC 3.4.21.14)). Serine proteinases contain an essential serine residue at the catalytic site and are characterized by the catalytic triad of His 57, Asp 102 and Ser 195 (chymotypsin numbering). Differences in substrate specificity are determined by changes within the S_1 subsite of the active site.

The cysteine proteinase group includes the plant proteinase papain (EC 3.4.22.2) and the cytosolic calcium-activated calpains (EC 3.4.22.17). In the case of papain, the thiol group of Cys 25 and the side chains of His 159 and Asn 175 are essential for the catalytic process. In contrast to the serine proteinases, cysteine proteinases require reducing agents to stabilise or enhance their activity. The lysosomal cathepsins, including cathepsins B, H and L belong to the papain family of cysteine proteinases. They are probably the most active cellular proteinases and have a wide specificity for peptide bonds (Bond and Butler, 1987). The substrate specificity of the different cathepsins is determined by interactions between the S_2 subsite and the substrate P_2 residue.

The aspartic proteinase group includes pepsin (EC 3.4.23.1), renin (EC 3.4.99.19), lysosomal cathepsin D (EC 3.4.23.5) and human immunodeficiency virus (HIV) proteinase. The aspartic proteinases are characterised by having maximal activity at low pH values. The catalytic process of these proteinases involves two aspartic residues in close proximity to each other in the centre of the substrate binding cleft. The substrate specificity of aspartic proteinases is extremely diverse (Neurath, 1990).

Metalloproteinases generally maintain optimum activity in the neutral pH range and contain a metal ion, usually zinc, which serves as a catalyst for hydrolysis. Removal of the metal ion by chelators inactivates these enzymes. This group contains the mammalian matrix metalloproteinases (collagenases and gelatinases) and bacterial thermolysin.

1.3.2 : Proteinase inhibitors

Many natural protein inhibitors of proteinases have been discovered and characterised including the serpins (serine proteinase inhibitors) , cystatins (cysteine proteinases) and TIMP (tissue inhibitor of metalloproteinases). Synthetic inhibitors are readily available which can be used to distinguish the group to which the proteinase under study belongs. The organophosphate diisopropylphosphofluoridate (DFP) and 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride (AEBSF) are commonly used inhibitors of serine proteinases. Pepstatin A, a pentapeptide-like compound, is routinely used to identify aspartic proteinases. Various metal chelators, including ethylenediaminetetraacetic acid (EDTA) and 1,10 phenathroline, which has a high affinity for zinc, are used to characterise metalloproteinases. The peptide epoxide L-*trans*-epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) is a potent, irreversible inhibitor of cysteine proteinases. Specific

synthetic peptide inhibitors are also available which can be used to distinguish between sub-classes of proteinases. Examples of these include Benzoyl(Z)-Phe-Ala-diazometylketone(CHN_2) and Z-Phe-Phe-CHN_2, which are specific, irreversible inhibitors of cathepsin Ls and Bs (Salvesen and Nagase, 1990).

1.4 : Proteinases of parasitic helminths

Proteinases can play many roles in host-parasite interactions. It is clear from examination of the life cycle of helminth parasites that proteinases perform key functions in their development. Several proteinases are elaborated at different stages of the life cycle to perform functions such as facilitating tissue invasion, evading the host immune response, triggering or contributing to morphological changes and allowing metabolism within the host.

Proteolytic enzymes must have originated very early in evolution, since all organisms require them for digestion and metabolism of their own proteins. In the early days of life, a limited set of peptidases with a broad substrate specificity, could have accounted for this activity (Jensen, 1976). It is reasonable to assume that from this limited set of ancestral enzymes a complex series of proteinases evolved by a process of gene duplication and divergence. Subsequent to this divergence, the enzymes would have acquired a higher degree of specificity by tailoring their catalysis to a discrete suite of peptide bonds located at specific sites in protein substrates. At the same time, the expression and distribution of these new proteinases could also be restricted, generating specialised proteinases for specific purposes (Neurath, 1984; Creighton and Darby, 1989).

This diversification of proteinases has undoubtedly contributed to the success of parasitic helminths. The main selection pressure that initiated the parasite-host relationship may have been the ease by which nutrients could be acquired from the host by the parasite (Halton, 1997) and hence it is found that much of the predominant proteolytic activity in helminth parasites are involved in this function. However, the development of mechanisms to facilitate the parasites migration through host tissue, to defend against host immunological attack, and to otherwise ensure the completion of their life cycle would obviously have presented other selection pressures. It is not unlikely that the diversification and specialisation of proteinases has played a major role in the evolution of these processes. It would, therefore, appear that proteinases are essential for the survival of parasites and would represent suitable candidate molecules to which novel vaccine and/or rational drug design could be targeted.

1.5 : Schistosome Proteinases

1.5.1 : Proteinases of eggs and miracidia

The schistosome egg releases soluble proteins through its shell, which include proteolytic enzymes. Various activities have been reported from eggs, however, the genes encoding these activities have yet to be characterised. Early reports of proteinase activity from *S. mansoni* eggs centred around three partially-purified thiol proteinases, all of which cleaved the cathepsin B-specific substrate Z-Arg-Arg-aminofluorocourmarin(AFC) (Asch and

Dresden, 1979; Sung and Dresden, 1986). These activities were localised to glands in the miracidium and were suggested to be involved in the passage of eggs through the wall of the intestine or bladder or in the penetration of the smooth muscle, extracellular matrix of the snail intermediate host by the hatched miracidiae. A novel cysteine proteinase activity from miracidiae and sporocysts of *S. mansoni* has been reported by Yoshino *et al.* (1993). The partially purified activity specifically hydrolyses Z-Phe-Arg-NHMec, diagnostic of cathepsin-L like activity. This activity was also suggested to be involved in penetration and infection of the snail.

Day *et al.* (1995) demonstrated that soluble extracts of eggs of *S. japonicum* contained cathepsin L-like activity with physical characteristics that were very similar to those of adult worms. They suggested that this activity may contribute to the pathology associated with schistosome eggs either directly, through its proteolytic activity or indirectly through immunological responses directed against the proteinase(s).

Leucine aminopeptidase (LAP) activity, a metalloproteinase, has been reported in secretions of *S. mansoni* eggs and hatching fluid. Bogitsh (1983) and Xu and Dresden (1986) demonstrated the presence of LAP activity in the germ cell of *S. mansoni* eggs, in the epidermis of the miracidium within the egg and also in the vitelline membrane surrounding the miracidium. They suggested that LAP plays a role in the hatching of the schistosome egg, possibly through the degradation of the egg shell to allow the release of the miracidium.

Doenhoff and co-workers have recently reported the purification of a serine proteinase, termed Sm480, from soluble extracts of *S. mansoni* eggs. This protein was

highly immunogenic and cross-reactive antigens were detected in miracidiae, cercariae and adult worms (Curtis *et al.*, 1996; Doenhoff, 1998).

While the genes encoding these proteinases have yet to be characterised, it is not unlikely that these may also be expressed in other life cycle stages, based on similarities of enzyme activities. Indeed, whereas complete cDNA sequences encoding cathepsin L (Smith *et al.*, 1994; Michel *et al.*, 1995) and cathepsin D (Becker *et al.*, 1995) have been characterized from adult schistosomes, partial cDNA sequences have been isolated for both of these from egg stage cDNA libraries as part of the ongoing World Health Organisation-sponsored *Schistosoma* Genome Initiative (Johnston, 1997).

1.5.2 : Proteinases of cercariae

The first step in human infection by schistosomes is the penetration of intact skin by free-living cercariae. The cercarial acetabular glands, the contents of which are released via ducts to the exterior following stimulation with skin lipids, contain multiple proteinases (Stirewalt, 1974; Minard *et al.*, 1997).

Various serine proteinases have been characterised from cercariae, however, of all schistosome proteinases reported to date, the best characterised is the cercarial elastase. This 28/30 kDa proteinase was purified from cercarial secretions and was shown to represent 90% of the total caseinolytic/gelatinlytic activity in these secretions (Marikovsky *et al.*, 1988b). Both skin penetration and glycocalyx (tail) shedding are believed to be facilitated by the activities of cercarial elastase (Marikovsky *et al.*, 1990a, 1990b; McKerrow *et al.*, 1991). Furthermore, a surface localised, schistosomular serine

proteinase with characteristics similar to cercarial elastase has been implicated in immunomodulatory roles including the cleavage of host complement components (Marikovsky *et al.*, 1988a, 1988b, 1990a, 1990b) and of host IgG and IgE bound to immune effector cells via Fc receptors (Auriault *et al.*, 1981; Verwaerde *et al.*, 1988).

Cercarial elastase has been considered a potential target for immunological intervention as blocking of its action may interfere with the infection by cercariae and the subsequent development of schistosomules (Pierrot *et al.*, 1996; Darani *et al.*, 1997; Doenhoff, 1998). Immunisation with the proteinase elicited specific antibody responses which partially protected against challenge infection with *S. mansoni* cercariae (Darani *et al.*, 1997; Doenhoff, 1998).

In contrast to serine-like proteinases, other classes of proteolytic enzymes of cercariae have not been studied comprehensively. However, cathepsin L-like and cathepsin B-like activity have been recently described in extracts of *S. mansoni* cercariae (Dalton *et al.*, 1997a). These activities were also localised to the post acetabular glands and expression of these proteinases continues after the cercariae transform into schistosomules within the definitive host, unlike the cercarial elastase or other serine proteinase described which are exclusively expressed by cercariae (Dalton *et al.*, 1997a).

1.5.3 : Proteinases of adults and schistosomules

Schistosome cDNA libraries have been principally constructed using mRNA from adult schistosomes and hence more is known about genes encoding proteinases in the adult stage than in any of the other stages of the life cycle. Of the proteinase genes from adults, a number encode proteinases that appear to play a role in the nutrition of the parasite and

these enzymes are discussed in detail in a later section (see section 1.6). In addition to proteinases involved in nutrition, a number of other proteinases have been described which appear to play roles in modulating the host immune response to the parasite.

Calpains are calcium-activated, cysteine proteinases, that are implicated in the regulation of cytoskeletal proteins, receptor proteins and protein kinases as well as in housekeeping roles such as intracellular protein turnover and membrane biogenesis in mammalian tissues. Calpains have a discrete large subunit of approximately 80 kDa and a small subunit of 30 kDa (Croall and Demartino, 1991). The large 80 kDa subunit appears to bear the proteolytic activity of calpains while the 30 kDa subunit plays a regulatory role in activating the large subunit (Goll *et al.*, 1992).

Calpain activities have been purified from soluble extracts of *S. mansoni* (Siddiqui *et al.*, 1993) and cDNAs encoding the large subunit isolated (Karcz *et al.*, 1991; Andersen *et al.*, 1991). The schistosome calpain has been localised to the surface epithelial layer and underlying muscle fibres and it appears that it is involved in the biogenesis of the schistosome surface membrane (Siddiqui *et al.*, 1993). Immunological studies demonstrated that calpain was the target antigen of a murine CD4+ T lymphocyte clone, named clone B, that transfers protective immunity against challenge infection with *S. mansoni* (Jankovic *et al.*, 1996). Mouse recipients of clone B displayed significant protection against cercarial challenge and, moreover, could kill schistosomula *in vitro*. Calpain is therefore considered a useful target for development of a schistosome vaccine. However, the recombinant 80 kDa subunit induced modest levels of protection (29-39%) against challenge infection of *S. mansoni* in mice (Hota-Mitchell *et al.*, 1997).

1.6 : Proteinases involved in host haemoglobin degradation

Maturing and adult schistosomes feed on host red blood cells (RBC) and it is estimated that male and female adult *S. mansoni* ingest 39,000 and 330,000 RBC per hour, respectively (Timms and Bueding, 1959). Female adults are the egg-producing partner and would therefore require more energy than the male. Host haemoglobin (Hb) derived from the ingested RBC is essential for growth, development and reproduction of schistosomes (Timms and Bueding, 1959; Zussman *et al.*, 1970; Jaffe *et al.*, 1971). Upon ingestion by the schistosome, RBC are lysed by the action of haemolysin(s) within the esophagus of the parasties (Bogitsh, 1978; Kasschau *et al.*, 1986). Hb released from RBC flows into the cecum of the schistosome where it is degraded to dipeptides or free amino acids which are either taken up or diffuse into the gastrodermal cells lining the schistosome cecum (Bogitsh, 1989) (Fig 1.2).

Nearly 40 years ago Timms and Bueding (1959) identified a proteolytic activity with a pH optimum of 3.9 and a substrate preference for Hb in extracts of adult *S. mansoni*. They concluded that 'the action of one or several proteolytic enzymes present in the alimentary canal of the worms' may be responsible for the degradation of Hb. Since then, an expanding number of proteinases, active at acidic and neutral pH, have been implicated in Hb digestion. However, the precise molecular identity of these enzymes and the proteolytic pathways involved in the degradation of Hb still remain unresolved. Nonetheless, it was believed that inhibition of these enzymes would result in profound anti-parasitic effects. Accordingly, elucidation of how schistosomes digest Hb may identify new targets or strategies for designing anti-schistosomal drugs or developing vaccines.

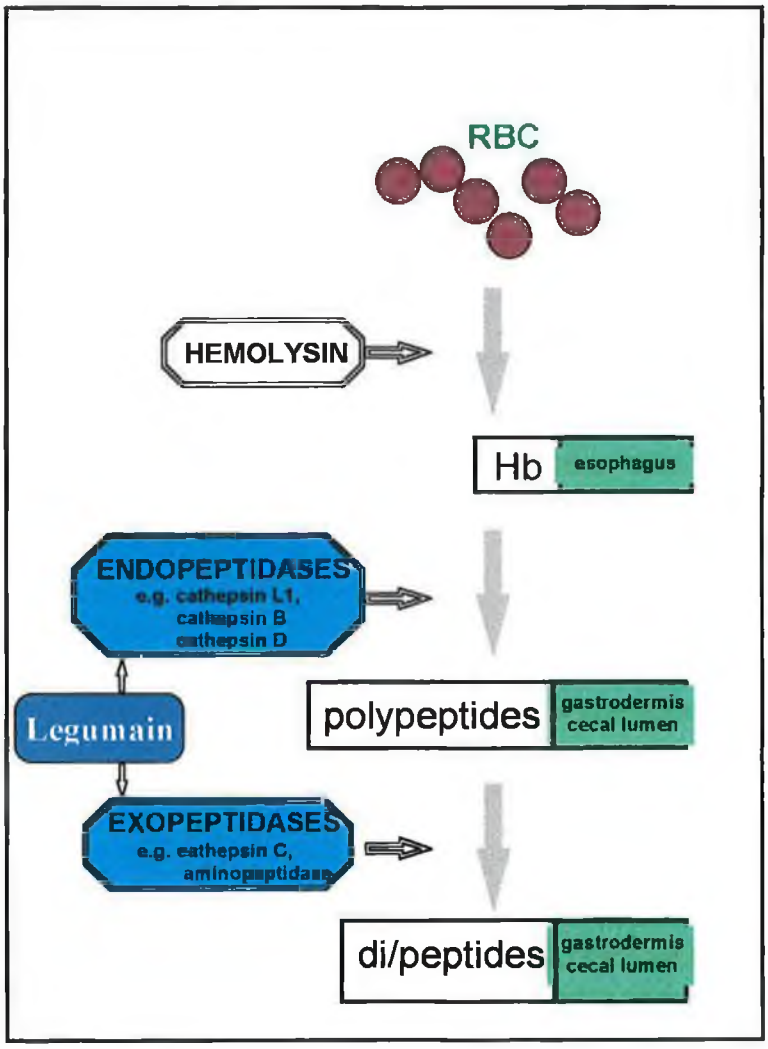


Fig. 1.2 : Schematic representation of putative mechanism of catabolism of host haemoglobin (Hb) from ingested red blood cells (RBC) by schistosomes (Adapted from Brindley *et al.*, 1997).

1.6.1 : Cathepsin B

Cathepsin B (EC 3.4.22.1) is a category of papain-like, cysteine proteinases, located in lysosomes in mammalian cells. The schistosome cathepsin B, known also as Sm31, was first reported by Ruppel and co-workers while investigating the diagnostic potential of schistosome antigens (Ruppel *et al.*, 1985). A 31 kDa protein was identified on immunoblots to be highly immunogenic in infected humans and mice and this protein was subsequently localised to the schistosome gut. Subsequently, a cDNA encoding Sm31 was isolated and sequence alignments revealed that the Sm31 gene encoded a cathepsin B, cysteine proteinase (Klinkert *et al.*, 1987, 1989). At that time, Sm31 gained prominence as an important enzyme in the degradation of host hemoglobin and was thus termed 'hemoglobinase' (McKerrow and Doenhoff, 1988).

Gotz and Klinkert (1993) expressed active, recombinant Sm31 in insect cells using baculovirus. Klinkert *et al.* (1994) reported a three-dimensional model for *S. mansoni* cathepsin B based on the co-ordinates of the crystal structure of human cathepsin B. While the enzymes are similar in structure, examination of the models complexed with specific inhibitors revealed likely differences in inhibition profiles. It was confirmed that Z-Trp-Met-CHN_2 is a more potent inhibitor of schistosome than of human cathepsin B, whereas the reverse is true with CA-074 (N-(L-3-trans-carboxyoxirane-2-carbonyl)-L-isoleucyl-L-proline), a derivative of the epoxide E-64.

Lipps *et al.* (1996) expressed recombinant procathepsin B from *S. mansoni* in *Saccaromyces cerevisiae*. The recombinant enzyme was produced as the inactive proenzyme which required exposure to porcine pepsin in order to be processed to the

mature, active zymogen. The recombinant protease had a pH optimum of 6.0 against Z-Arg-Arg-NHMec and was inactive below pH 4. The recombinant yeast-expressed Sm31 does not show a marked preference for Hb as substrate although it cleaves Hb at a few sites. Based on NH2-terminus analysis of digestion products of Hb, Lipps *et al.* (1996) calculated a consensus cleavage sequence of 6X1*181 (one, hydroxl, small aliphatic residue; six, aliphatic; eight, hydrophobic; *, cleavage point) (P3 P2 P1*P1' P2' P3').

As mentioned earlier, it had been widely suggested that Sm31 was responsible for the degradation of hemoglobin within the gut of *S. mansoni*. However, recent studies indicate that this enzyme may not function alone in this process. It now appears that a battery of proteases are involved, including Sm31, in addition to other cysteine proteases including cathepsin L (below), cathepsin C and possibly schistosome legumain, as well as a cathepsin D aspartic protease (Fig 1.2).

1.6.2 : Cathepsin L

Cathepsin L (like cathepsin B) is a cysteine proteinase of the papain superfamily. However, the prosegment of cathepsin Ls exhibits the ERFNIN motif of Karrer *et al.* (1993), which, together with other biochemical and structural characteristics, distinguishes cathepsin L and other cathepsins from cathepsin B. The ERFNIN motif is a distinctive set of conserved amino acid residues separated by three amino acids within a long, conserved alpha helix, constituting the core of a globular portion of the propeptide (Groves *et al.*, 1996; Coulombe *et al.*, 1996). The ERFNIN motif is present in all papain, cathepsin L-like and cathepsin H-like cysteine proteinases but absent from the cathepsin B-like enzymes. In mammals, cathepsin L (EC 3.4.22.15) represents a major component of the lysosomal

protein degradative system (Kirschke and Barrett, 1987). Human cathepsin L is also of particular interest because of evidence of its action in tumour invasion and metastasis (Yagel *et al.*, 1989).

Activity ascribable to cathepsin L has been characterized from adult and other stages of schistosomes. Dalton and colleagues analysed the cysteine proteinase activity in soluble extracts and ES products of *S. mansoni* and *S. japonicum* using class-specific, synthetic peptidyl substrates including several that could discriminate between classes of papain-like, cysteine endoproteinases, namely cathepsin L (Z-Phe-Arg-NHMec, Z-Phe-Val-Arg-NHMec), cathepsin B (Z-Arg-Arg-NHMec; Z-Arg-NHMec) and cathepsin H (Z-Arg-NHMec). They concluded that the dominant proteinase activity between pH 4 to pH 6 was cathepsin L-like (Smith *et al.*, 1994; Day *et al.*, 1995; Dalton *et al.*, 1996), in contrast to cathepsin B as reported by other workers (Chappell and Dresden, 1986; Lindquist *et al.*, 1986; Wasilewski *et al.*, 1996). This activity against cathepsin L substrates at acidic pH is many fold greater than that for cathepsin B substrates, although cathepsin B activity with a pH optimum of ~6 against Z-Arg-Arg-NHMec is apparent (Dalton *et al.*, 1996).

The activity of schistosome cathepsin L in tissue extracts of adult *S. mansoni* against the substrate Z-Phe-Arg-NHMec can be totally blocked by nanomolar concentrations of the diazomethanes Z-Phe-Ala-CHN_2 and Z-Phe-Phe-CHN_2. Z-Phe-Ala-CHN_2 (K_i 50% = 50 nM at pH 5) is a more potent inhibitor than Z-Phe-Phe-CHN_2 (K_i 50% = 0.4 μm), which contrasts with mammalian cathepsin L where Z-Phe-Phe-CHN_2 is the more potent inhibitor (Smith *et al.*, 1994; Day *et al.*, 1995). Cathepsin L activity with

similar biochemical properties has been detected in extracts of schistosome eggs (Day *et al.*, 1995).

cDNAs encoding two cysteine proteinases with similarity to vertebrate cathepsins Ls have been characterized from *S. mansoni* (Smith *et al.*, 1994; Michel *et al.*, 1995) and homologues of both genes reported from *S. japonicum* (Day *et al.*, 1995). Using PCR with degenerate primers designed to anneal to regions of cysteine proteinase genes encoding the active site residues Cys25 and Asn179, Smith *et al.* (1994) isolated the first cathepsin L from adult *S. mansoni* cDNA. The full length transcript encoded a preproenzyme of 319 amino acid residues, including the signal peptide of 8 residues, a prosegment of 96 residues, and the mature enzyme of 215 residues. The apparent molecular mass of the mature enzyme is 24.3 kDa, with pI of 4.95. This proteinase has been termed SmCL1 (*S. mansoni* Cathepsin L1) (Smith *et al.*, 1994; Day *et al.*, 1995) since a second cathepsin L gene has been reported from adult *S. mansoni* (Michel *et al.*, 1995). A cDNA encoding the apparent homologue, SjCL1, from *S. japonicum* has also been reported. While the cDNA encoding SjCL1 did not encode the full-length preproenzyme, its deduced amino acid sequence revealed the mature proteinase would include 215 amino acid residues, of similar length to SmCL1, with a predicted molecular mass of 24.1 kDa and pI of 5.63. The deduced amino acid sequences of SjCL1 and SmCL1 are 92% identical, and their eight S2 subsite residues are identical, indicative of identical substrate preferences (Day *et al.*, 1995).

The cDNA encoding the second cathepsin L-like proteinase has been reported from *S. mansoni* (Michel *et al.*, 1995) and *S. japonicum* (Day *et al.*, 1995). This proteinase was termed SmCL2 and SjCL2, in each species (Day *et al.*, 1995). The cDNA

encoding the *S. mansoni* proteinase was isolated by subtractive hybridisation undertaken to locate female specific transcripts in adult *S. mansoni*. The *S. japonicum* homologue was isolated by heterologous DNA screening of a *S. japonicum* cDNA library using the radiolabelled cDNA sequence from *S. mansoni* as the probe. The preprocathepsin SmCL2 contains 330 amino acids, including a signal peptide of 18, a pro-peptide of 97, and a mature enzyme of 215 amino acid residues. Its apparent molecular mass is predicted to be 24.3 kDa, with a pI of 9.25. Active site labelling and immunoblotting of schistosome extracts demonstrated that native SmCL2 is ~31 kDa rather than 24.3 kDa as predicted from its deduced amino acid sequence, indicating that it may be glycosylated (Michel *et al.*, 1995; Wasilewski *et al.*, 1996). The amino acid sequence deduced from the cDNA encoding SjCL2 is comprised of 332 amino acid residues, with the signal peptide being comprised of 16 residues, the prosegment of 99 residues, and the mature enzyme of 217 residues. The apparent molecular mass of SjCL2 is predicted to be 24.5 kDa, with a pI of 8.98. The deduced amino acid sequences of SjCL2 and SmCL2 are 78% identical, and share five of eight S2 subsite residues (Day *et al.*, 1995).

SmCL2 is over-expressed in females compared to male schistosomes and, in particular, is expressed in the reproductive system of the female parasite, and in the gynecophoric canal of the male. Based on this localisation, Michel *et al.* (1995) postulated that its role may be associated with the activation of phenyloxidase, an enzyme involved with cross-linking of eggshell proteins. They postulated, further, that SmCL2 may play a role in altering the viscosity of the seminal and uterine fluids.

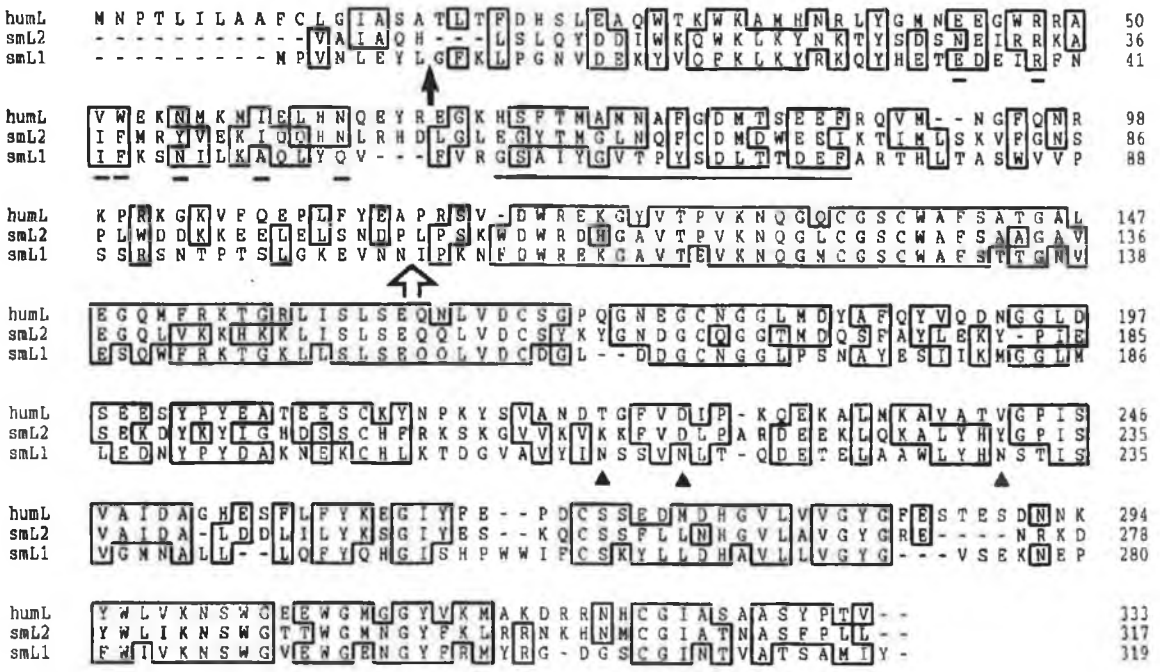


Fig 1.3 : Alignment of the deduced amino acid sequence of SmCL1 and SmCL2 with that of human cathepsin L (humL). Boxes denote conserved residues, and gaps have been introduced to maximise alignment. The putative signal peptide cleavage site is indicated by the solid arrow. The open arrow indicates the cleavage site between the propeptide and the mature enzyme. The ERFNIN-like motif and the conserved block in the propeptide are underlined. Cathepsin L1 potential N-glycosylation sites are indicated with solid triangles. (Adapted from Dalton *et al.*, 1996).

The low level of identity (43%; see Fig. 1.3) of the deduced amino acid sequences between SmCL1 and SmCL2 indicates that they are distinct proteins rather than products of alleles of the same gene (Dalton *et al.*, 1996). Also, SmCL1 and SmCL2 differ in five of the six residues that constitute the S_2 subsite of the active site, which determines the specificity of the proteinase for its substrate; SmCL1 contains Leu-67, Pro-68, Gly-133, and Ala 162 and Val 208, whereas cathepsin L2 contains Thr-69, Met-70, Ala-135, Gly 163 and Asn 209 (refer to Fig 5.1).

Moreover, SmCL1 contains 3 potential N-linked glycosylation sites in it mature region. SmCL2, by contrast, has a single potential site at Asn-209, however, since this Asn residue is contained within the S_2 subsite, it is unlikely that this site would be glycosylated. Together, these structural differences suggest that SmCL1 and SmCL2 perform separate functions (Dalton *et al.*, 1996).

1.6.3 : Cathepsin D

Cathepsin D is a member of the aspartic proteinase sub-class of endoproteinases. cDNAs encoding cathepsin D have been isolated from *S. japonicum* (Becker *et al.*, 1995) and *S. mansoni* (Wong *et al.*, 1997). The schistosome proteinases are 48-55% identical to cathepsin D from mammalian species. However, the schistosome cathepsin Ds contain a carboxyl extension of about 43 residues which is absent from the mammalian cathepsin D. Other differences include the absence of a β-hairpin loop 3, which is cleaved during maturation of vertebrate cathepsin Ds to yield heavy and light chain subunits, differences in residues involved in substrate binding and specificity and in phosphorylation (Becker *et al.*, 1995; Wong *et al.*, 1997).

Cathepsin D has been localised to the gastrodermal cells of both *S. mansoni* and *S. japonicum*, with higher levels of expression observed in female compared to male worms (Bogitsh and Kirschner, 1986, 1987; Bogitsh *et al.*, 1992; Bogitsh and Brindley, unpublished data). Cathepsin D-like activity has been observed in soluble extracts and ES products of *S. japonicum* and *S. mansoni* (Becker *et al.*, 1995). This activity was optimal at pH 3.0 to 4.0 and could be blocked by pepstatin. Further, the extracts and ES products digested human haemoglobin at a similar pH optimum and this activity was also inhibited by pepstatin (Becker *et al.*, 1995). Together, these results indicate that that schistosome cathepsin D plays a central role in haemoglobin catabolism. Cathepsin D represents a potential target for anti-schistosomal therapies in view of its function and also differences between schistosome D and the homologous human cathepsin D may be exploitable in the design of anti-proteinase drugs.

1.6.4 : Cathepsin C

Cathepsin C, otherwise known as dipeptidylpeptidase I (DPP I) (EC 3.4.14.1) is a member of the papain family of cysteine proteinases. However, unlike papain or cathepsin L, DPP I is an exopeptidase which hydrolyses the removal of dipeptides from the amino terminus of the substrate. Mammalian cathepsin C exhibits an array of biochemical properties distinct from other papain-like enzymes including a dependence on halide ions for enzymatic activity. Furthermore, cathepsin C enzymes possess prosegments comprised of approx. 200 amino acid residues, compared to 60 to 100 residues of other papain-like proteinases (McGuire *et al.*, 1997).

Bogitsh and Dresden (1983) detected a cathepsin C-like activity in extracts of *S. mansoni* and *S. japonicum* and these activities were localised to the gastrodermis and the gut luminal surface. More activity was present in female worms compared to males and they proposed that cathepsin C was involved in haemoglobin degradation. Butler *et al.* (1995) and Hola-Jamriska *et al.* (1998) have reported the sequence of a cathepsin C from *S. mansoni* and *S. japonicum* respectively. Northern blot analysis showed that *S. japonicum* cathepsin C was expressed in greater quantities in female as compared to male worms (Hola Jamriska *et al.*, 1998). Exoproteinases such as cathepsin C probably play a downstream role of degradation of Hb, acting on cleavage fragments released after degradation by endoproteinases (Fig 1.2).

1.6.5 : Schistosome legumain (Asparaginyl endopeptidase)

Asparaginyl endopeptidases are members of a novel family of cysteine proteinases which are termed legumains (EC 3.4.22.34), since they were first characterised from seeds of leguminous plants. These enzymes cleave peptide bonds on the carboxyl side of Asn residues except where the Asn occurs at the NH_2 terminus or at the second position from from the NH_2 terminus or when the Asn is glycosylated. Legumains function in the post-translational modification of storage proteins in legume seeds by cleaving asparaginyl peptide bonds between pro-peptides and mature proteins (Ischii, 1994).

Like cathepsin B Sm31 proteinase, a second schistosome antigen, termed Sm32, was first identified as an immunogenic component of soluble extracts of schistosomes (Ruppel, 1985a, 1987). Davis *et al.* (1987) isolated the cDNA encoding Sm32, however, its identity was in question as it showed no obvious similarity to other proteinases in the

databases. The sequence of Sj32, the *S. japonicum* homolog of Sm32, has also been reported (Mercklebach *et al.*, 1994). However, the identity of Sm32/Sj32 was clarified when sequences encoding asparaginyl endopeptidases from seeds of legumes were shown to be similar to Sm32 (Ischii, 1994; Dalton *et al.*, 1995a). Subsequently, Dalton *et al.* (1995b) demonstrated the presence of asparaginyl endopeptidase activity in soluble extracts of *S. mansoni* using diagnostic substrates developed by Kembhavi *et al.* (1993). The schistosome activity showed similar substrate specificity and inhibitor profile to the asparaginyl endopeptidase of the jack bean *Canavalia ensiformis*.

Monoclonal antibodies have been used to localise *S. mansoni* legumain to the epithelium of the gut in schistosomula and adults, to the ventral surface of adult males and to the cephalic glands of cercariae (Zhong *et al.*, 1995). Historically, schistosome legumain has been considered to be centrally involved in the degradation of Hb, however, its function remains unclear. Its low specific activity against synthetic peptide substrates (Dalton *et al.*, 1995b) argues against a central role in Hb digestion, given the enormous numbers of RBC that schistosomes ingest (Lawrence, 1973). Further, Hb-degrading activity of the cysteine proteinases reputed to be involved in Hb digestion can be blocked by E-64 and diazomethanes (Wasilewski *et al.*, 1996; Becker et al, 1995), neither of which are potent inhibitors of schistosome legumain (Dalton *et al.*, 1995b).

Dalton and Brindley (1996) suggested that, as with plant legumains (Ischii, 1994), schistosome legumain may process other proteins. Indeed, it may be involved in activation of the proteolytic enzymes, such as cathepsins L, D and C, which are directly involved in Hb degradation (Fig. 1.2). In support of this theory, the primary sequence of the cathepsins L1, L2, B, C and D contain asparagine residues in vicinity of the cleavage point

of the pro- and mature forms of the enzymes (Dalton and Brindley, 1996). If this is the case, Dalton and Brindley (1996) suggested that drugs targeted at schistosme legumain may block the activation of other proteolytic enzymes directly involved in Hb digestion and thereby deliver a profound anti-schistosomal effect. Such a strategy would have to be selective, however, as a legumain has recently been reported from mammalian tissues (Chen *et al.*, 1997).

1.7 : Aims of this project

The dominance of cathepsin L activity in extracts of schistosome tissues and in ES products (Dalton *et al.*, 1996) suggest that cathepsin Ls may be involved (with other proteinases) in Hb degradation. Recent studies on the related parasitic trematode *Fasciola hepatica* (liver fluke) have demonstrated that cathepsin Ls are involved in immune evasion, tissue penetration and nutrition mechanisms in these parasites (Smith *et al.*, 1993; Carmona *et al.*, 1993; Dowd *et al.*, 1994). Since SmCL2 has been localised to the reproductive organs and other sites distinct from the gut and gastrodermis of the schistosome (Michel et al, 1995), it is unlikely to be involved in Hb catabolism. One of the aims of this project was to determine if SmCL1 possibly functions in the digestion of host Hb in *S. mansoni*.

A second aim of this project was to compare and contrast the biochemical characteristics of SmCL1 and SmCL2 with regards to their substrate specificity. As described earlier, SmCL1 and SmCL2 have many structural differences and the two enzymes are likely to perform separate functions (refer to section 1.6.2). Moreover, SmCL1 shows only 44.3% identity to human cathepsin L, while SmCL2 is 52.3% identical

(Dalton *et al.*, 1996; see Fig. 1.3). These enzymes are considered potential targets at which to direct anti-schistosomal therapies. Indeed, the potential anti-schistosomal effects of drugs targeted at cysteine proteinases has been demonstrated by Wasilewski *et al.* (1996) using morpholinourea-Phe-Ala-CHN_2 and analogues. While these drugs would be inhibitors of both cathepsin L and B, they produced dramatic reductions in schistosome worm loads and fecundity in infected mice and were lethal to cultured schistosomula. Active site mapping using peptide substrates would provide leads into the design of specific peptide inhibitors of these enzymes.

Because parasite extracts contain many proteolytic activities, and additionally because of the paucity of material that can be obtained, it is virtually impossible to isolate individual proteinases directly from schistosome parasites. In order to study the biochemical characteristics of these enzymes, the immediate aim of this project was to express the proteinases in an active form in a recombinant expression system. The cathepsin L proteinases of *Fasciola hepatica* have been functionally expressed in the brewers yeast *Saccharomyces cerevisiae* (Dowd *et al.* 1997; Roche et al, 1997). The same system was employed for expression of the schistosome enzymes in the present study.

Chapter 2

Functional expression of *Schistosoma mansoni* cathepsin L1 (SmCL1) in *Saccharomyces cerevisiae*; purification of the recombinant enzyme, and localisation of the native enzyme in schistosome tissues

2.1 : Materials

Bachem

Synthetic peptide substrates and inhibitors

Biomeda

'Crystal/Mount' medium

Biorad

5-bromo-5-chloro-3-indoyl phosphate (BCIP), DC protein assay kit, goat anti-rabbit IgG conjugated to alkaline phosphatase, nitro-blue tetrazolium (NBT), prestained molecular weight markers.

Boehringer Mannheim

DIG glycan detection kit

Calbiochem

Goat anti-rabbit IgG-flourescein isothiocyanate (FITC) conjugate pre-adsorbed to bovine, horse, human and mouse sera

Difco Bacto

Yeast nitrogen base

Gibco BRL, Life technologies Ltd

Synthetic oligonucleotides

Pharmacia LKB Biotechnology

Sephacryl S200 HR, DEAE Sepharose, nitrocellulose

Promega

Agarose, Deoxytrinucleotides, DNA molecular weight markers, restriction enzymes, DNA gel extraction kit, T4 DNA ligase, Taq DNA polymerase, Wizard DNA miniprep system.

Queensland Institute of Medical Research

DNA sequencing (Applied Biosystems)

Sakura Finitek

Tissue-Tek O.C.T. medium

Sigma Chemical Company

2-mercaptoethanol, bovine serum albumin (BSA), bromophenol blue, Coomassie brilliant blue R, D-glucose, dithiothreitol (DTT), ethidium bromide, gelatin, peptone, sodium dodecyl sulphate (SDS), streptomycin, tetracycline, Tween 20, uracil, yeast extract.

2.2 : Methods

2.2.1 : Cloning of *Schistosoma mansoni* cathepsin L1 into yeast expression vector

The isolation of the complete preprocathepsin SmCL1 from an adult *S. mansoni* cDNA library has been previously described (Smith *et al.*, 1994). Primers designed to anneal to the 5' and 3' termini of the SmCL1 cDNA, termed SmCL1F and SmCL1R were synthesised.

SmCL1F: CGC AAG CTT ATG CCT GTA AAC CTC GAG TAC

SmCL1R: CGC AAG CTT CCC CTA GTA GAT CAT CGC TGA

The primers included *Hin*d III recognition sequences (underlined). These primers were used to re-amplify SmCL1 by PCR from the original template which had been cloned into pGEM-T as described by Smith *et al.* (1994) for insertion into the yeast expression vector pAAH5 (Ammerer, 1983). Cycling conditions in the PCR were as follows : 5 min at 94°C initial denaturing step, 25 cycles of denaturation at 94°C, primer annealing at 50°C for 1 min and extension at 74°C for 1 min, and a final extension step of 5 min at 74°C. Amplified fragments were electrophoretically separated in a 1% agarose-Tris-acetate gel. The gel was stained with ethidium bromide and the a band of the expected size (approx. 1 kb) corresponding to the SmCL1 amplified product was excised from the gel and eluted from the agarose using a Gel Extraction Kit (Promega) according to the protocols

provided by the manufacturer. The purified fragment was digested with *Hin*d III and ligated to *Hin*d III linearized pAAH5 as similarly described by Sambrook *et al.* (1989).

2.2.2 : Transformation of *E. coli* and preparation of plasmid DNA

Escherichia coli strain MC1061 was used as the host for pAAH5 plasmid and recombinant construct propagation and manipulations. Fresh competent cells were prepared using the calcium chloride method and transformation was performed according to standard protocols (Sambrook et al. 1989). *E. coli* cells were normally cultured on Luria Bertani (LB) medium. One hundred µg/ml ampicillin was added to plates and liquid media of cultures of cells harbouring plasmids. Positive or recombinant colonies were picked from the plates and replicated in fresh media. Plasmid DNA was prepared using the Wizard miniprep DNA purification system (Promega) according to the suppliers protocols. An internal *Eco* R1 site within the SmCL1 cDNA was used to determine the correct orientation of the SmCL1 insert within the vector by restriction mapping. DNA was visualised by electrophoresis on agarose-TAE gels. A clone with the correct orientation for expression was isolated and named pAAH5.SmCL1.

2.2.3 : DNA sequencing

The complete DNA sequence of selected recombinant clones was determined by an automated method (Applied Biosystems) at the DNA Sequencing Unit, The Queensland Institute of Medical Research, Brisbane, Australia.

2.2.4 : Transformation and culturing of *Saccharomyces cerevisiae*

Saccharomyces cerevisiae strain DBY746 (Mat α *his*3-D1-*leu*2-3 *leu*2-112 *ura*3-52 *trp*1-289a) (Yeast Genetic Stock Centre, Department of Biophysics and Medical Physics, University of California, Berkeley, CA, USA) was transformed with the pAAH5.SmCL1 construct or pAAH5 containing no insert by the lithium acetate method (Carter *et al.*, 1987). Yeast transformants were cultured on selective minimal media (Bacto Yeast Nitrogen Base 6.7 g/l,, D-glucose 20 g/l, uracil 20 mg/ml in 0.1M phosphate buffer, pH 5.5, polymerised with 20 g/l Difco no.1 agar) at 30°C for 3-4 days. DBY746 strain was routinely maintained in buffered complex media (YEPD): Yeast extract 10 g/l, peptone 20 g/l, D-glucose 20 g/l in 0.1M phosphate buffer pH 6.5.

2.2.5 : Expression and purification of yeast-expressed recombinant cathepsin L1 (ySmCL1).

Recombinant yeast cells were grown in selective minimum media at 30°C in flasks with vigorous agitation and the OD_{600} was monitored. Initially, ten clones were cultured in 300 mls of medium and analysed for expression of recombinant SmCL1 (ySmCL1). Cultures were harvested when they reached the early stationary phase. The cells were collected by centrifugation, and the supernatant assayed for cathepsin L proteinase activity using the fluorogenic substrate Z-Phe-Arg-NHMec (see below). One clone producing the most enzyme was used for further analysis. For large-scale expression of yeast recombinant SmCL1 (ySmCL1), yeast cells were grown in an automative fermentor (New Brunswick, model 101) in selective minimal media at 30°C until the OD_{600} reached 1.4. Yeast cells were removed by centrifugation and the supernatant stored at 4°C.

Ten litres of pAAH5.SmCL1-transformed yeast supernatant was concentrated at 4°C to 20 ml in an Amicon 2000A concentrator using an Amicon YM3 membrane (3000 Da molecular mass cut-off). The concentrate was applied to a Sephacryl S200HR (Pharmacia) gel filtration column (2.6 x 74.5 cm) equilibrated in 0.1M Tris-HCl, pH 7.0, at 4°C. Proteins were eluted from the matrix with the 0.1M Tris-HCl, pH 7.0, and fractions (5 ml) containing cathepsin L activity, measured using the fluorogenic substrate Z-Phe-Arg-NHMec (see below), were pooled. The pooled fractions were concentrated to 3 ml, dialysed against 20 mM Tris-HCl, pH 7.0, and applied to a QAE-Sephadex A50 column (10 cm x 2.5 cm) (Pharmacia), equilibrated in the same buffer. The column was washed with the equilibration buffer (375 ml) after which bound molecules were eluted in a 0 - 500 mM NaCl gradient. Fractions (5 ml) containing cathepsin L activity were pooled, concentrated as above, and stored at -20°C.

2.2.6 : Fluorimetric enzyme assays

Routine assays to detect cathepsin L-like activity in supernatents and purified fractions were carried out using the fluorogenic peptide Z-Phe-Arg-NHMec as substrate. The substrate was stored as a 1 mg/100 µl stock solution in dimethylformamide. Assays were carried out using a final concentration of 10 µM substrate in 0.1 M sodium phosphate buffer, pH 6.5, containing 1mM dithiothreitol, in a volume of 1 ml. The solutions were incubated at 37°C for 1 h before the reaction was terminated by the addition of 0.2 ml of 1.7 M acetic acid. The amount of NHMec released was measured with a fluorimeter (370

nm excitation and 440 nm emission). One unit of enzyme activity was defined as that amount which catalyzed the release of 1 μmole NHMec/min at 37°C.

2.2.7 : Adult worm extract (WE) and Excretory/Secretory (ES) products

Adult worms, freshly perfused from the mesenteric veins of mice, were sonicated in PBS using a Branson Sonicator set at repeating duty cycle 2.5 and output of 25%. Prior to five x 1 min sonication cycles, cellular debris was removed by centrifugation at 12,000 xg for 30 min at 4°C. Supernatants were removed, aliquoted and stored at -20°C until required. For the preparation of ES products, freshly perfused worms were cultured overnight at 37°C in RPMI-1640 medium supplemented with 20 μg/ml streptomycin and 20 μg/ml tetracycline. Medium was removed and concentrated twenty fold in Centricon 10 cartridges (Amicon) according to the manufacturers instructions and the concentrated supernatant stored at -20°C.

2.2.8 : Protein concentration estimation

Measurements of protein concentration were accomplished using the DC Protein Assay kit from Bio-Rad according to the manufacturers instructions. Bovine serum albumin was used as a protein standard.

2.2.9 : SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were analysed by one dimensional 12% SDS-PAGE gels according to the method of Laemmli (1970). Samples were prepared in reducing sample buffer (0.12 M

Tris-HCl, pH 6.8, 5% (w/v) SDS, 10% (w/v) glycerol, 0.01% (w/v) Bromophenol Blue and 5% 2-mercaptoethanol) and boiled for 5 min. Gels were stained for protein with Coomassie brilliant blue R.

2.2.10 : Immunoblotting

Adult schistosome worm extracts (WE), ES products, concentrated supernatant from yeast transformed with pAAH5.SmCL1 and purified recombinant enzyme were separated by reducing SDS-PAGE and electrophoretically transferred to nitrocellulose paper using a semi-dry electroblotting system. Following blocking of non-specific binding in 1% bovine serum albumin in TBST (10 mM Tris, pH 8, containing 150 mM NaCl, 0.1% Tween), the nitrocellulose membrane was incubated in anti-SmCL1 or non-immune rabbit serum at 1 in 500 dilution in the same buffer. Alkaline phosphatase-conjugated anti-rabbit IgG was used to detect the bound immunoglobulin using nitro-blue tetrazolium (NBT) and 5-bromo-4-chloro-indolyl phosphate (BCIP) as substrates for the alkaline phosphatase.

2.2.11 : Zymography

Purified ySmCL1 was analysed for proteinase activity in non-denaturing gels containing co-polymerized gelatin according to the method of Dalton and Heffernan (1989). Briefly, proteins were applied to 10% non-SDS polyacrylamide gels, with the seperating gel containing 0.1% gelatin. After electrophoresis, the gels were washed for 2 X 15 mins in 0.1 M phosphate buffer, pH 6.5. The gels were then incubated in the same buffer overnight at 37°C. Gels were subsequently stained with coomassie and destained in 20% (V/V) methanol, 10% acetic acid.

2.2.12 : Glycosylation studies

Glycoproteins were detected using the DIG Glycan Detection Kit (Boehringer Mannheim) according to the manufacturers instructions. Briefly, proteins were separated on 12% SDS-polyacrylamide gels and transferred to nitrocellulose as described above (see section 2.2.10). Filters were then incubated in 10 mM sodium metaperiodate, in 0.1 M sodium acetate buffer pH 5.5 for 20 min at room tempeture. After washing in phosphate buffered saline (PBS) the filter was incubated in 1 µl of digoxigenin-0-3-succinyl-ε-aminocaproic acid hydrazide dissolved in 5 ml sodium acetate buffer, pH 5.5 for 1 h at room tempeture. Digoxigenin (DIG), a spacer linked steroid hapten, covalently binds to aldehyde groups of sugars of glycoconjugates. DIG labeled glycoproteins are subsequently detected in an enzyme immunoassay using a DIG-specific antibody conjugated with alkaline phosphatase. Transferrin and creatinase were used as positive and negative controls, respectively.

2.2.13 : Hemoglobin proteolysis by ySmCL1

Hemoglobin was prepared as described previously (Becker *et al*, 1995). Hemoglobin (150 µg) was incubated with ySmCL1 (20 µg) at 37°C for 18 h in the presence of 1 mM dithiotreitol. Digestions was carried out in the following buffers; 0.1M sodium acetate pH 4.0 and 4.5; 0.1M sodium citrate, pH 5.0 and 5.5 and 0.1M sodium phosphate, pH 6.0 and 6.5, with the ionic strength of each buffer equalized to 100 mM using NaCl. Following the incubation, the reaction products were analyzed by 15% SDS-PAGE under non-reducing conditions and staining with Coomassie Brilliant Blue R.

2.2.14 : Preparation of rabbit anti-SmCL1 serum

Recombinant SmCL1 was expressed in *Escherichia coli* and purified by affinity chromatography as described in Chapter 4 (section 4.2.2). The bacterially-expressed, recombinant protein (termed bSmCL1) was used to immunise a rabbit and produce a highly specific anti-bSmCL1 serum (section 4.2.4).

2.2.15 : Immunolocalisation of SmCL1 in adult worms.

Mixed-sex adult worms were perfused from mice and then embedded in Tissue-Tek O.C.T. medium (Sakura Finetek) after which 10 μM sections were cut using a cryostat microtome. The sections were mounted on glass slides and air dried for 4 h. Sections were fixed in ice-cold acetone for 2 min, washed in PBS and incubated in goat normal serum diluted 1:5 in PBS for 30 min to inhibit non-specific background with the secondary antibody (below). After washing in PBS, sections were incubated in rabbit anti-bSmCL1 or control serum diluted 1:200 in PBS/1% bovine serum albumin (BSA) for 1 hour. Sections were washed in PBS, and incubated in goat anti-rabbit conjugated to fluorescein isothiocyanate (Calbiochem) diluted 1:100 in PBS/1% BSA (The conjugated antibody had been pre-adsorbed to bovine, horse, human and mouse sera by the manufacturer). All incubations were performed at room temperature. After further washing in PBS, sections were mounted in "Crystal/Mount" medium (Biomeda), viewed under UV light on an Olympus BX60 microscope, and photographed.

2.3 : Results

2.3.1 : Cloning of SmCL1 in pAAH5 expression vector

In order to analyse the biochemical properties of *S. mansoni* cathepsin L1, the cDNA encoding the complete pre-pro-cathepsin L was subcloned into the *Saccharomyces cerevisiae* expression plasmid pAAH5. pAAH5 is a shuttle vector with the yeast replication region of the 2-micron circle and the *E. coli* replication region of pBR322. The yeast LEU2 gene for leucine synthesis provides a selective marker in *S. cerevisiae* and the ampicillin resistance gene is used for selection of transformants in *E. coli.* The *Hin*d III cloning site is flanked at the 5' side by the promotor and the untranslated leader of the yeast alcohol dehydrogenase gene ADC1 containing the ribosome-binding site (Ammerer, 1983). The SmCL1 insert provided the translation and termination codons itself. The signals required for post-translational processing and intracellular sorting of the proenzyme are encoded by the prosegment-encoding sequences within the SmCL1 gene (Fig 2.1).

The complete cDNA encoding pre-pro SmCL1 was amplified using two primers (SmCL1F and SmCL1R) with *Hind* III recognition sequences incorporated. A band of the expected size (approx. 1 kb) was amplified and this fragment digested with *Hind* III and ligated into *Hind* III linearized pAAH5 plasmid. *E. coli* strain MC1061 cells were transformed with resulting constructs and transformed clones were selected on LB medium containing ampicillin. Ten clones harbouring plasmids were selected and analysed for the presence of ligated insertions. Plasmid DNA was purified from the clones, digested with *Hin*d III and

visualised on agarose gels. An example of a clone containing the expected 1 kb insert is shown in Fig. 2.2, lane 2.

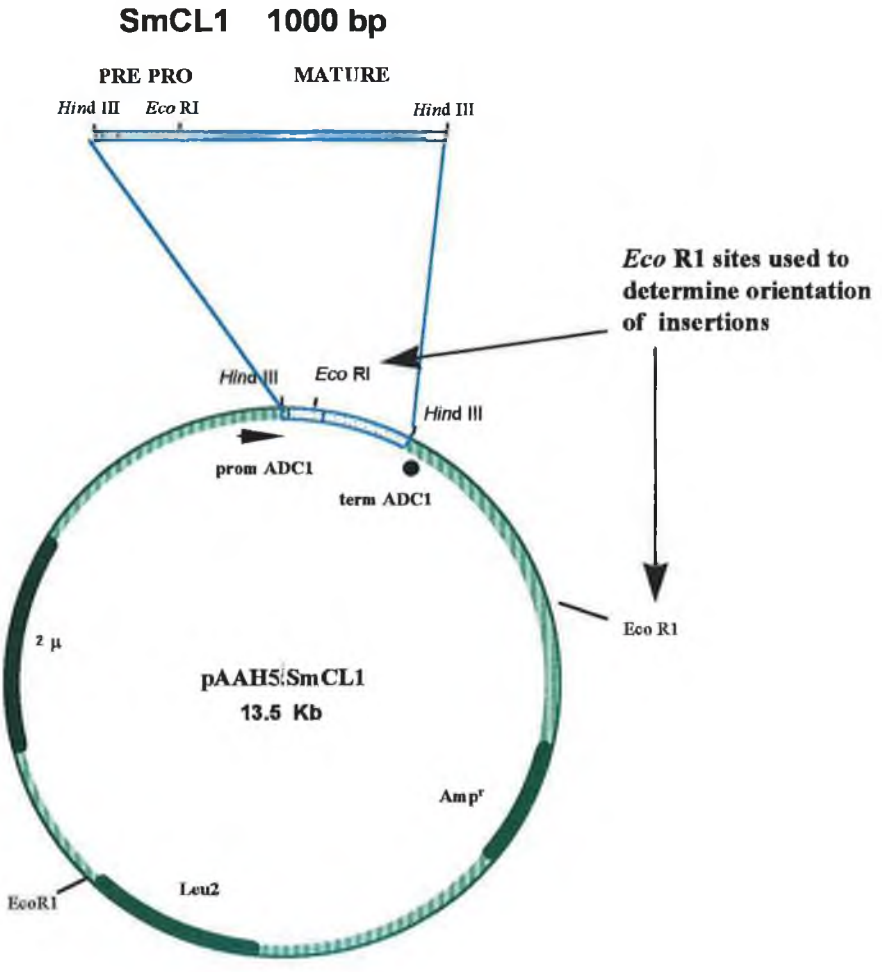


Fig 2.1 : Diagrammatic representation pAAH5.SmCL1 construct

An internal *Eco* R1 site within the SmCL1 cDNA allowed us to determine the correct orientation of the SmCL1 insert within the vector (Fig. 2.1). Plasmids containing the insert from a number of clones were digested with *Eco* R1 and examined for the presence of a 2.4 Kb band indicating that the SmCL1 cDNA had ligated into the vector in the

correct orientation for expression (Fig. 2.2, lane 3). Constructs containing inserts that had ligated in the wrong orientation resulted in the presence of a 1.6 kb band instead of the 2.4 kb fragment. A clone with the correct orientation for expression was selected for further experiments and this clone was termed pAAH5.SmCL1. Sequence analysis of the complete insert confirmed that no mutations had occured during PCR amplification.

Fig 2.2 : Cloning of SmCL1 into the yeast expression vector, pAAH5

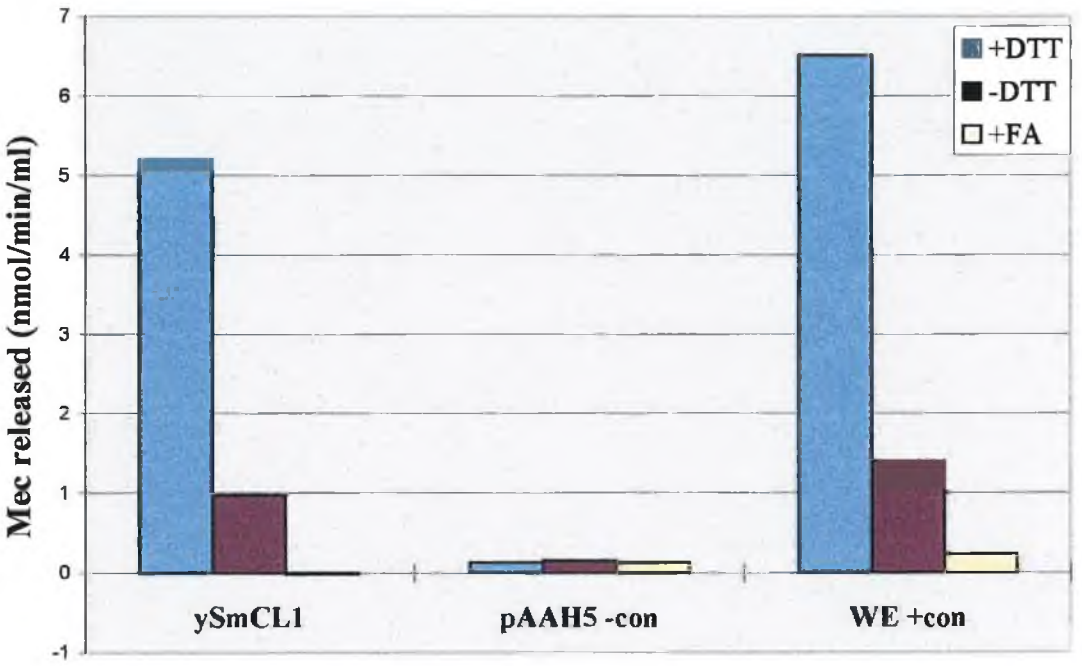
TAE agarose gel showing cloning and orientation of insert. Lane 1; 1 kb molecular weight ladder. Lane 2; digestion of pAAH5.SmCL1 construct with *Hin*d III reveals the expected 1 kb fragment; Lane 3; digestion with *Eco* R1 shows the 2.4 kb band demonstrating the insert had ligated into the vector in the correct orientation for expression.

2.3.2 : Expression of SmCL1 in yeast

Yeast strain DBY746 cells were transformed with the pAAH5.SmCL1 construct or the pAAH5 plasmid containing no insertion as a negative control for expression. Recombinant yeast clones were selected on minimal medium and ten clones were chosen for initial studies. The selected clones were cultured in shake flasks until cells entered early stationary phase and supernatants were assayed for cathepsin L activity using the fluorogenic peptide substrate benzyloxycarbonyl-phenylalanine-arginine-7-amido-4-methylcoumarin (Z-Phe-Arg-NHMec). This substrate is efficiently cleaved by cathepsin Ls in the presence of a reducing agent, in this case dithiothreitol (DTT), releasing the fluorogenic leaving group methylcourmarin (Mec). This fluorescence can be used to quantify the level of cathepsin L activity (Barrett and Kirschke, 1981).

Supernatants harvested from cells transformed with the pAAH5.SmCL1 construct showed significant activity against Z-Phe-Arg-NHMec (Fig 2.3). This activity was enhanced over five-fold by DTT and was completely inhibited by the cathepsin L-specific inhibitor benzyloxy carbonyl-phenylalanine-alanine-diazomethylketone (Z-Phe-Ala-CHN_2), results characteristic of cathepsin L-like, cysteine proteinase activity. Supernatants of yeast transformed with pAAH5 plasmid exhibited no activity, while *S. mansoni* crude worm extract, which contains cathepsin L activity (Dalton et al., 1996) showed a similar profile to supernatants from cells transformed with pAAH5.SmCL1. A clone displaying the highest level of cathepsin L activity was chosen and used for further studies. Expression was scaled up by subsequently culturing this clone in an automative 10 L fermenter. The fermenter allowed optimisation of yeast growth by control of parameters such as

temperature, aeration and agitation resulting in enhanced expression levels of the recombinant SmCL1.



Legend : FA = Z-Phe-Ala-CHN_2

ySmCL1 = supernatant from yeast cells transformed with pAAH5.SmCL1

pAAH5 -con = supernatant form yeast transformed with pAAH5 control plasmid

WE +con = *S. mansoni* crude worm extract positive control

Fig 2.3 : Fluorimetric enzyme assay for the presence of cathepsin L activity in supernatants of yeast transformed with pAAH5.SmCL1

Supernatants of yeast cells transformed with pAAH5.SmCL1 or plasmid pAAH5 were assayed using Z-Phe-Arg-NHMec as substrate in the presence or absence of the reducing agent DTT or the specific cathepsin inhibitor Z-Phe-Ala-CHN_2 (Results are the means of experiments performed on ten clones).

2.3.3: Purification of recombinant ySmCL1

Yeast-expressed recombinant SmCL1 (ySmCL1) was purified from yeast culture media by gel filtration on Sephadex S200HR followed by ion exchange chromatography on QAE-Sephadex matrix. Fractions were assayed for cathepsin L activity using Z-Phe-Arg-NHMec as substrate as outlined above. Z-Phe-Arg-NHMec-cleaving activity resolved as two peaks on the S200 HR matrix (Fig 2.4A). Fractions corresponding to both peaks were separately pooled and concentrated by ultrafiltration using an Amicon 50A stir-cell concentrator fitted with a 3 kDa cut-off membrane. Although the total activity was greater in the first peak (peak I) compared to the second peak (peak II), further purification was performed with the enzyme pool of peak II since it contained proteinase with much higher specific activity (peak I, 0.169 U/mg, peak II, 0.55 U/mg) (Table 1). The activity in both peaks was completely inhibited by Z-Phe-Ala-CHN_2 and enhanced by DTT (data not shown).

When the peak II activity was applied to the ion exchange QAE-Sephadex matrix, a large proportion of the proteolytic activity did not bind to the resin and was collected in the run-through fractions (Fig 2.4B). Subsequent elution of the column with a 0-500 mM NaCl gradient resulted in the release of proteolytic activity over a wide range of fractions (fig. 2.4B). However, this pool exhibited a low specific activity of 0.098 U/mg. In contrast, the proteinase in the run-through from QAE Sephadex exhibited a specific activity of 2.5 U/mg which represented a 42-fold enrichment compared to the fermentation supernatant (Table 1).

TABLE 1 : Chromatographic purification of yeast-expressed recombinant SmCL1 on gel filtration and ion exchange matrices.

	Protein (mg)	Activity (U)	Sp. activity (U/mg)	Yield (%)	Purification fold
Culture medium	53.1	3.24	0.06	100	1
S200HR Peak I	12.52	2.12	0.169	65.4	2.82
S200HR Peak II	2.03	1.12	0.55	34.6	9.17
QAE-Sepadex Pool	0.18	0.45	2.5	14.0	42.0

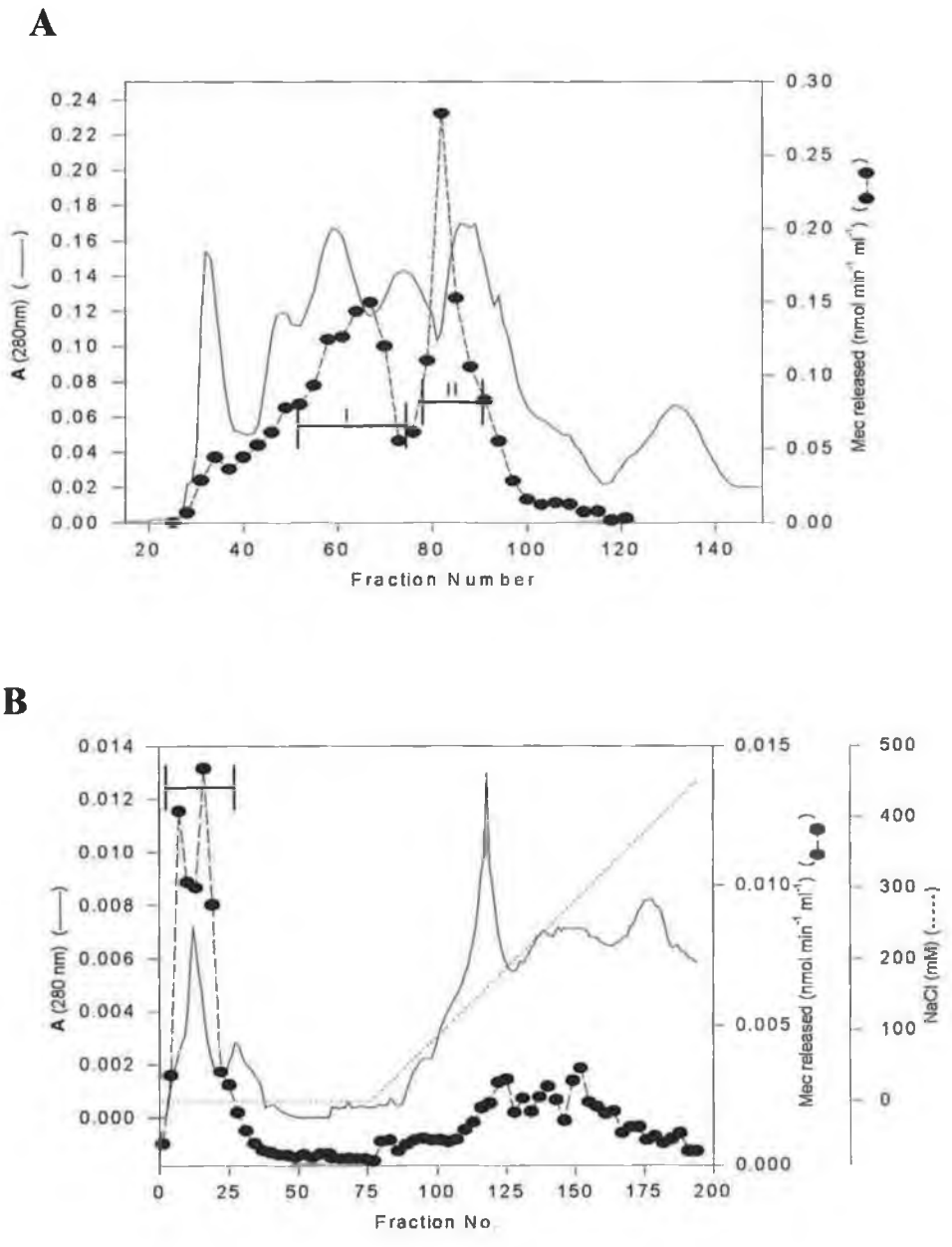


Fig 2.4 : Purification of ySmCL1

Profiles of elution from a Sephacryl S200HR column (A) and QAE-Sephadex column (B). Protein elution was monitored by absorbance at 280nm (lines) and cathepsin L activity was assayed using the synthetic, fluorogenic substrate Z-Phe-Arg-CHN_2 (dots). The salt gradient (0-500 mM NaCl) during QAE-Sephadex chromatography is represented by the dotted line and fractions pooled are indicated.

2.3.4 : SDS-PAGE and immunoblotting

The run-through peak of the QAE-Sephadex column was divided onto three pools: the fractions from the ascending part of the peak (pool 1, fractions 2-10, Fig. 2.4B), those corresponding to the highest point of the peak (pool 2, fractions 11-15) and those from the descending portion (pool 3, fractions 16-20). These pools were concentrated on an Amicon 50A concentrator as before and analysed along with the concentrated yeast culture media and peak II from the S200HR column by SDS-PAGE (Fig 2.5, panel A). The gel demonstrated that a protein of 45 kDa was enriched close to purity (a minor band was seen at 40 kDa) using a purification scheme of gel filtration and anion exchange resins (Fig. 2.5A, lane 5).

Immunoblot analysis of replicate samples as those used in Fig. 2.5A demonstrated that the 45 kDa protein was reactive with antibodies prepared against bacterially expressed, recombinant SmCL1 (anti-bSmCL1) (Fig. 2.5B) but did not react with pre-immunisation control serum (Fig. 2.5C). This data verified the identity of the 45 kDa protein as the recombinant SmCL1.

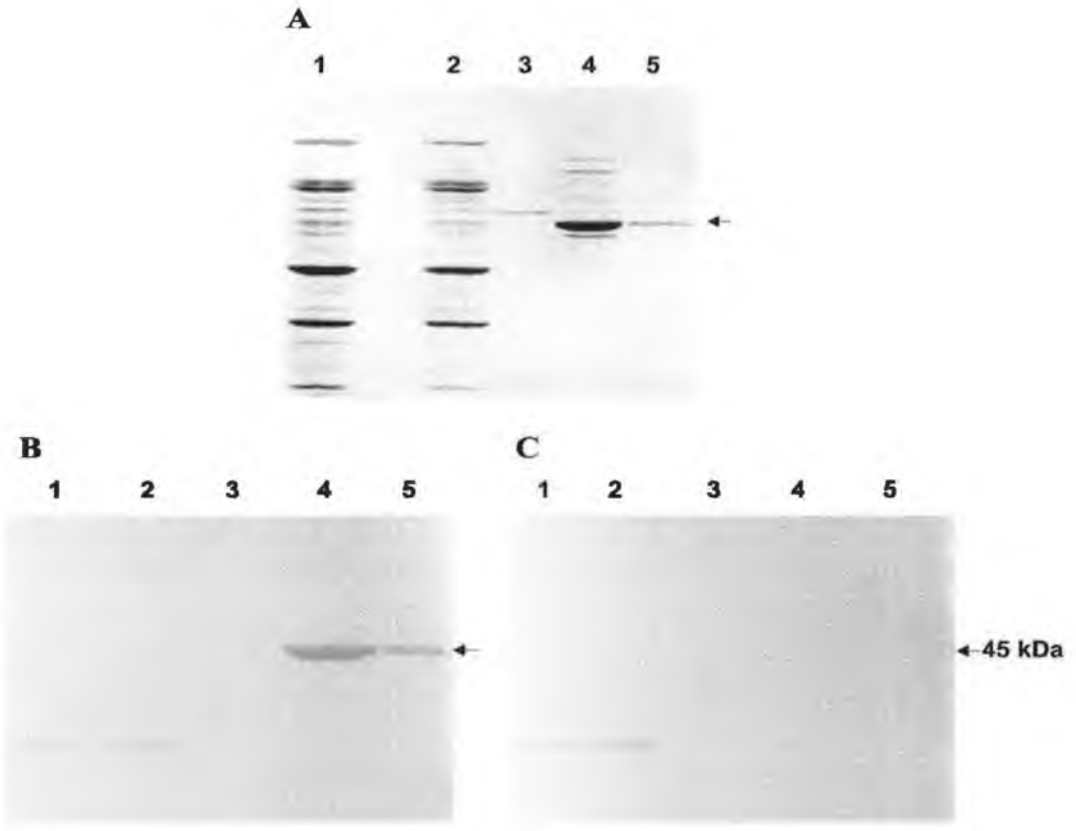


Figure 2.5 : Chromatographic purification of ySmCL1 on gel filtration and anion exchange matrices. Ten to 20 μg of protein of concentrated culture media (lane 1), S200HR peak II (lane 2), and QAE-Sephadex run-through pool I (lane 3), pool II (lane 4) and pool III (lane 5) were separated by 12% SDS-PAGE under reducing conditions. Gels were either stained in Coomassie Brilliant Blue-R (Panel A) or transferred to nitrocellulose and probed with rabbit anti-bSmCL1 serum (Panel B) or control serum (Panel C).

2.3.5 : Glycosylation of ySmCL1

SDS-PAGE analysis showed that the molecular size of ySmCL1 (45 kDa) was larger than the predicted size for the mature SmCL1 (24.1 kDa) or the proenzyme (35 kDa). Since the SmCL1 sequence revealed that this proteinase contains 3 potential N-linked glycosylation sites, the purified ySmCL1 was examined for the presence of N-linked sugar residues using an enzyme immunoassay-based, glycan detection system. The recombinant protein showed a positive reaction for the presence of glucan, as did the positive glycosylated control transferrin (Fig 2.6). In contrast, the unglycosylated control creatinase did not react with the digoxigenin probe.

2.3.6 : Zymogram analysis of ySmCL1

Purified ySmCL1 and soluble extracts of adult *S. mansoni* were analyzed for cysteine proteinase activity by gelatin-substrate PAGE. Two prominent gelatinolytic bands were observed in extracts of adult *S. mansoni* parasites, but were apparent only in the presence of the reducing agent cysteine (Fig. 2.7, Panel A,). These activities have been previously characterized in schistosome extracts to cathepsin L-like cysteine proteinase(s) (Dalton *et al.*, 1996). Proteolytic activity was observed in lanes in which ySmCL1 was applied and this activity resolved as a single band, which migrated more slowly than the two proteinases in the schistosome extracts. Like the activities in the schistosome extracts, activity of ySmCL1 was enhanced markedly by the reducing agent, cysteine (Fig. 2.7, panel B). The optimum pH for the gelatinolytic activity of ySmCL1 was determined to be pH 6.5 (data not shown).

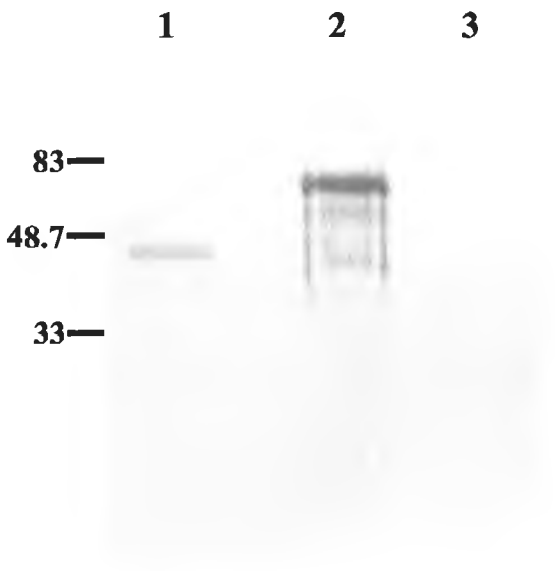


Fig 2.6 : Detection of glycoproteins using DIG-Glycan assay

ySmCL1 QAE Pool III (lane 1), positive control transferrin (lane 2) and negative control creatinase were separated by SDS-PAGE, transferred to nitrocellulose and assayed for the presence of glycan using the DIG-glycan detection kit. Molecular weight markers are indicated on left.

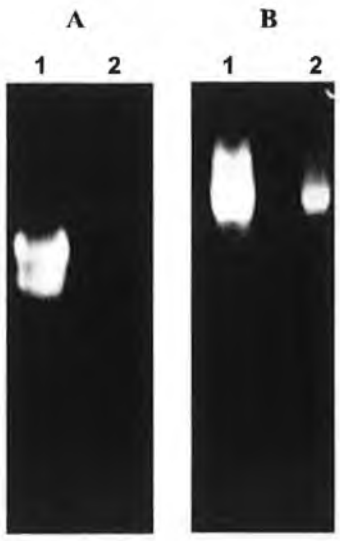


Figure 2.7 : Gelatinolytic activity of ySmCL1.

Soluble extracts of adult *S. mansoni* (Panel A) and ySmCL1 (Panel B) were analysed by 10% gelatin-substrate PAGE (zymography) at pH 6.5 in the presence (lanes 1) and absence (lanes 2) of 10 mM cysteine.

2.3.7 : Haemoglobin Proteolysis by ySmCL1

As SmCL1 has been implicated in the degradation of host haemoglobin, the ability of ySmCL1 to digest human haemoglobin was examined. Haemoglobin was incubated with ySmCL1 in buffers of varying pH in the presence of the reducing agent DTT and incubated overnight at 37°C. Reaction products were separated by SDS-PAGE and visualised by staining with Coomassie brilliant blue-R.

ySmCL1 cleaved human haemoglobin, and based on the smeared appearance of the digested products, cleaved this substrate at more than one site (Fig 2.8). In contrast to peptide substrates (see section 4.3.4) and gelatin (section 2.3.6) where ySmCL1 showed a pH optimum for activity of pH 6.5, it most efficiently cleaved haemoglobin at pH 4.5. Indeed, under the present assay conditions, haemoglobin was not digested at pH 5.0 or above.

2.3.8 : Immunolocalisation of native SmCL1 in adult worms.

Thin frozen sections of adult male and female *S. mansoni* were probed with rabbit pre-immunization and anti-bSmCL1 sera in order to determine the site of expression of SmCL1 in worm tissues. A series of 10 μm-thick longitudinal, diagonal and transverse sections were examined by light microscopy. No reactivity was observed on sections probed with pre-immunization serum (Fig. 2.9, Panel C). By contrast, immunofluorescent labeling was observed on sections probed with anti-bSmCL1 serum at a) the tegument of adult worms of both sexes, with very prominent reactivity at and immediately below the tegument (Fig. 2.9, Panel A); and b) in the gastrodermal cells lining the lumen of the schistosome gut (Fig. 2.9, Panel B). No reactivity was evident at other sites or organs in the adult worms.

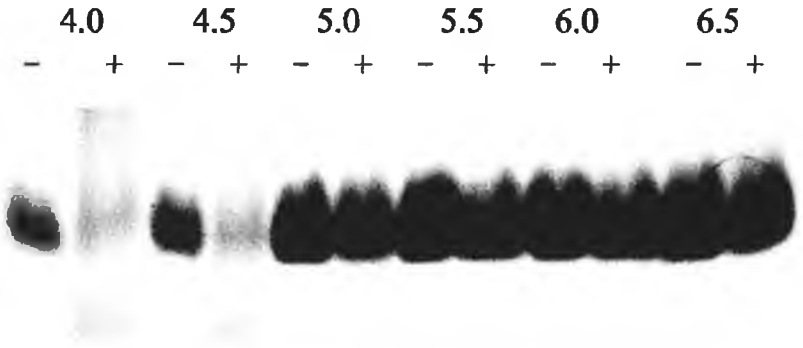
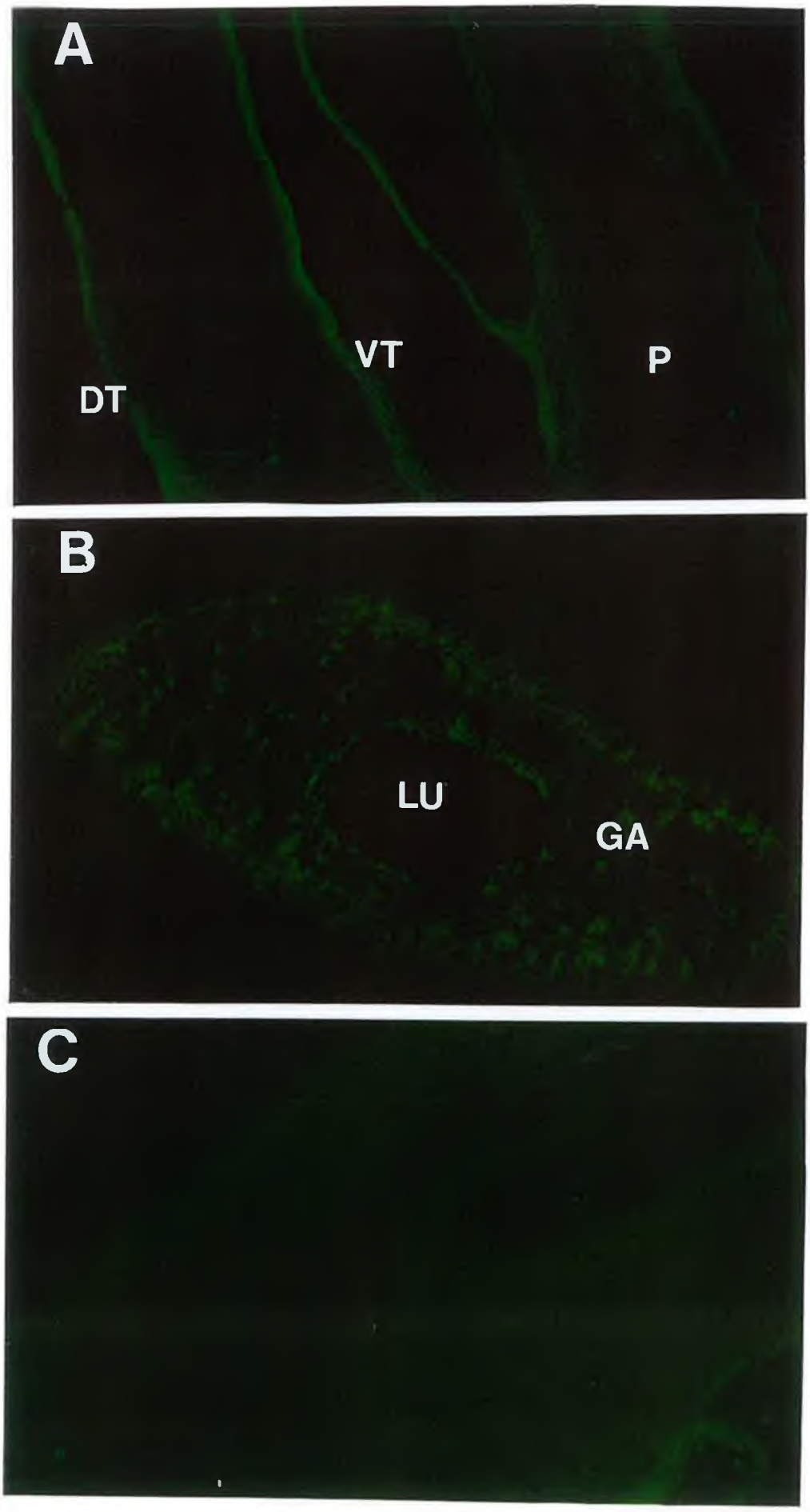


Figure 2.8 : Hydrolysis of native human haemoglobin by ySmCL1. Native human haemoglobin was incubated with (+) and without (-) purified ySmCL1. Reactions were carried out at 37°C, at different pHs (indicated above) and in the presence of 1 mM DTT. After incubation for 18 h, the reaction products were resolved by 15% SDS-PAGE under non-reducing conditions and stained with Coomassie brilliant blue-R.

Figure 2.9 : Immunolocalisation of native SmCL1 in adult schistosomes.

Longitudinal sections of male (Panel A) and transverse sections of female (Panel B) adult worms probed with rabbit anti-SmCL1 serum followed by labeling with anti-rabbit-fluorescein conjugate. Control, pre-immunisation serum did not show any specific reactivity with schistosome tissues (Panel C). Intense reaction was observed in the tegument (Panel A) and in the gastrodermal cells lining the gut (Panel B); VT = ventral tegument; DT = dorsal tegument; P = parenchyma; LU = gut lumen; GA = gastrodermis, (see overleaf).


A
VT
P
DT
B
LU
GA
C

Chapter 3

Functional expression of *Schistosoma mansoni* cathepsin L2 (SmCL2) in *Saccharomyces cerevisiae*

3.1 : Materials

Pharmacia

Quickprep mRNA Purification kit, Superose 12 gel filtration column.

Promega

1,5-bromo-4-chloro-3-indoyl-β-D-galactosidase (X-Gal), isopropylthio-β-D-galactoside (IPTG), pGEM-T vector system.

Sigma Chemical Co.

Ampicillin, l-3-carboxy-2-3-*trans*-epoxypropionyl-leucylamido(4-guanido)-butane (E-64)

Stratagene

ZAP-cDNA Synthesis kit

3.2 : Methods

3.2.1 : cDNA preparation

mRNA was isolated from mature mixed-sex adult worms using a Quickprep mRNA Purification kit (Pharmacia) according to the manufacturers instructions. In brief, total RNA was isolated from adult worms and bound to oligo dT resin. Following elution of mRNA from the resin and ethanol precipitation of the mRNA, cDNA was synthesised using a ZAP-cDNA synthesis kit (Stratagene). cDNA was purified by phenol-chloroform extraction, precipitated in ethanol and stored at -20°C.

3.2.2 : Cloning of SmCL2 cDNA

The cDNA sequence encoding pre-pro SmCL2 has been published by Michel *et al.* (1995). In this study, the full-length SmCL2 cDNA was isolated from total adult *S. mansoni* cDNA by PCR employing the primers SmCL2F and SmCL2R. These primers were designed based on the published SmCL2 sequence (GenBank accession no. Z32529) and contained *Hin*d III restriction sites (underlined).

SmCL2F : 5' CGCAAGCTTATGAAAGTTTTCTTCTTCTTTTTTCAATTATTATTT
CAGTTGCTATCGCCCAA 3'

SmCL2R : 5' TTAAAGCTTCTAAAATGTAATGAATAG 3'

Cycling conditions were as described in Chapter 2 (section 2.2.1). Reactions were analysed on 1% agarose gels and bands of the expected size of 1 kb were excised from the gel and purified as described in Chapter 2 (section 2.2.1). The purified DNA fragments were ligated to pGEM-T vector (Promega) according to the protocols provided by the manufacturer. Constructs were transformed into *E. coli* strain DH5α using the calcium chloride method (Sambrook *et al.*, 1989). LB plates containing 100 mg/ml ampicillin, 0.5 mM IPTG and 40 mg/ml X-Gal were used for blue/white selection of recombinant plasmids derived from pGEM-T. Clones were screened for the presence of inserts by PCR using λGT11 universal primers. Clones containing inserts of approximately 1 kb were sequenced as before to check for the presence of the SmCL2 cDNA.

3.2.3 : Subcloning SmCL2 cDNA into yeast expression vector

The SmCL2 cDNA sequence contains an internal *Hin*d III at nucleotide number 61 in its pro-region. As the yeast expression vector pAAH5 contains a unique *Hin*d III cloning site, it was necessary to perform partial restriction digests on the cloned cDNA in pGEM-T in order to obtain the full length prepro-SmCL2 cDNA for insertion into pAAH5. This was performed by incubating the pGEM-T construct containing the amplified SmCL2 cDNA in sequential serial dilutions of *Hind* III for 1 hour at 37°C. Digested fragments were separated on a 1.2% agarose gel and the fragment of approximately 1 kb, corresponding to the full-length prepro-SmCL2, was excised from the gel. This 1 kb DNA fragment was extracted from the agarose and the fragment cloned into pAAH5 as for SmCL1. Clones

were examined by PCR in order to evaluate which clones contained inserts with the correct orientation for expression. The PCRs were performed utilising the primers used for initial cloning (SmCL2F and SmCL2R) and primers which were synthesised based on sequences within the promoter region (termed PROM) and terminator region (named TERM) of pAAH5.

PROM : GTT CCA GAG CTG ATG AGG

TERM : CCA GGT AGA CAA GCC GAC

A clone with the correct orientation for expression was isolated and named pAAH5.SmCL2. This recombinant plasmid DNA was then used to transform *S. cerevisiae* strain DBY746 cells as described in Chapter 2 (section 2.2.2).

3.2.4 : Purification of yeast-expressed recombinant cathepsin L2 (ySmCL2).

Recombinant yeast clones were monitored for production of ySmCL2 using the peptide substrate Z-Phe-Arg-NHMec as described for ySmCL1 (Chapter 2, section 2.2.5). For purification of ySmCL2, 10 litres of culture supernatant was concentrated to 20 mls using an Amicon 2000A ultrafiltration stir-cell concentrator fitted with a YM3 (3 kDa cut-off) membrane. This concentrate was applied to a Sephacryl S300HR (Pharmacia) gel filtration column (2.6 x 60 cm), previously equilibrated in PBS at 4°C. Proteins were eluted from the matrix in PBS and fractions tested for cathepsin L activity using Z-Phe-Arg-NHMec as

substrate. Fractions containing cathepsin L activity were pooled, concentrated to 500 µl, and applied to a Superose 12 gel filtration column (Pharmacia) equilibrated in PBS. Proteins were eluted from this matrix in PBS and fractions containing cathepsin L activity pooled and stored at -20°C.

3.2.5 : Preparation of rabbit anti-SmCL2 serum

Recombinant SmCL2 was expressed in *Escherichia coli* and purified by affinity chromatography as described in Chapter 4 (section 4.2.3). The bacterially-expressed, recombinant protein (termed bSmCL2) was used to immunise a rabbit and produce a highly specific anti-bSmCL2 serum (refer to Chapter 4, section 4.2.4).

3.2.6 : SDS-PAGE and immunoblotting

Proteins were analysed by 12% SDS-PAGE and immunoblotting using anti-bSmCL2 serum as described in Chapter 2, sections 2.2.9 and 2.2.10 respectively.

3.2 : Results

3.3.1: Isolation and cloning of SmCL2 cDNA

The cDNA encoding the pre-pro-SmCL2 had previously been isolated and cloned by Michel *et al.* (1995). Using the sequence published in GenBank (accession no. Z32529), primers were designed to anneal to the 5' and 3' ends of the cDNA, and these were used to amplify pre-pro-SmCL2 by PCR from adult *S. mansoni* total cDNA. A band of the expected size, 1 kb, was amplified (Fig. 3.1, panel A). This band was excised, purified and ligated into pGEM-T. Clones were selected by blue/white screening and the presence of inserts was confirmed by PCR using universal λGT11 primers (Fig 3.1, panel B). Selected clones were subjected to DNA sequencing and the clones containing the full-length SmCL2 cDNA were identified. Sequencing also confirmed that no mutations had occurred during PCR amplification.

3.3.2 : Subcloning into pAAH5 yeast expression vector

The yeast expression vector pAAH5 contains a unique *Hin*d III cloning site for insertion of cDNAs for expression (see Fig 2.1). As the SmCL2 cDNA contains an internal *Hin*d III restriction site (Fig. 3.2A), it was necessary to perform partial restriction digests in order to clone the full-length fragment into pAAH5 plasmid DNA. When the pGem-T construct containing the cloned SmCL2 cDNA was incubated in *Hin*d III, the fully digested fragment of 920 bp was observed (Fig. 3.2B, lanes 1-3). However, as the concentration of *Hin*d III became limiting in the reaction, a minor band of approximately 1 kb, corresponding to the full-length cDNA, was detected and purified (Fig. 3.2B, lane 4). This

fragment had *Hin*d III overhangs on its 5' and 3' ends while the internal *Hin*d III site remained undigested, and this fragment was subsequently used for ligation into pAAH5.

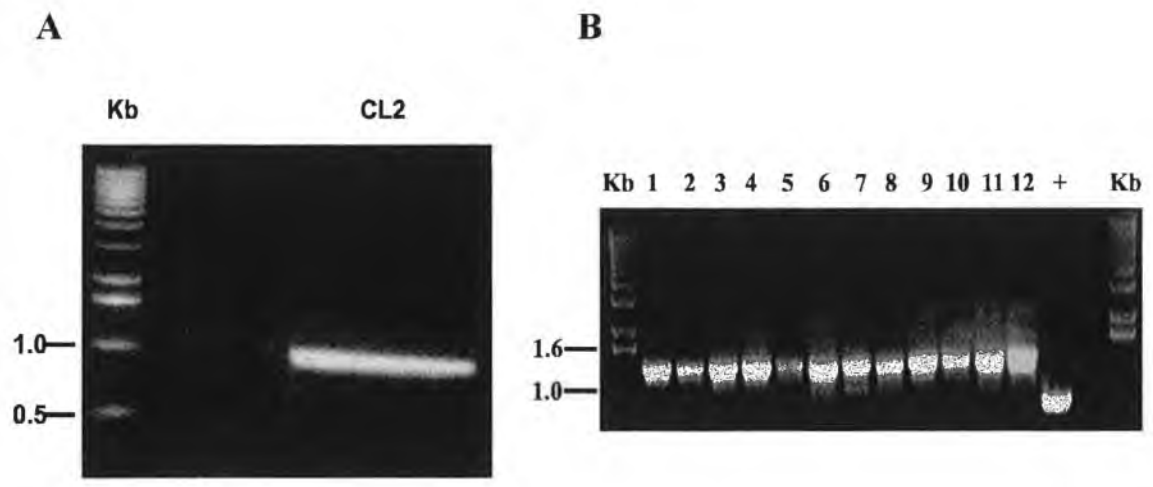


Fig. 3.1 : Isolation and cloning of SmCL2 cDNA.

Panel A; Agarose gel analysis of PCR amplification of cDNA encoding SmCL2 from adult *S. mansoni* cDNA. Panel B; cloning of PCR product into pGEM-T. Twelve selected clones were analysed for the presence of inserts by PCR using universal λGT11 forward and reverse primers. A positive control insert (+, provided with the pGEM-T cloning kit) was included and DNA size markers are indicated.

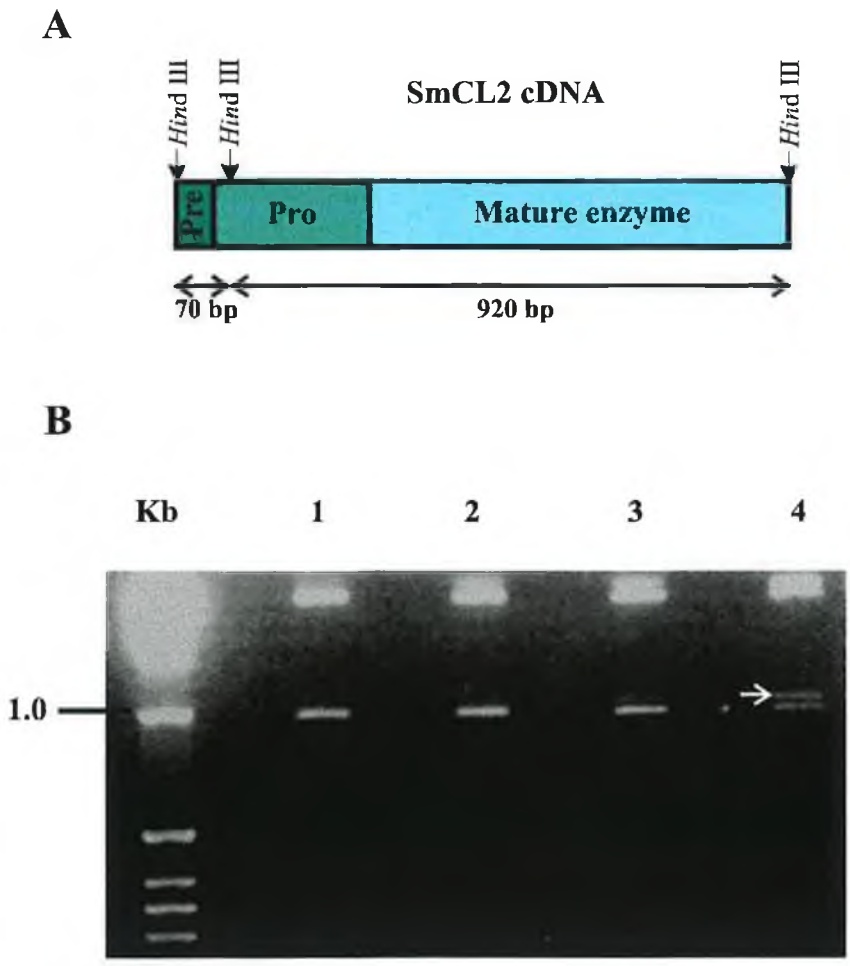


Fig. 3.2 : Partial restriction digests on SmCL2 cDNA in pGEM-T to isolate full-length fragment for insertion into pAAH5.

Panel A; Schematic diagram of SmCL2 cDNA fragment for insertion. Panel B; DNA was incubated in serial dilutions of *Hind* III (lane 1, 5 units enzyme; lane 2, 1 unit; lane 3, 0.2 units, lane 4, 0.05 units) and separated on a 1.5% agarose gel. The full-length fragment is indicated by the arrow.

The full length SmCL2 *Hin*d III-digested fragment was ligated into *Hin*d III-linearised pAAH5 and this construct was transformed into *E. coli* strain MC1061. Recombinant clones were selected on LB medium containing ampicillin and analysed for the presence of inserts by PCR employing primers based on sequences in the promotor (termed PROM) and terminator (TERM) regions of the pAAH5 vector. An example of a clone containing an insert of the expected size of 1 kb is shown in Fig. 3.3B, lane 1.

Orientation of inserts within the vector was assessed by PCR employing the SmCL2R primer used for initial cloning from cDNA and the PROM and TERM primers. A 1.3 kb fragment amplified in reactions containing a combination of the SmCL2R and PROM primers demonstrated the inserts which had ligated in the correct orientation for expression (an example is shown in Fig. 3.3, lane 2). No amplification occured in reactions employing the SmCL2R and TERM primers, confirming the correct orientation for expression in these clones (Fig. 3.3, lane 3). A clone, termed pAAH5.SmCL2, with the correct orientation for expression was selected for further experiments. The insert from this clone was sequenced and it was confirmed that the full-length SmCL2 cDNA had been cloned into the vector and this insert had indeed ligated in the correct orientation for expression.

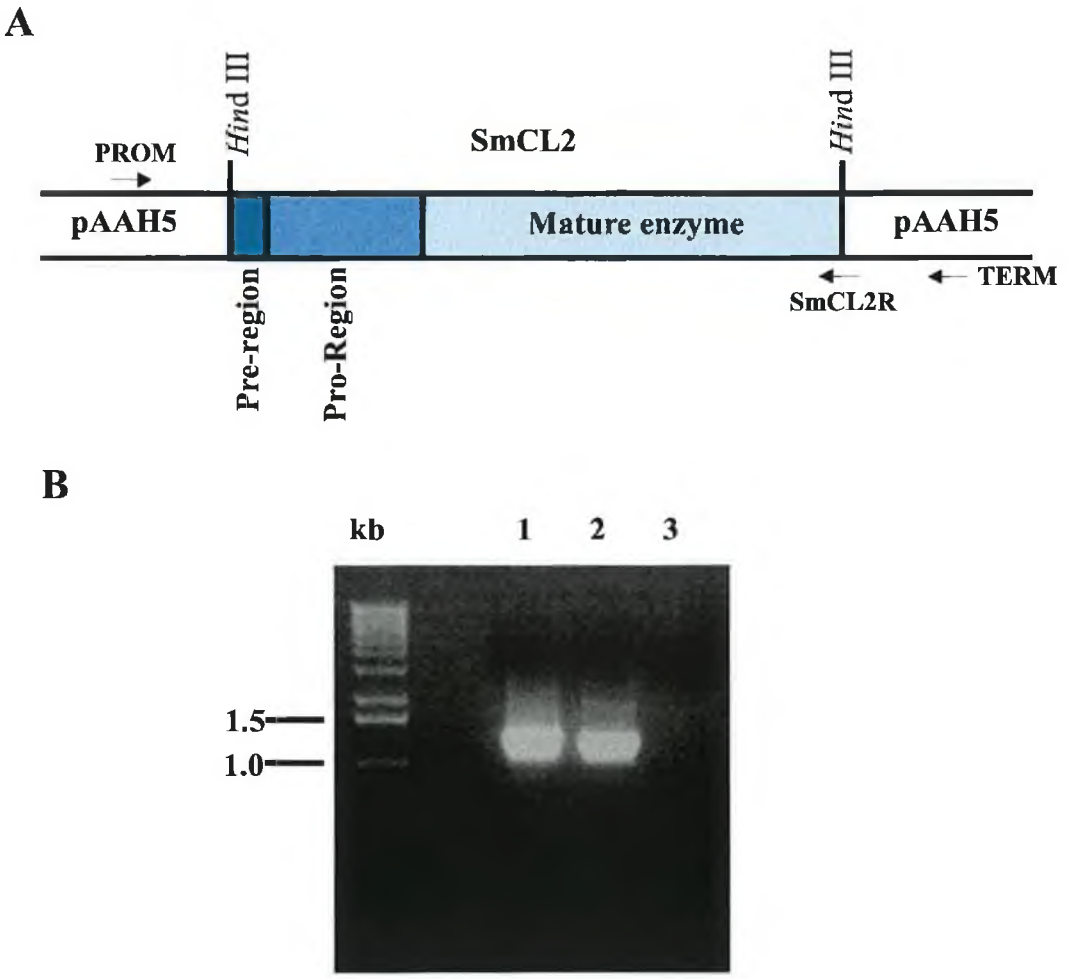


Fig. 3.3 : Cloning of SmCL2 into pAAH5 and assessing orientation of insert.

Panel A; Diagrammatic representation of construct. Panel B; PCR reactions to determine orientation of insert within pAAH5 of selected clone. Lane 1, reaction performed using PROM and TERM primers; lane 2, PROM and SmCL2R primers used; lane 3, TERM and SmCL2R primers used.

3.3.4 : Expression of SmCL2 in yeast

Yeast strain DBY746 cells were transformed with the pAAH5.SmCL2 construct or the pAAH5 vector containing no insert and recombinant clones selected on minimal medium. Five clones transformed with the pAAH5.SmCL2 construct and five transformed with pAAH5 were selected and grown in 300 mls of minimal medium in shake flasks. Cultures were grown until cells entered early stationary phase and supernatants were assayed for cathepsin L activity using Z-Phe-Arg-NHMec as substrate.

Supernatants harvested from cells transformed with the pAAH5.SmCL2 construct showed significant activity against Z-Phe-Arg-NHMec (Fig 3.4). This activity was enhanced two-fold by DTT (1mM) and was completely inhibited by the cathepsin L-specific inhibitor Z-Phe-Ala-CHN_2 (2 μM), and by the general cysteine protease inhibitor l-3-carboxy-2-3-*trans*-epoxypropionyl-leucylamido(4-guanido)-butane (E-64) (10 μM). These results are characteristic of the presence of cathepsin L-like, cysteine protease activity. Yeast transformed with pAAH5 control plasmid exhibited no activity, while *S. mansoni* crude worm extract, which contains cathepsin L activity (Dalton *et al.*, 1996) showed a similar profile to supernatants from cells transformed with pAAH5.SmCL2.

A clone displaying the highest level of cathepsin L activity was chosen and used for further studies. Expression was scaled up by subsequently culturing this clone in an automative 10 L fermentor as described in Chapter 2 (section 2.2.5). Ten litres of culture media was concentrated to 20 mls by ultrafiltration and then stored at -20°C until further required.

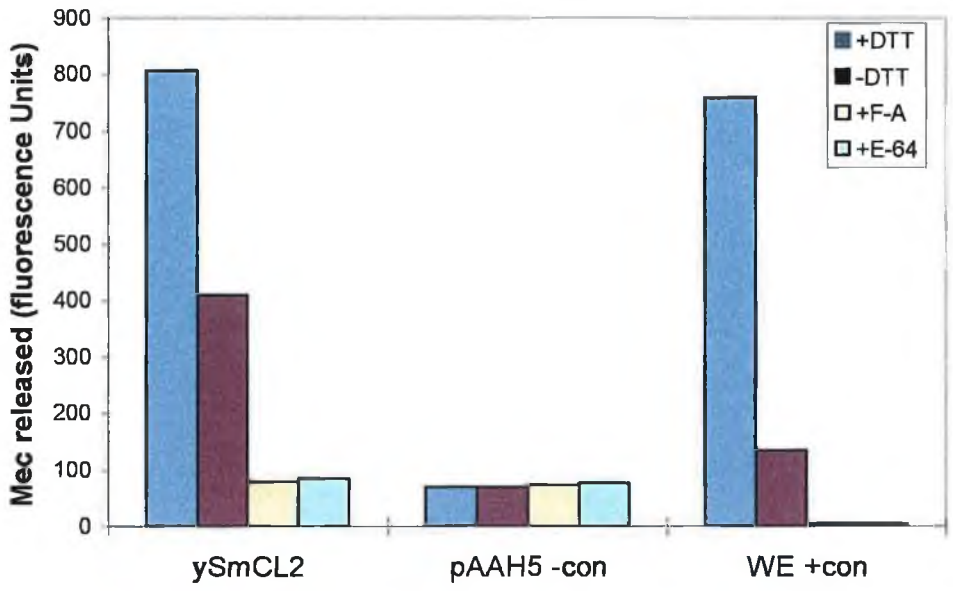


Fig 3.4 : Expression of functionally active, recombinant SmCL2 in yeast.

Supernatants of yeast cells transformed with pAAH5.SmCL2 (ySmCL2) or control plasmid pAAH5 (pAAH5 -con) were assayed using Z-Phe-Arg-NHMec as substrate in the presence or absence of the reducing agent, DTT (1mM) and the specific cathepsin L inhibitor Z-Phe-Ala-CHN_2 (F-A) (2 μM) or general cysteine protease inhibitor E-64 (10 μM) (Results are the means of experiments performed on five clones). *S. mansoni* crude worm extract (WE +con), which contains cathepsin L activity, was used as a positive control in the assay.

3.3.5 : Purification of yeast-expressed recombinant SmCL2

Yeast-expressed recombinant SmCL2 (ySmCL2) was purified from yeast culture medium using two gel-filtration columns with different separation ranges. The enzyme was first resolved on a Sephacryl S300HR column (effective separation range 10-1500 kDa). Z-Phe-Arg-NHMec-cleaving activity separated as 2 peaks (Fig. 3.5A). Peak II had a higher specific activity than peak I (Table 3.1) and this was subsequently used for further experiments.

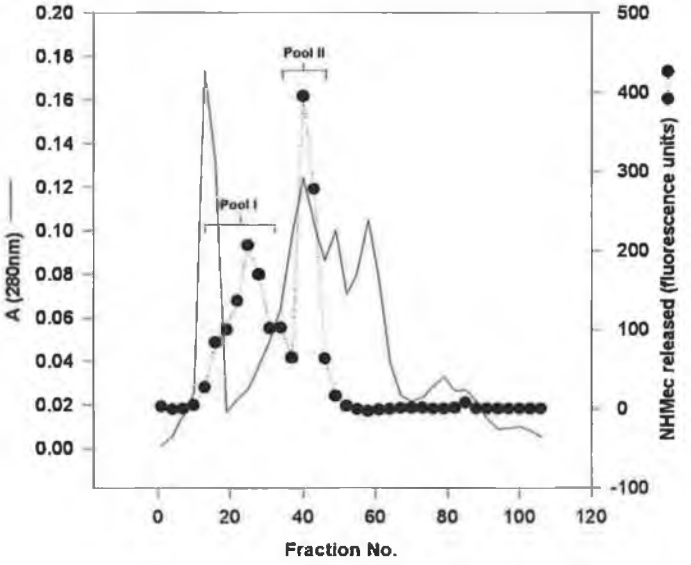
Peak II was applied to the Superose 12 column (effective separation range 1-300 kDa) and the cathepsin L activity separated as a single peak which eluted from the column after the main protein peak (Fig. 3.4B). The activity eluted from the Superose 12 was enhanced by DTT and inhibited by Z-Phe-Ala-CHN_2 (data not shown).

3.3.6 : SDS-PAGE and immunoblotting

Pooled fractions from the S300 HR and the Superose 12 gel filtration columns were analysed by 12% SDS-PAGE and immunoblotting using anti-bSmCL2. Gels stained with Coomassie Blue-R revealed a protein, of approx. 32 kDa in size, had been purified close to homogeneity (Fig 3.6A). A protein of approximately 18-20 kDa was also evident.

The 32 kDa protein reacted strongly on immunoblots with anti-bSmCL2 serum, and the protein was seen to be present in concentrated supernatant and the S300 HR pool (Fig. 3.6B). The 18-20 kDa protein did not react with the anti-bSmCL2. A 32 kDa protein was also reactive with anti-bSmCL2 in *S. mansoni* crude worm extract. No bands were detected on replicate filters probed with pre-immune control serum (not shown).

A



B

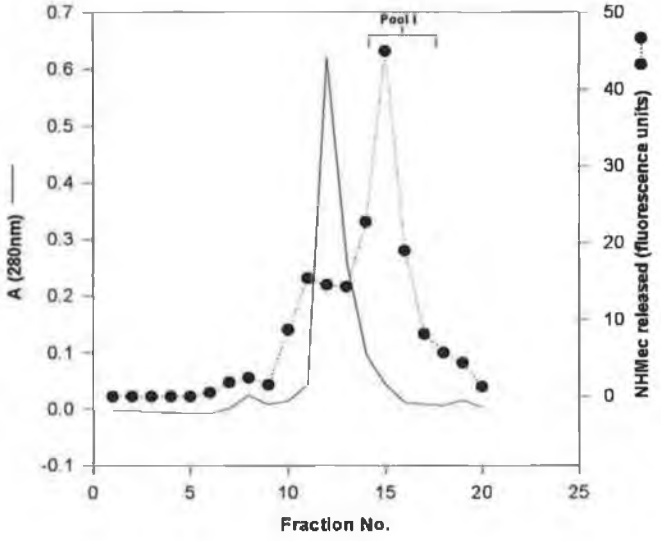


Fig 3.5 : Purification Profiles of ySmCL2. Profiles of elution from a Sephacryl S300HR column (Panel A) and Superose 12 column (Panel B) monitored by absorbance at 280nm (lines) and activity against Z-Phe-Arg-NHMec (dots). Pooled fractions are indicated.

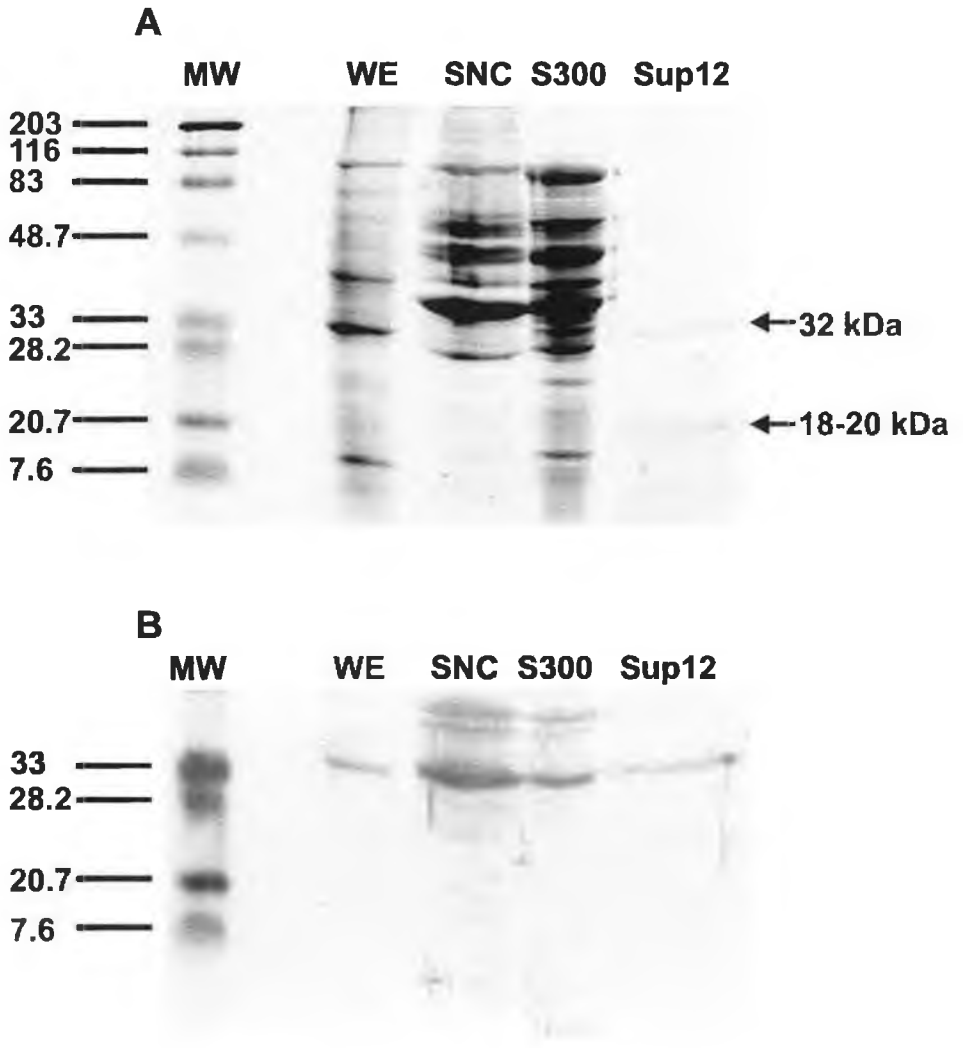


Fig 3.6 : Purification of ySmCL2 on gel filtration matrices

S. mansoni crude worm extract (WE), crude concentrated supernatant (SNC) from yeast culture expressing ySmCL2, S300 HR pool II (S300) and Superose 12 pool I (Sup12) were separated by SDS-PAGE and stained with Coomassie Blue-R (Panel A) or transferred to nitrocellulose and probed with anti-bSmCL2 serum (Panel B). Molecular weight standards (MW) are illustrated in kDa.

Table 3.1: Purification of ySmCL2

	Protein (mg)	Activity (U)	Sp. activity (U/mg)	Yield (%)	Purification fold
Culture medium	270.3	1.35	0.005	100	1
S300 Pool I	10.2	0.51	0.05	37.78	10
S300 Pool II	2.7	0.79	0.29	58.52	58.0
Superose 12	0.06	0.075	0.8	5.56	160.0

Chapter 4

Comparative biochemical analysis of recombinant ySmCL1 and ySmCL2 and investigation into expression of the native enzymes in schistosomes

4.1 : Materials

Advanced Biotechnologies

Total RNA Isolation Reagent

Qiagen

Ni-NTA Superflow nickel chelate resin, pQE.30 *E. coli* expression vector, QIAquick-spin columns, RGS-4H monoclonal antibody.

Sigma Chemical Company

Guanidine hydrochloride, imidazole, kanamycin, urea.

Superfos Biosector

Quil A/Saponin adjuvant.

4.2 : Methods

4.2.1 : Cloning into *E. coli* expression vector

The cDNAs encoding the mature portion of SmCL1 and SmCL2 were cloned into the bacterial expression vector pQE.30 (Qiagen) as recommended by the manufacturers. The pQE.30 vector contains an affinity tag consisting of six histidine residues (6xHis tag) which is placed at the carboxy-terminus of the recombinant protein. Recombinant proteins can therefore be purified in one step on nickel chelate (Ni-NTA) resin, utilising the histidines affinity for nickel ions (Janknecht *et al.*, 1991). Also, a motif consisting of Arg, Gly, and Ser (RGS) is added to the N-terminus of the 6xHis tag. A monoclonal antibody raised to the RGS-4His epitope (Qiagen) can be used to screen for the presence of the recombinant protein on immunoblots. Recombinant constructs (termed SmCL1.pQE.30 and SmCL2.pQE.30) were transformed into *E. coli* strain M15 and transformed clones selected on LB plates containing 100 μg/ml ampicillin and 25 μg/ml kanamycin.

4.2.2 : Expression of recombinant proteins in *E. coli*

LB medium containing 100 μg/ml ampicillin and 25 μg/ml kanamycin was innoculated with cells transformed with SmCL1.pQE.30 or SmCL2.pQE.30 and incubated at 37°C with shaking until the optical density (OD_{600}) reached 0.8. Expression of recombinant proteins was induced by addition of IPTG to a final concentration of 1 mM, and the cells were harvested 5 h later by centrifugation (this time point was found to give optimal expression levels of recombinant proteins). The solubility and cellular location of

recombinant bSmCL1 and bSmCL2 were determined according to the manufacturers instructions. Cell pellets were stored at -20°C until further required.

4.2.3 : Purification of recombinant proteins

The pellets of SmCL1.pQE.30- or SmCL2.pQE.30-transformed cells were resuspended in 0.1 M sodium phosphate-0.01 M Tris-HCl (pH 8.0) containing 6 M guanidine hydrochloride at 5 ml per gram of cell pellet and sonicated for 8 min (duty cycle 25%, output 2.5) (Branson Sonifier 250, Branson Ultrasonics). The extracts were then centrifuged at 14,000 g for 30 min and the supernatants decanted. The supernatants were then incubated with 2 ml of Ni-NTA Superflow resin (Qiagen) for 1 h at room temperature. The resin was packed into a column and washed with 5 volumes of 0.1 M sodium phosphate-0.01 M Tris-HCl (pH 7.2) containing 8M urea. Recombinant proteins were eluted with a linear gradient of imidazole (0-250 mM imidazole prepared in washing buffer) at 0.5 ml/min. One ml fractions were collected and analysed by 12% SDS-PAGE. Immunoblot analysis was also performed using the RGS-4his monoclonal antibody (Qiagen) engineered onto the C-terminus of the recombinant protein expressed in pQE30.

4.2.4 : Anti-serum production

Purified proteins (termed bSmCL1 and bSmCL2) were dialysed overnight in PBS at room temperature and used as antigens to raise an anti-bSmCL1 and anti-bSmCL2 sera in New Zealand White rabbits. The rabbits were immunised five times with 20 µg of bSmCL1 or bSmCL2 suspended in QuilA adjuvant (Superfos Biosector) with intervals of 3 weeks

between boosts. Bleeds were taken from rabbits at regular intervals (every 2-3 weeks) and used on immunoblots to check for the production of antibodies against the recombinant proteins.

4.2.5 : Characterisation of ySmCL1 and ySmCL2 proteinase activity.

Substrate specificity and kinetics studies were carried out on purified ySmCL1 and ySmCL2 using a bank of fluorogenic peptide substrates. The kinetic constants, k_{cat} and K_m were obtained by non-linear regression analysis using the Enzfitter program (Leatherbarrow, 1987). Active site titration using the cysteine proteinase inhibitor l-3-carboxy-2-3-*trans*-epoxypropionyl-leucylamido(4-guanido)-butane (E-64) and the fluorogenic substrate Z-Phe-Arg-NHMec was performed to determine the molar concentration of active ySmCL1 and ySmCL2 (Barrett *et al.*, 1982).

For determination of the optimum pH of proteinase activity, the following buffers were used at a concentration of 50 mM; glycine, pH 2.5 to 3.0 and 9.1 to 10.0, sodium acetate, pH 3.5 to 5.5, sodium phosphate, pH 5.5 to 7.5 and Tris-HCl, pH 7.5 to 9.0. The ionic strength of each buffer was equalised to 100 mM using NaCl. Assays were conducted using Z-Phe-Arg-NHMec as substrate as described earlier (Chapter 2, section 2.2.6).

4.2.6 : Reverse transcription (RT)-PCR analysis

Eggs and miracidiae were isolated according to the method of Dalton *et al.* (1997b) and cercariae isolated as outlined by Dalton *et al.* (1997a). Total RNA was prepared from adult worms, cercariae, eggs and miracidiae using Total RNA Isolation Reagent

(Advanced Biotechnologies) as per maufacturers instructions. RNA was purified by phenol/chloroform extraction followed by isopropanol precipitation and ethanol washes as described by the manufacturers. First strand cDNA was synthesised from 1 μg of total RNA using gene specific primers (termed L1RTR for SmCL1, L2RTR for SmCL2, and TPIR for triosephosphate isomerase (TPI) expression control, see below) with Superscript II reverse transcriptase (Gibco BRL) using the conditions recommended by the manufacturer. RNA was removed by hydrolysis before PCR by adding NaOH was to a final concentration of 0.4 M to each reaction and incubating the mix at 65°C for 30 min. Acetic acid was then added to a final concentration of 0.4 M to neutralise the reaction mix. cDNA was purified on QIAquick-spin columns (Qiagen) and PCR reactions were carried out using 2 μl of purified first strand cDNA. Thirty cycles were performed, each composed of a denaturation step for 1 min at 94°C, primer annealing for 1 min at 50°C and extension for 1 min at 72°C. Reaction products were separated on a 1.2% agarose gel.

Primers used in RT-PCR reactions were as follows :

For SmCL1 : L1RTF : CAA GGA ATG TGT GGC TCT TGT

L1RTR : GGA TGA CTA ATT CCA TGT TG

For SmCL2 : L2RTF : ATA TGG GAA TGA TGG ATG TCA

L2RTR : CTA AAA TGT AAT GAA TAG

For TPI : TPIF : GTT GGG GGG AAC TGG AAA ATG AA

TPIR : CAA AGG GTG CAT TCA CCG GAG AA

4.2.7 : Immunoblot analysis of soluble extracts and ES products

Soluble extracts of female and male adult worms and cercariae were prepared as described in Chapter 2 (section 2.2.7). 10 μg of total protein from each extract was separated on 12% SDS-PAGE gels and transferred to nitrocellulose. Filters were probed with either anti-bSmCL1 or anti-bSmCL2 and developed using anti-rabbit IgG-alkaline phosphatase conjugate as similarly described in chapter 2 (section 2.2.10).

4.3 : Results

4.3.1: Expression and purification of recombinant cathepsin Ls from *E. coli*

The mature forms of SmCL1 and SmCL2 were expressed in *E. coli*, utilising the expression vector pQE.30. This vector contains a galactose-inducable promotor which allows high cell densities to be reached before expression of recombinant proteins is induced. After induction of expression, samples of cultures were taken at regular time intervals and cells analysed for the presence of an induced protein. In the case of SmCL2, a protein of approximately 32 kDa in size was seen to be expressed in cells one hour after induction with IPTG (Fig. 4.1). The intensity of this band was also seen to increase over time.

The protein solubility and cellular location of the bSmCL1 and bSmCL2 was determined and neither protein was found in the soluble portion of the cytoplasm or to be secreted into the periplasmic space of the bacterial cell (data not shown). Rather, the recombinant proteins were expressed in insoluble inclusion bodies in the cytoplasm. Proteins were therefore solubilised in guanidine.HCl and purified under denaturing conditions.

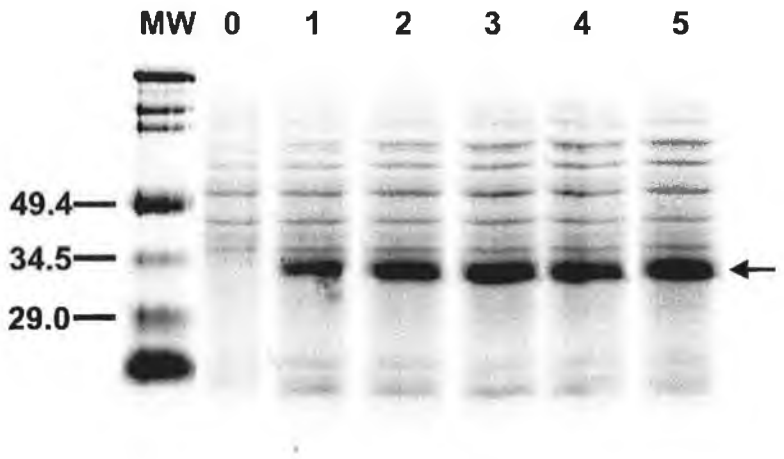


Fig. 4.1 : Induction of expression of recombinant bSmCL2 in *E. coli* cells

E. coli cells harbouring the expression plasmid were induced to express bSmCL2 by adding IPTG and samples were taken at 0, 1, 2, 3, 4 and 5 hours after induction. Cell pellets were boiled in 100 µl of SDS-PAGE reducing sample buffer, separated on 12% gels and stained with Coomassie brilliant blue-R. The arrow indicates the protein of interest and pre-stained molecular weight markers (MW) are indicated in kDa.

Affinity chromatography on nickel chelate resin was employed as the recombinant proteins contained a hexa-his tag at their C-terminus. A linear gradient of imidazol was used to elute bound proteins from the resin and fractions containing protein (monitored by OD_{280}) were analysed by SDS-PAGE. For bSmCL2, a protein of approx. 32 kDa was seen to elute from the resin and this protein was >95% pure (Fig. 4.2A). This protein eluted between 75 and 100 mM imidazole. A similar profile was observed for elution of recombinant SmCL1, except its molecular size was approximately 24 kDa (data not shown).

Both the 24 kDa bSmCL1 and 32 kDa bSmCL2 were analysed on immunoblots and the proteins reacted strongly with the monoclonal antibody to the polyhistidine ligand (Fig. 4.2B), verifying that the affinity purified molecules were indeed the recombinant proteins.

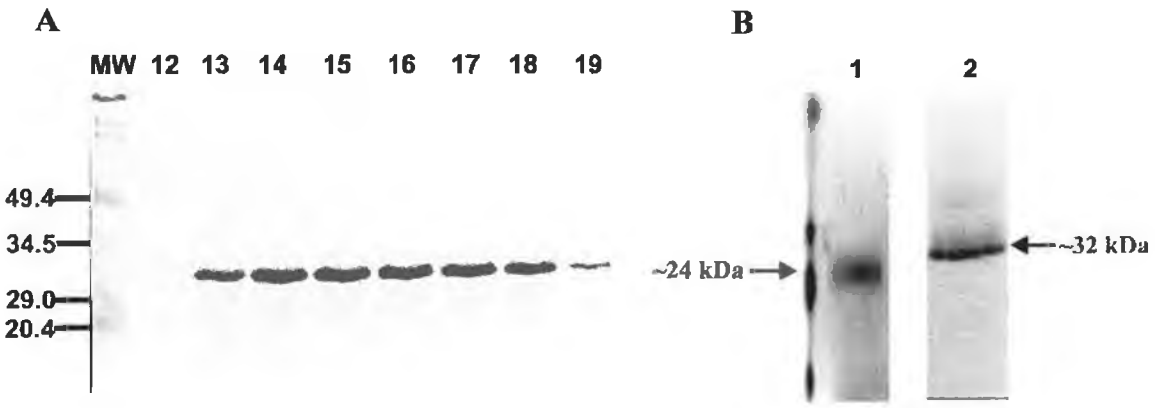


Fig. 4.2 : Purification of recombinant cathepsin Ls from *E. coli* cells by affinity chromatography

Panel A; Profile of elution of bacterially-expressed SmCL2 from Ni-NTA resin. Fractions eluted from the resin (fraction numbers indicated) were analysed by SDS-PAGE to determine the presence and purity of recombinant protein. Panel B; Immunoblot analysis of purified, bacterially-expressed SmCL1 (Panel 1) and SmCL2 (Panel 2) using the monoclonal antibody to the polyhistidine ligand.

4.3.2 : Anti-serum preparation

Recombinant, bacterially-expressed SmCL1 (bSmCL1) and SmCL2 (bSmCL2) were used as the antigens to prepare mono-specific rabbit anti-sera. Rabbits were given boosts at 2-3 week intervals and the production of antibodies against the recombinant antigens analysed by immunoblotting. Both in the case of bSmCL1 and bSmCL2, the production of antibodies was not detected until after the third boost. However, strong signals were detected, both against native and recombinant antigens, about 10-14 days after the fourth boost (Fig 4.3).

The anti-bSmCL1 serum was seen to react strongly with the 24 kDa bSmCL1 protein (Fig. 4.3A, lane 3). It also reacted with a protein of approximately 33 kDa in *S. mansoni* adult worm extract (Fig. 4.3A, lane 1) and purified, 45 kDa yeast expressed SmCL1 (ySmCL1) (Fig. 4.3A, lane 2).

The anti-bSmCL2 serum reacted with a the 32 kDa bSmCL2 (Fig. 4.3B, lane 3). It also reacted with a protein in crude extracts of adult worms (Fig. 4.3B, lane 1) and purified ySmCL2 (Fig. 4.3B, lane 2), and both these proteins also had an apparent molecular weight of 32 kDa.

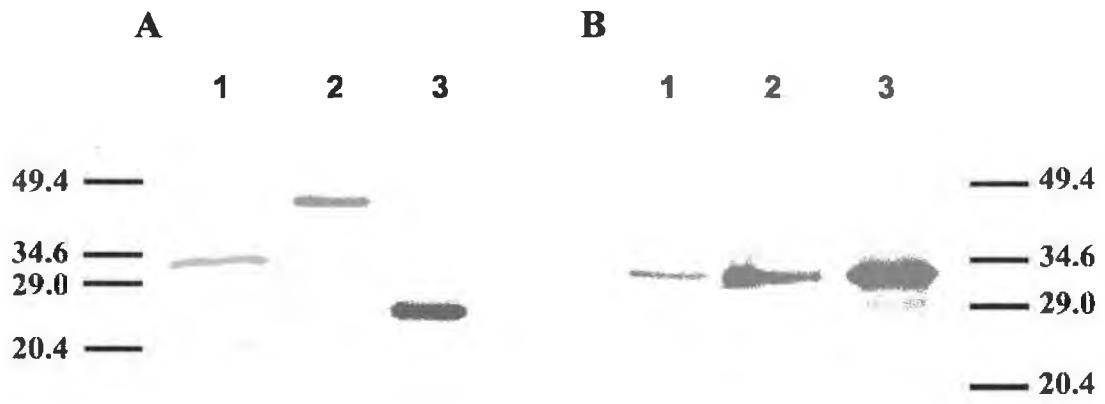


Fig 4.3 : Production of anti-serum against recombinant SmCL1 and SmCL2 antigens.

Anti-serum raised against bSmCL1 or bSmCL2 was used to probe recombinant and *S. mansoni* expressed native antigens of SmCL1 (Panel A) or SmCL2 (Panel B). Lane 1; crude *S. mansoni* adult worm extract. Lane 2; purified ySmCL1 (Panel A) or ySmCL2 (Panel B) and lane 3; purified bSmCL1 (Panel A) or bSmCL2 (Panel B). Molecular weight standards are indicated in kDa.

4.3.3 : Substrate specificity of ySmCL1 and ySmCL2.

The substrate specificity of ySmCL1 and ySmCL2 was characterised using a panel of 14 fluorogenic peptide substrates. Kinetic constants were obtained for peptide substrates that are classically used for the characterisation of cathepsin proteinases, namely, Z-Phe-Arg-NHMec and Z-Arg-Arg-NHMec (Table 4.1 and Barrett and Kirschke, 1981). Both SmCL1 and SmCL2 had a high affinity (k_{cat}/K_m) for Z-Phe-Arg-NHMec, SmCL2 having approximately ten times higher affinity for this substrate than SmCL1. Both enzymes have a low affinity for Z-Arg-Arg-NHMec, with SmCL2 exhibiting negligible cleavage. The preference for a hydrophobic residue in the P_2 position is consistent with the classification of SmCL1 and SmCL2 as cathepsin Ls (Barrett and Kirschke, 1981).

Both enzymes also have a high affinity for Boc-Val-Leu-Lys-NHMec, with SmCL1 having three-times higher affinity for this substrate over Z-Phe-Arg-NHMec in contrast to SmCL2 which has 3 times less affinity for Boc-Val-Leu-Lys-NHMec over Z-Phe-Arg-NHMec. It was found that SmCL2 does not cleave any other of the substrates examined while SmCL1 has a broader range of specificity. In particular, SmCL1 has a high affinity for H-Leu-Val-Tyr-NHMec (k_{cat}/K_m= 51.89 mM^{-1} s^{-1}) and Suc-Leu-Tyr-NHMec (k_{cat}/K_m= 44.12 mM^{-1} s^{-1}). Also, SmCL1 exhibited low affinity for Tos-Gly-Pro-Arg-NHMec and Boc-Phe-Val-Arg-NHMec.

Table 4.1 : A comparison of the kinetic constants of ySmCL1 and ySmCL2 on synthetic fluorogenic peptide substrates. '*' indicates substrate cleavage of insufficient magnitude to allow calculation of enzyme kinetic results.

	Cathepsin L1			Cathepsin L2		
Substrate	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
Z-Phe-Arg-NHMEC	11.81	0.424	35.90	3.2	1.31	406.3
Boc-Val-Leu-Lys-NHMEC	2.73	0.319	116.8	5.6	0.82	147.5
H-Leu-Val-Tyr-NHMEC	2.1	0.107	51.89	*	*	*
Suc-Leu-Tyr-NHMEC	4.9	0.214	44.12	*	*	*
Tos-Gly-Pro-Arg-NHMEC	6.8	0.016	2.38	*	*	*
Boc-Phe-Val-Arg-NHMEC	9.8	0.02	1.32	*	*	*
Z-Arg-Arg-NHMEC	3.27	0.004	1.10	*	*	*
Z-Phe-Ser-Arg-NHMEC	7.84	$5.0x10^{-5}$	0.0064	*	*	*
Tos-Gly-Lys-Arg-NHMEC	*	*	*	*	*	*
Tos-Gly-Pro-Lys-NHMEC	*	*	*	*	*	*
Tos-Leu-Gly-Arg-NHMEC	*	*	*	*	*	*
Tos-Val-Pro-Arg-NHMEC	*	*	*	*	*	*
Suc-Leu-Leu-Val-Tyr-NHMEC	*	*	*	*	*	*

4.3.4 : pH activity profile of ySmCL1 and ySmCL2

ySmCL1 exhibited a broad bell-shaped activity profile against Z-Phe-Arg-NHMec (Fig. 4.4). Its pH optimum for activity is approximately 6.5, but the enzyme retained over 50% of its activity as low as pH 4.5 and as high as pH 8.0. By contrast, ySmCL2 showed a much narrower range of activity against Z-Phe-Arg-NHMec. Its pH optimum is 5.5, but the enzyme was virtually inactive above pH 7.0. However, it did show a greater tolerance for acidic conditions compared to ySmCL1, retaining 50% of its activity at pH 3.2.

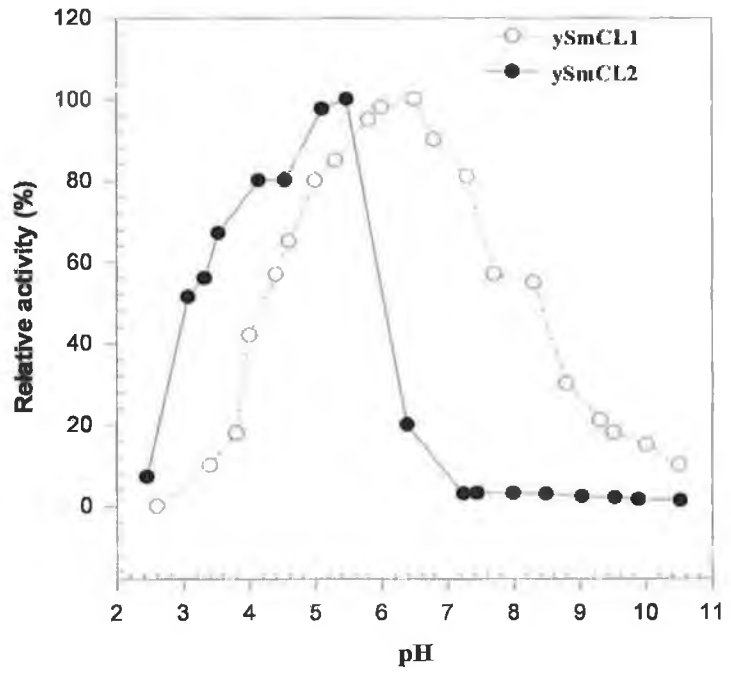


Fig. 4.4 : pH profile of activity for recombinant ySmCL1 and ySmCL2.

The activity of the recombinant proteases against Z-Phe-Arg-NHMec was measured at different pHs. Results are the means of triplicate experiments and are plotted as relative activity.

4.3.5 : RT-PCR analysis of life-cycle stages

RT-PCR analysis of cDNA isolated from various life-cycle stages of *S. mansoni* was performed to investigate the expression of SmCL1 and SmCL2. These studies revealed that both enzymes were expressed by both males and females of the adult stage (Fig. 4.5). SmCL1 was expressed by cerceriae (Fig. 4.5, Panel A, lane 6) but not by miracidiae or eggs (Fig. 4.5, Panel A, lane 4 and 5) while SmCL2 was not expressed by any of the larval stages (Fig. 4.5, Panel B, lanes 4-6).

TPI, a constitutively-expressed housekeeping enzyme (Hooker and Brindley, 1996), is expressed by all stages (Fig.4.5, Panel C). No bands are observed in reactions containing the TPI specific primers but no reverse transcriptase (Fig. 4.5, Panel D), confirming that no contaminating genomic DNA was present. Also, no bands were present in reactions where no cDNA was added, confirming no unspecific reactions had occured (Fig. 4.5, Panels A, B and C, lane 8).

4.3.6 : Immunoblot analysis of soluble extracts and ES products

Protein extracts from various stages of the schistosome life cycle were analysed by immunoblotting using anti-bSmCL1 or anti-bSmCL2 sera as probes. In the case of SmCL1, a protein was observed at 33 kDa in male and female adults (Fig 4.6, panel A, lanes 1 and 2) which seemed to correspond to the mature enzyme. This protein was present at approximately double the amount in females as opposed to males. A band at 43 kDa was also observed in males, females and cercariae (Fig. 4.6, panel A, lane 3) which may represent the pro-SmCL1. Various intermediate bands were also present which could

be breakdown products of the pro-enzyme or differentially glycosylated forms of the enzyme. A strongly staining band at 33 kDa was also observed in adult worm ES products, indicating that SmCL1 was secreted into the culture medium.

SmCL2 was also observed in male and female adult schistosomes (Fig. 4.6, panel B, lanes 1 and 2), although the mature form of 32 kDa was present in females at approximately five times the amount as that observed in males. An apparent pro-enzyme at 39 kDa was also observed in males and females. SmCL2 was also present in ES, but in comparatively lower amounts than SmCL1 (Fig. 4.6, panel B, lane 4).

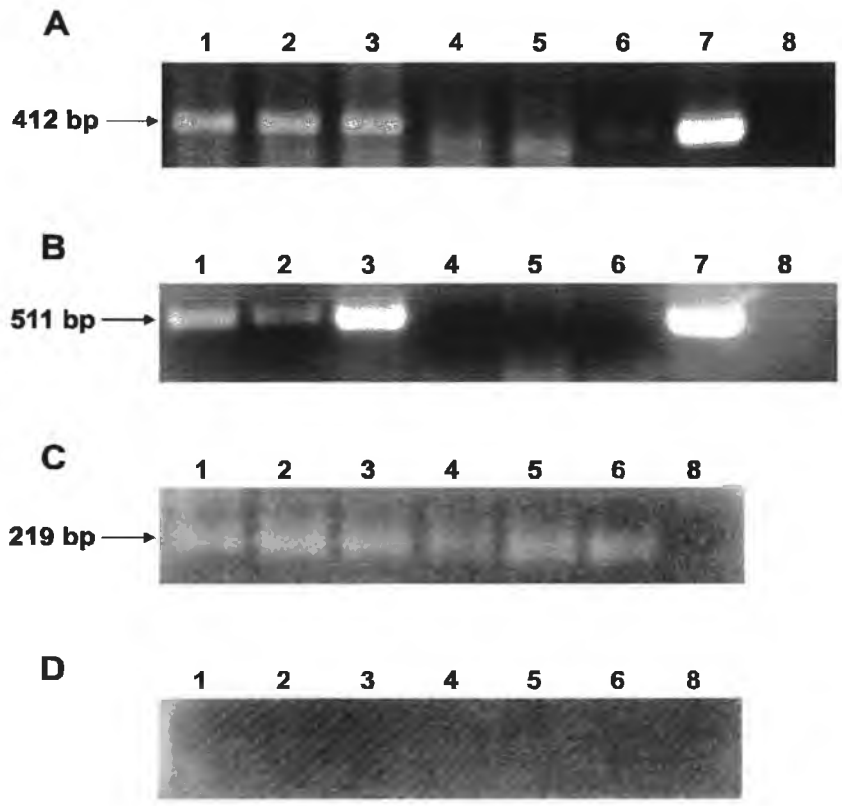


Fig. 4.5 : RT-PCR analysis of SmCL1 and SmCL2 expression in life-cycle stages of *S. mansoni*.

Mixed-sex adults (lane 1), male adults (lane 2), female adults (lane 3), eggs (lane 4), miracidiae (lane 5) and cercariae (lane 6) were analysed for the presence of mRNA encoding for SmCL1 (Panel A) or SmCL2 (Panel B) by RT-PCR. SmCL1 (Panel A, lane 7) and SmCL2 (Panel B, lane 7) cDNAs were used as positive template controls. Control reactions with no added template DNA were also included (lane 8, all Panels). RT-PCRs using primers that anneal to cDNA of the constitutively-produced triosephosphate isomerase (TPI) were performed as a positive control (Panel C) while Panel D designates the control PCR reaction carried out using a blank reverse transcriptase to check for contaminating genomic DNA.

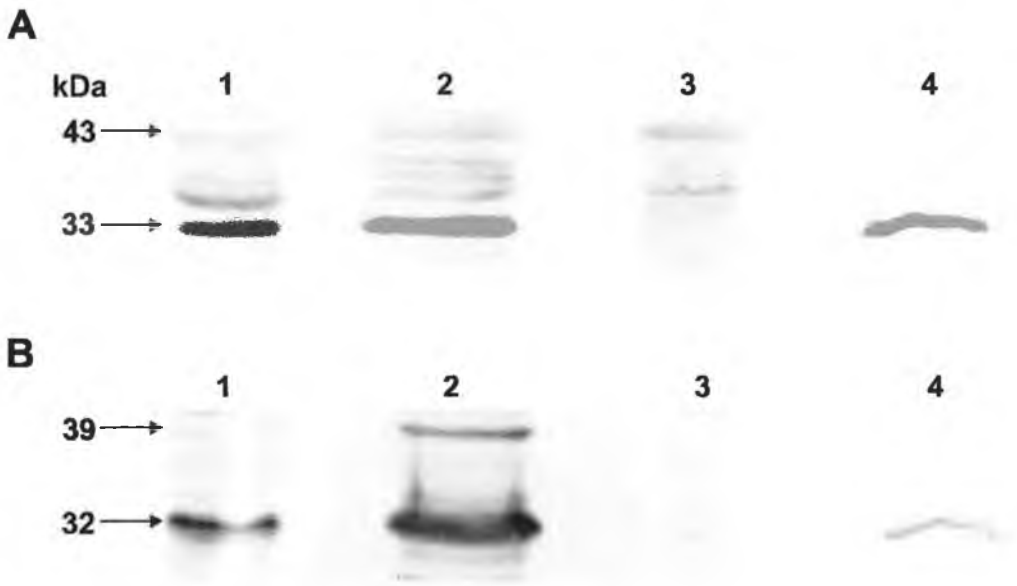


Fig. 4.6 : Immunoblot analysis of soluble extracts and ES products of *S. mansoni* for the presence of native SmCL1 and SmCL2.

Soluble protein extracts of male worms (lanes 1), female worms (lane 2), cercariae (lane 3), and ES products (lane 4) were transferred to nitrocellulose filters and probed with rabbit anti-bSmCL1 (Panel A) or rabbit anti-bSmCL2 (Panel B). Replicate filters probed with control pre-immune sera did not exhibit any reactivity (not shown).

Chapter 5

Discussion

5.1 : Introduction

Proteinases can play many roles in host-parasite interactions. It is clear from examination of the life cycle of helminth parasites that proteinases perform key functions in their development. Indeed, it can be argued that the evolution and diversification of proteinases has contributed to the success of parasitic helminths. The main selection pressure that initiated the parasite-host relationship may have been the ease by which nutrients could be acquired from the host by the parasite (Halton, 1987) and hence it is found that much of the predominant proteolytic activity in helminth parasites is involved in this function. However, the development of mechanisms to facilitate the parasites migration through the host tissue, to defend against host immunological attack, and to otherwise ensure the completion of their life cycle would obviously have presented other selection pressures. It is not unlikely that the diversification and specialisation of proteinases has played a major role in the development of these processes.

Schistosomes acquire their amino acid nutrients by degrading host haemoglobin and several proteinases have been implicated to be involved in this process. The dominance of cathepsin L activity in extracts of schistosome tissues and in ES products of adult worms (Dalton *et al.*, 1996a) suggest that these enzymes (along with other proteinases) may be centrally involved in this mechanism. Indeed, inhibitors of cysteine proteinases were shown to prevent haemoglobin digestion by schistosomula and decrease their viability *in vitro* (Wasilewski *et al.*, 1996; Zerda *et al.*, 1988). Moreover, treatment of *S. mansoni*-infected mice with these inhibitors not only reduced worm burden but delivered anti-fecundity affects, i.e. reduced egg output by females (Wasilewski *et al.*, 1996). Cysteine proteinases, including cathepsin L-like proteinases, are therefore

considered important targets to which novel anti-schistosome chemotherapy and/or immunoprophylaxis could be directed.

Elucidation of the precise physiological role of proteinases of schistosomes has been hampered by the difficulty in obtaining homogenous enzymes. It is not practical to isolate the enzymes directly from schistosomes given the difficulty in obtaining large numbers of the parasites. Furthermore, many of the enzymes have similar molecular sizes and substrate specificity's (for example, the cathepsin L1, cathepsin L2, cathepsin B and cathepsin C proteinases). In order to study the biochemical properties of these proteinases of *S. mansoni* it is therefore necessary to develop means by which the recombinant molecules can be expressed and purified in a functional form.

Several cysteine proteinases have been expressed in bacterial systems. However proteins overexpressed in *E. coli* frequently form insoluble aggregates or inclusion bodies in the bacterial cytoplasm (Kopitar et al, 1996; Smith and Gottesman, 1989; Taylor *et al.*, 1992; Kuhelj *et al.*, 1995; Valesco et al; 1994; Dolinar et al; 1995). This is due to the absence of mechanisms for post-translational modification of the recombinant enzyme, resulting in an improperly folded product. Bacterial cells also lack secretion mechanisms present in eukaryotic cells causing build up of recombinant proteins within the cell cytoplasm. Recombinant cysteine proteinases expressed in bacteria aggregated in inclusion bodies and could only be solubilised by strong denaturants in the presence of a reducing agent (Kopitar et al, 1996; Smith and Gottesman, 1989; Taylor *et al.*, 1992; Kuhelj *et al.*, 1995; Valesco et al; 1994; Dolinar et al; 1995). Solubilised proteins were then refolded and the propeptide cleaved from the proenzyme by either auto-activation or addition of a second proteinase (e.g. pepsin or subtilisin) to yield the mature, active zymogen.

In order to circumvent the problems of solubility and improper folding experienced with bacterial expression systems, various mammalian and insect cell systems have been employed to express recombinant cysteine proteinases (Bromme *et al.*, 1996; Gotz and Klinkert, 1993; Kane, 1993; Tao *et al.*, 1994). While these systems expressed the recombinant proteinases in an active form, yields were generally much lower than those obtained in the bacterial systems. Also, these systems are much more complex, costly and laborious.

Yeast expression systems are a popular alternative to bacterial and mammalian systems for the production of functionally active heterologous proteinases. As a unicellular microorganism, yeast retains the advantages of bacterial systems in their ease of manipulation and growth. On the other hand, yeast cells exhibit many features of cell structures normally associated with higher organisms and a eukaryotic subcellular organisation capable of accurate post-translational processing and modification of complex recombinant proteins. Heterologous proteins can be directed into the secretory pathway in yeast which can be advantageous as it avoids the build-up of potentially toxic proteins in the cell. Moreover, the secretion of the protein offers purification advantages in that extraction of the recombinant protein from cells is not necessary and there are much fewer contaminating proteins present in the culture medium compared to the cell extract. Secretion is of particular relevance for the production of recombinant cysteine proteinases as disulphide bond formation occurs during passage through the endoplasmic reticulum and Golgi body (Schekman and Novic, 1982).

For these reasons, the bakers yeast *Saccharomyces cerevisiae,* strain DBY746, was used in this study to express SmCL1 and SmCL2. This expression system has been

successfully used to produce substantial quantities of each of the two forms of cathepsin L from the related parasitic trematode *Fasciola hepatica* (Dowd *et al.*, 1997; Roche *et al.*, 1997). Most importantly, in these cases the recombinant *F. hepatica* cathepsin Ls were secreted into the yeast culture medium in an active form, avoiding the need for activation steps.

5.2: Expression of SmCL1 and SmCL2 in *S. cerevisiae*

Functional expression of recombinant proteins in yeast requires the recognition of trafficking signals within the cDNA (Moir and Davidow, 1991). In mammalian cells, cathepsin L proteinases are synthesised as preproenzymes that are sequentially processed to the mature molecules during their passage through the endoplasmic reticulum, Golgi complexes and lysosomes or secretory vesicles.

Proteins which enter the secretory pathway in eukaryotic cells have an amino-terminal signal peptide composed of 10-15 hydrophobic amino acid residues (Garoff, 1985; Kelly, 1985). Proteins containing these signal peptides, or pre-regions, are synthesised on ribosomes that are situated on the surface of the rough ER. The hydrophobic signal peptide is thought to embed into the lipid bilayer of the ER membrane while the rest of the protein is being synthesised (Rapoport, 1985). When the protein is fully synthesised, cleavage of the signal peptide by a specific signal peptidase releases the protein into the lumen of the ER, where the protein folds into its 3-dimensional conformation (Lodish, 1988). Most of the proteins containing this hydrophobic signal sequence will pass from the ER to the Golgi apparatus via transport vesicles.

Proteins destined for secretion are packaged into secretory vesicles, which are formed by budding from the trans-Golgi network. The sorting signal which directs proteins into these vesicles is largely unknown, but it is thought to reside in the polypeptide chain of the protein (Griffiths and Simons, 1986). However, it has been suggested that motifs within the pro-peptide and also at the C-terminus of proteinases are involved in signalling proteinases into the secretory pathway. Mutations within the propeptide and deletion of C-terminal sequences of proPC2, a subtilisin-like serine proteinase, prevented its secretion from transfected mammalian cells (Taylor *et al.*, 1997). It has also been suggested that proteins secreted from eukaryotic cells are transported by a non-selective "default pathway"; i.e. any protein in the ER will automatically be carried through the Golgi apparatus to the cell surface unless it is selected for transport elsewhere (Rothman, 1987).

Many lysosomal proteins are synthesised with N-terminal extensions, or propeptides, that are removed during or soon after delivery to lysosomes. In the case of mammalian cathepsin L and other cathepsins, one function of the propeptide is to regulate the enzymatic activity of the protein, rendering the protein inactive when it is present (Mason *et al.*, 1989). The proregion of mammalian cathepsin L is also required for proper folding, stability and exit from the ER (Tao *et al.*, 1994).

cDNAs encoding two distinct cathepsin L proteinases from *S. mansoni* and their homologues in *S. japonicum* have been isolated (Smith *et al.*, 1994; Michel *et al.*, 1995; Day *et al.*, 1995). Examination of the primary amino acid sequence of these cathepsin Ls revealed that these enzymes are synthesised as prepro-enzymes and would seem to possess

structures similar to those observed for the mammalian enzymes (Smith *et al.*, 1994, Michel *et al.*, 1995, Day *et al.*, 1995; Dalton *et al.*, 1996a).

The cathepsin L proteinases from the parasitic trematode *F. hepatica* were localised to secretory vesicles of the gut epithelial cells, indicating that these enzymes may be directed into the secretory pathway in these cells (Smith *et al.*, 1993; Dowd *et al.*, 1994). Moreover, these molecules were also shown to be the major proteolytic activity present in culture medium in which these worms were incubated. Cathepsin L like activity has been also shown to be present in culture medium in which adult *S. mansoni* worms were incubated (Dalton *et al.*, 1996a), suggesting that one or both of the *S. mansoni* cathepsin Ls may also enter the secretory pathway in schistosome cells.

In this study, yeast cells were transformed with an expression plasmid pAAH5 containing the cDNAs encoding the full-length pre-procathepsin SmCL1 or SmCL2. Analysis of the culture supernatants harvested from cells transformed with SmCL1 or SmCL2 indicated that cathepsin L-like activity was secreted into the culture medium in both cases. This suggests that both proteinases are trafficked through the normal secretory pathway in yeast and that the preproenzymes contain all the information necessary for processing and secretion. Similar results were observed for the functional expression of *Fasciola hepatica* pre-procathepsin Ls expressed in the same yeast system as that used to express SmCL1 and SmCL2 (Dowd *et al.*, 1997; Roche *et al.*, 1997). Indeed, expression of *F. hepatica* cathepsin L1 in secretion-defective yeast mutant strains verified that this enzyme was secreted through the normal secretory pathway and not through a non-specific pathway or by leakage from the cell (Roche *et al.*, 1997).

There have been several previous reports of the expression of cysteine proteinases in *S. cerevisiae*, including mammalian cathepsin B (Rowan *et al.*, 1992), cathepsin S (Bromme *et al.*, 1993), papain (Vernet *et al.*, 1993) and *S. mansoni* cathepsin B (Lipps *et al.*, 1996). In all these cases the enzymes have been expressed as fusion proteins with the yeast α-factor pre- or pre-pro signals. It is noteworthy that these recombinant proteinases were secreted from the yeast as inactive proenzymes and have required a subsequent activation step to obtain functionally active enzyme (Rowan *et al.*, 1992; Bromme *et al.*, 1993; Vernet *et al.*, 1993; Lipps *et al.*, 1996). In the case of *F. hepatica* cathepsin Ls expressed in *S. cerevisiae*, the recombinant proteinases which were purified from the yeast culture medium were identical in molecular weight to the native enzymes, indicating that the recombinant proteinases were processed in the yeast cells in a similar fashion to the native proteinases (Dowd *et al.*, 1997; Roche *et al.*, 1997).

Studies on the processing of the pro-enzyme to the mature form of mammalian cathepsins indicate that an endoproteinase(s) is involved in the initial cleavage of the pro-peptide and that the final processing steps are performed by exopeptidases that clip amino acids from the N-terminus of the molecule as far as the molecular structure of the mature enzyme will allow (Rowan *et al.*, 1992; Mach *et al.*, 1993; Isidoh and Kominami, 1994). These cleavages are thought to begin in the trans-Golgi network and they continue in the secretory vesicles.

Asparaginyl residues were shown to occur near the cleavage point between the pro-peptide and mature enzymes of cathepsin L1, cathepsin L2, cathepsin B, cathepsin D and cathepsin C of schistosomes (Dalton and Brindley, 1996, 1997). This observation led these authors to propose that a novel cysteine proteinase, an asparaginyl endopeptidase

termed schistosome legumain (also known as Sm32) was the endoproteolytic activity involved in the cleavage of the pro-peptide from the mature protein of these proteinases. Asparaginyl residues are also present in the vicinity of the cleavage point between the pro-region and mature protein of *F. hepatica* cathepsin L1 and L2, and this parasite also expresses an asparaginyl endopeptidase (Dowd and Dalton, unpublished data). These observations suggest that trematodes possess a common mechanism , involving asparaginyl endopeptidases, for the processing of cathepsin proteinases. In addition, since mammalian cathepsins do not contain these asparagine residues close to the cleavage point of their pro-regions, these enzymes likely undergo different processing to the trematode cathepsins.

A recent entry into the public databases (Genbank accession number U32517) revealed that *S. cerevisiae* possesses a putative asparaginyl endopeptidase. If this enzyme is involved in the maturation of secreted proteinases, it may explain why ySmCL1 and ySmCL2 are processed to their active form in yeast. It is noteworthy that *S. mansoni* cathepsin B expressed in *S. cerevisiae* was secreted as an inactive pro-form, unlike ySmCL1 and ySmCL2 (Lipps *et al.*, 1996). However, SmCL1 and SmCL2 were expressed utilising their native pre-region as a signal sequence, whereas the cathepsin B was expressed as a fusion protein, where the pro-cathepsin B was fused to yeast α-factor secretion signal. It is possible that the cathepsin B fusion protein was processed in a different fashion to prepro-ySmCL1 and prepro-ySmCL2. The α-factor signal may have targeted the protein differently to its native pre-region and the protein may not have been exposed to the action of the asparaginyl endopeptidase.

Both in the case of ySmCL1 and ySmCL2, yields of recombinant protein secreted into the culture medium were quite low in comparison to yields obtained for the expression of other recombinant cathepsins. There are a number of reasons which may explain why such poor yields were obtained. Firstly, the pAAH5 expression vector utilises the constitutive, glycolytic alcohol dehydrogenase (ADH 1) promoter of *S. cerevisiae*. Glycolytic promoters are most active at the start of fermentation, when the sugar concentration is highest. This can result in the production a recombinant protein which is toxic to the cell which can hinder cell growth and produce a selective pressure for a reduction in expression. Secondly, both ySmCL1 and ySmCL2 were processed to their active form in yeast, and it is possible that in this form, the enzymes may have been toxic to the yeast cell. Thirdly, as ySmCL1 was glycosylated by the yeast cells (discussed in section 5.3), glycosylation of the protein may have been a rate limiting step in the secretion of the protein, thereby reducing levels of expression. Finally, as pAAH5 is an autonomously replicating plasmid, it was necessary to grow the transformed yeast in selective medium. Selection pressure must be retained throughout growth otherwise the yeast cells can lose their plasmids. However, growth in selective medium resulted in low levels of biomass accumulation and correspondingly low levels of expression.

Attempts by Lipps *et al.* (1996) to express cathepsin B in *S. cerevisiae* using the constitutive α-factor promoter also resulted in low yields of recombinant protein (approximately 20 μg of recombinant protein per liter of culture). However, in this same report, high levels of expression (up to 10 mg/liter of culture) were achieved when the enzyme was expressed the under the control of a galactose-inducible promoter. The use of inducible promoters allows high cell numbers to be achieved before expression of the

recombinant protein is induced, and any deleterious effects of the recombinant protein on cell growth are avoided. The recombinant cathepsin B was also produced and secreted in its inactive pro-form, which is unlikely to be as toxic to the yeast cell as the mature, active product. Furthermore, the cloned cDNA encoding the cathepsin B was mutated to delete a potential glycosylation site. Expression of SmCL1 and SmCL2 using a similar system to that described for cathepsin B may increase yields of these proteinases significantly.

Alternatively, systems other than *S. cerevisiae* for expression of SmCL1 and SmCL2 could be investigated. In recent times, a number of non-*Saccharomyces* yeast strains have emerged as hosts for recombinant protein production. Human cathepsin L has been expressed in the methylotropic yeast *Pichia pastoris* and yields of 20 mg of recombinant enzyme per litre of culture were obtained, which were sufficient for crystallisation studies to be performed on this molecule (Coulombe *et al.*, 1996a, 1996b). Furthermore, *P. pastoris* expression systems allow the expression cassette to be integrated into the yeast genome, thereby avoiding stability problems associated with autonomously replicating plasmids and allowing growth in complete, non-selective medium. This allows much higher biomass levels to be obtained compared to growth in selective medium, with correspondingly higher levels of expression of the recombinant protein.

Various bacterial expression systems have been used to express cathepsin proteinases as described earlier, and although the proteins have been expressed in denatured form in insoluble inclusion bodies, many of the enzymes have been solubilised and refolded to yield active product. While only a small percentage of the denatured protein is recovered in active form (e.g. Smith and Gottesman (1989) reported less than 0.1% of the total recombinant human cathepsin L produced in *E. coli* was recovered

actively), the expression levels in bacterial systems are so high that the yield of renatured active *proteinase* is usually comparable with that obtained in yeast and mammalian systems.

5.3: Purification of recombinant cathepsin Ls

The recombinant ySmCL1 and ySmCL2 were purified from the yeast culture medium using conventional chromatography techniques. ySmCL1 was purified by gel-filtration followed by anion exchange chromatography. Z-Phe-Arg-NHMec-cleaving activity eluted from the gel-filtration column as two peaks, the second peak being used for subsequent purification as it contained proteinase with higher specific activity. Similar elution profiles from gel filtration were observed for the purification of recombinant *Fasciola hepatica* cathepsin Ls expressed in yeast (Dowd *et al.*, 1997; Roche *et al.*, 1997). It is possible that the elution of the enzymes in two peaks may be due to aggregation of some of the recombinant enzyme to yeast proteins.

ySmCL1 was further purified on an anion exchange column. The major proportion of the proteolytic activity did not bind to the resin at pH 7.0, but was collected in the run-through fractions. This was unexpected as SmCL1, with a predicted isoelectric point (pI) of 5.6, would be expected to have an overall negative charge at pH 7.0 and hence bind to an anion exchanger. It is possible, however, that the true pI of SmCL1 is higher than the value predicted from the primary sequence as pI values of native proteins are determined by the charges of the amino acid residues that are exposed on the outer surface of the molecule. Nevertheless, while the ySmCL1 did not bind to the resin, the majority of the contaminating yeast proteins did bind and hence the ySmCL1 was further

purified. In fact, it was probably advantageous that ySmCL1 did not bind to the anion exchanger as elution conditions from ion exchange columns can decrease enzyme activity, especially if proteins bind tightly and harsh elution conditions are required to remove them from the resin.

Peaks containing cathepsin L activity were analysed by SDS-PAGE and immunoblotting using anti serum prepared against bacterially-expressed SmCL1 (anti-bSmCL1 serum). A protein was observed to be enriched as a result of the purification and this protein was highly reactive with anti-bSmCL1 serum, verifying its identity as the recombinant SmCL1. This protein migrated at approximately 45 kDa on gels, which is larger than the predicted size for the mature SmCL1 (24.1) or the proenzyme (35 kDa). Examination of the primary amino acid sequence of SmCL1 reveals that this proteinase contains three potential N-linked glycosylation sites in its mature region. Probing of the purified ySmCL1 for the presence of N-linked sugar residues revealed that the enzyme was indeed glycosylated and this may explain why it migrates higher than the expected size on gels. Immunoblot analysis of crude schistosome extracts indicated the presence of a 33 kDa protein and a 43 kDa protein, seemingly representing the mature-SmCL1 and pro-SmCL1 respectively (Fig. 4.6). The molecular sizes observed for the native enzyme would suggest it too may be glycosylated.

N-linked glycosylation occurs in the lumen of the ER in mammalian cells. The transfer of a core oligosaccharide (composed of N-acetyl-glucosamine, mannose and glucose and containing a total of 14 sugar residues) is transferred to target asparagine residues in a single enzymatic step almost as soon as the residue emerges in the ER lumen during protein synthesis on the ribosome (Hirschberg and Snider, 1987). Glycoproteins are

transferred from the ER via transfer vesicles to the Golgi body, where the oligosaccharide chains are 'trimmed' in the lumen of the Golgi. The purpose of N-linked glycosylation of glycoproteins, however, remains unknown. It was originally thought that N-linked oligosaccharides functioned in transport of glycoproteins through the ER and Golgi, although this now does not seem to be the case (Stanley, 1987). Indeed, most proteins retain their normal activities in the absence of glycosylation (Stanley, 1987). However, it is known that acid hydrolases destined for trafficking to lysosomes carry mannose 6-phosphate (M6P) groups which are added exclusively to the N-linked oligosaccharides of these enzymes in the Golgi (Pfeffer, 1988). Since all glycoproteins arrive in the Golgi with identical N-linked oligosaccharide chains, the signal for adding M6P units to oligosaccharides resides somewhere in the polypeptide chain of each protein (Lang *et al.*, 1984). The nature of this recognition sequence is still not known.

Yeast cells tend to hyperglycosylate glycoproteins in comparison to mammalian cells. The average length of oligosaccharide chains of glycoproteins from *S. cerevisiae* is 40 mannose residues, in comparison to 8-14 residues added by mammalian cells (Tarentino *et al.*, 1974). This would correlate with the molecular sizes observed for the ySmCL1 (45 kDa) and the native enzyme (33 kDa) expressed in schistosome tissues and may explain why the former molecule migrates higher than its predicted size on SDS-PAGE gels.

Modification of the N-linked glycosylation motif -Asn-X-Ser- (where X is any amino acid except proline) in the protein using site directed mutagenesis results in absence of glycosylation. This approach was adopted for expression of rat cathepsin B (Hasnain *et al.*, 1993), human cathepsin L (Kane, 1993) and *S. mansoni* cathepsin B (Lipps *et al.*,

1996) as it has been postulated that hyperglycosylation by yeast may result in an inactive enzyme. These proteinases were expressed and secreted as pro-enzymes, whereas ySmCL1 was secreted seemingly in its activated, mature form. Neither of the *F. hepatica* cathepsin Ls contain glycosylation motifs in their primary sequences (Roche *et al.*, 1997) and both of these recombinant proteinases were secreted as mature enzymes using the same expression system employed for expression of SmCL1. It is not clear how glycosylation (or the absence of glycosylation) affects the processing, targeting and functionality of these enzymes. However, hyperglycosylation did not seem to adversely affect the activity and secretion of ySmCL1.

ySmCL2 was purified from the culture medium using two gel-filtration columns with different separation ranges. When total culture supernatant was applied to the first gel filtration column, ySmCL2 resolved as two peaks similar to the profile observed for ySmCL1. Again, it is suggested to be due to a portion of the recombinant enzyme aggregating to yeast proteins. As described for ySmCL1, the second peak of activity was used for subsequent purification as it contained the higher specific activity. It was attempted to further purify ySmCL2 on ion exchange resins, however, it was found that no cathepsin L activity was recovered from either anion or cation exchange resins. The reason for this is unclear. Nonetheless, a second gel filtration column was employed instead of an ion exchanger and enzyme activity was eluted after the main protein peak on this column.

Analysis of the purified ySmCL2 by SDS-PAGE revealed that a protein of approximately 32 kDa had been purified close to homogeneity in fractions containing cathepsin L activity. This protein was also reactive with antiserum prepared against

bacterially-expressed SmCL2 (anti-bSmCL2 serum) on immunoblots, verifying its identity as the recombinant proteinase. This immunoreactive 32 kDa protein was also apparent in the unpurified crude concentrated supernatant from the fermentation and in the pool from the first gel filtration (S300 HR) column. Two additional proteins of approximately 38-40 kDa were also observed to react with the serum in both these lanes. The larger protein of approximately 40 kDa would seem to correspond to the pro-ySmCL2, whereas the protein at approximately 38 kDa may represent a semi-processed form of the pro-enzyme. Neither of these proteins were observed in the pool collected from the second gel-filtration (Superose 12) column, indicating that the mature enzyme was separated from the pro-enzyme in this step.

A protein of 32 kDa in size also reacted with anti-bSmCL2 in *S. mansoni* crude worm extract, although the predicted molecular mass of the mature SmCL2 is 24.3 kDa. However a protein of 32 kDa in size was also observed by Michel *et al.* (1995) to react with anti-SmCL2 serum in schistosome extracts, and they concluded that this protein represented mature SmCL2. They postulated that the protein may have an anomalous behaviour on SDS-polyacrylamide gels or that the proteinase could be glycosylated, as the proposed amino-acid sequence of the mature protein indicates the presence of one potential N-linked glycosylation site. However, expression of the mature SmCL2 in *E. coli*, which lacks the machinery for glycosylation of proteins, also resulted in a expression of a protein of approximately 32 kDa (Figs 4.1 and 4.2). This suggests that both the native enzyme and the yeast-expressed recombinant protein are not modified by glycosylation but do indeed run anomalously on SDS-PAGE gels. While SmCL2 does possess a single potential N-linked glycosylation site, examination of the deduced amino acid sequence of

SmCL2 shows that this site lies at Asn-204, in close proximity to the S_2 subsite Ala-205, and would, therefore, be unlikely to be exposed to the glycosylation machinery of the cell (Dalton *et al.*, 1996a).

While both ySmCL1 and ySmCL2 were purified close to homogeneity, substantial losses of enzyme were incurred during the purification of both proteinases. As explained earlier, a large proportion of each enzyme was lost due to the fact that both proteinases seperated as two peaks on the first gel filtration column. Losses of activity were also incurred in concentration steps performed between the stages of purification. It is likely that higher yields of enzyme would have been achieved if a one step, affinity-based purification scheme had been employed instead of gel filtration and ion exchange chromatography. A hexa-histidine affinity tag was engineered onto the carboxy-terminus of *S. mansoni* cathepsin B expressed in *S. cerevisiae*, and this enzyme was purified to homogeneity by elution from nickel chelte resin (Lipps *et al.*, 1996). The hexa-His tag did not seem to adversely affect the activity of the enzyme. One step purification of human cathepsin D was achieved by affinity chromatography using immobilised propeptide sequences (Wittlin *et al.*, 1998). The propeptides of cathepsin proteinases are known to bind their respective mature enzymes highly specifically and Wittlin *et al.* (1998) suggested that a similar strategy as that employed for the purification of cathepsin D could also be utilised for the simple, one-step purification of other families of proteinases.

5.4: Biochemical characterisation of ySmCL1 and ySmCL2

The S_1 subsite in the active site of cysteine proteinases is involved in the interaction with the P_1 site of the substrate, the formation of the transition states and the catalytic cleavage of the peptide bond. Interactions between the amino acids of the enzyme S_2 subsite and the residue in the P_2 position of the substrate determine the specificity of cysteine proteinases (Barrett and Kirschke, 1981). While analysis of the amino acid sequences in the mature region of SmCL1 and SmCL2 reveals that the two proteinases have only 43.8% identity (Dalton *et al.*, 1996a), the residues which constitute the active site of the enzymes are very similar. The S_1 subsite is highly conserved in cysteine proteinases, and this site is virtually identical between SmCL1 and SmCL2 (Fig. 5.1). The residues constituting the S_2 subsite of SmCL1 and SmCL2 are also very similar (Fig. 5.1) and the hydrophobic character of the residues forming this site is well conserved amongst the cathepsin L class. However, discrete differences are evident, which may account for different interactions with the substrate. In an effort to identify differences in the substrate specificity of SmCL1 and SmCL2, a panel of synthetic fluorogenic peptides of varying amino acids in the P_1 and P_2 positions were investigated as potential substrates for the purified, yeast-expressed ySmCL1 and ySmCL2 proteinases.

Mammalian cathepsin Ls have been shown to have a high affinity for the peptide substrate Z-Phe-Arg-NHMec and this substrate has classically been used to identify and characterise cathepsin L-like activities (Barrett and Kirschke, 1981; Mason *et al.*, 1985). Both ySmCL1 and ySmCL2 have a high affinity for Z-Phe-Arg-NHMec (k_{cat}/K_m = 38.86 $mM^{-1}s^{-1}$ & 406.3 $mM^{-1}s^{-1}$ respectively), and show little or no affinity for Z-Arg-Arg-

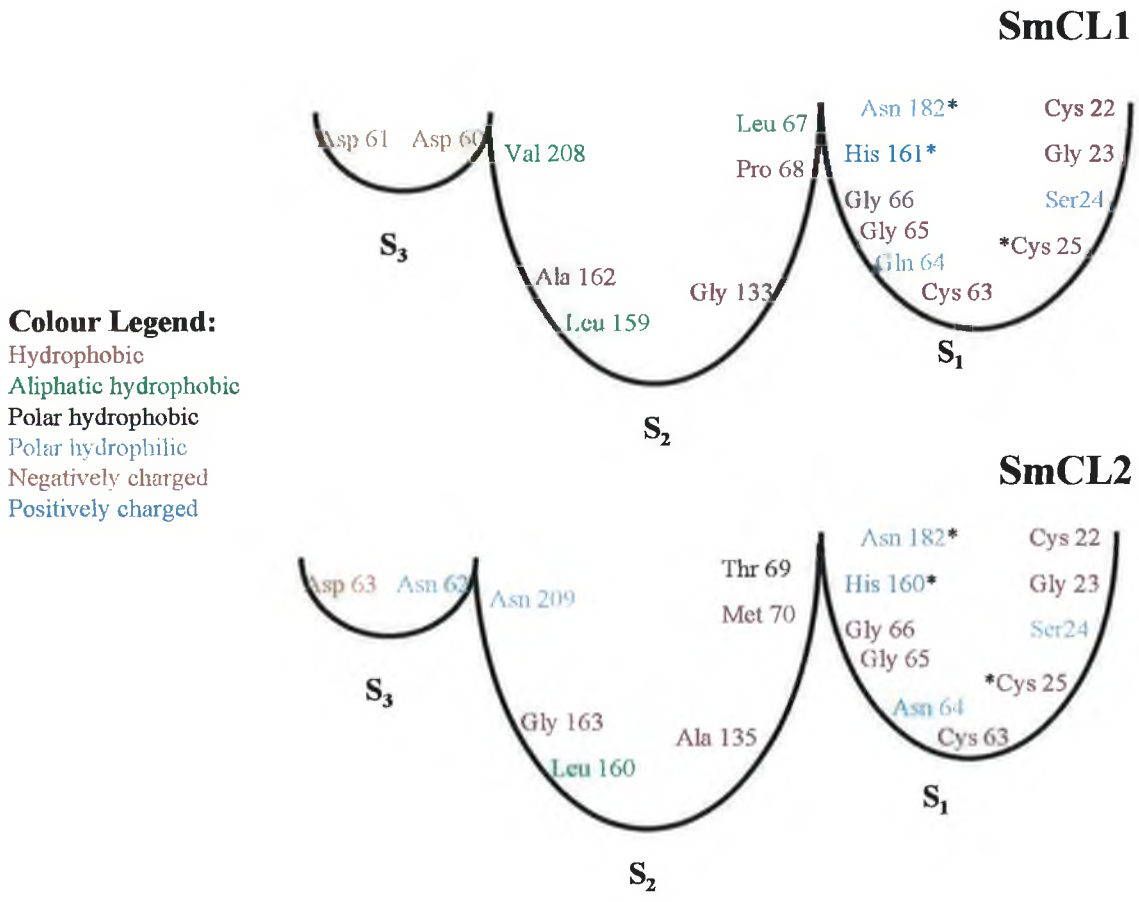


Fig 5.1. Schematic diagram of the amino acid residues which constitute the substrate binding pockets (S_1, S_2 and S_3 sites) of SmCL1 and SmCL2. '*' denotes residues of the catalytic triad involved in cleavage of the scissile bond. The numbering of amino acids is based on the papain system of Drenth *et al.* (1971). Homology modelling of the active sites was performed in collaboration with Dr. Ross Brinkworth, Centre for Drug Design, University of Queensland, Australia.

NHMec (k_{cat}/K_m =1.10 $mM^{-1}s^{-1}$ for ySmCL1 and no cleavage observed for ySmCL2), a substrate which is cleaved by cathepsin B-like enzymes. Cathepsin B-like proteinases contain a negatively-charged glutamate at the bottom of their S_2 subsite pocket, which favours the binding of basic P_2 residues like arginine. By contrast, the S_2 subsite of cathepsin Ls are generally very hydrophobic and open, which explains a specificity for the bulky aromatic side-chain of phenylalanine.

ySmCL2 shows a particularly high affinity for Z-Phe-Arg-NHMec (k_{cat}/K_m = 406.3 $mM^{-1}s^{-1}$), displaying over ten times higher affinity for this substrate in comparison to ySmCL1. SmCL2 shows higher homology to mammalian cathepsin Ls in comparison to SmCL1, both in its primary amino acid sequence and in the amino acids which make up the subsites of the active site (Dalton *et al.*, 1996a). Examination of the composition of the S_2 subsite of SmCL1 reveals that it is composed completely of hydrophobic amino acids, three of which are aliphatic (Leu 67, Leu 159, Val 208), whereas the S_2 subsite of SmCL2 is mainly composed of hydrophobic residues, with a hydrophilic end in the form of Asn 209. These differences would suggest that SmCL1 may bind aliphatic residues more efficiently in the P_2 position whereas SmCL2 may prefer aromatic residues, such as phenylalanine. Support for this hypothesis comes from a comparison of the affinities of ySmCL1 and ySmCL2 for Boc-Val-Leu-Lys-NHMec, relative to their affinities for Z-Phe-Arg-NHMec. Both enzymes show similar affinities for Boc-Val-Leu-Lys-NHMec (k_{cat}/K_m = 116.8 $mM^{-1}s^{-1}$ for ySmCL1 and 147.5 $mM^{-1}s^{-1}$ for ySmCL2). However, ySmCL1 cleaves Boc-Val-Leu-Lys-NHMec with approximately three times higher efficiency than its cleavage of Z-Phe-Arg-NHMec (k_{cat}/K_m = 38.86 $mM^{-1}s^{-1}$), while the opposite is true for ySmCL2 (k_{cat}/K_m = 406.3 $mM^{-1}s^{-1}$ for Z-Phe-Arg-NHMec). It would seem that ySmCL1

has a preference for the aliphatic leucine over the aromatic phenylalanine in the P_2 position.

A preference for Boc-Val-Leu-Lys-NHMec over Z-Phe-Arg-NHMec has also been observed for cathepsins L1 and L2 of *F. hepatica* (Dowd *et al.*, 1994, 1997; Roche *et al.*, 1997). Other cathepsin L-like cysteine proteinases such as the acid cysteine proteinase and the neutral thiol proteinase from the trematode parasite *Paragonimus westermani* (Yamakami and Hamajima, 1990; Yamakami *et al.*, 1995) and the acid-activatable cysteine proteinase of the cellular slime mould *Dictyostelium discoideum* (North *et al.*, 1996) also showed a marked preference for Boc-Val-Leu-Lys-NHMec. Boc-Val-Leu-Arg-NHMec, a substrate very similar to Boc-Val-Leu-Lys-NHMec, was cleaved much more efficiently than Z-Phe-Arg-NHMec by the cathepsin L-like cysteine proteinase (termed falcipain) from the malarial strain *Plasmodium falciparum* (Salas *et al.*, 1995). Interestingly, analogous to the proposed role of SmCL1 in schistosomes, falcipain has also been implicated to be involved in host haemoglobin digestion by malaria parasites. While Boc-Val-Leu-Lys-NHMec has been used as a substrate for chymotrypsin, there are no known reports of this peptide being used as substrate for mammalian cathepsin Ls. However, this substrate was shown to be cleaved by human osteoclast cathepsin K with four times less efficiency than Z-Phe-Arg-NHMec (Bossard *et al.*, 1996). It is tempting to speculate that Boc-Val-Leu-Lys-NHMec could potentially represent a substrate that could distinguish parasite cathepsin Ls from their mammalian homologues. Such a distinction could provide possible leads into the design of inhibitors directed specifically against the active sites of the parasite cathepsin Ls, and could thus have potential as novel chemothereputic agents.

Of the panel of substrates examined, SmCL2 only cleaved Z-Phe-Arg-NHMec and Boc-Val-Leu-Lys-NHMec, while SmCL1 cleaved a number of additional substrates. SmCL1 demonstrated a particularly high affinity for H-Leu-Val-Tyr-NHMec and Suc-Leu-Tyr-NHMec, indeed cleaving these substrates more efficiently than Z-Phe-Arg-NHMec. This is not altogether surprising, as SmCL1 would be expected to bind the aliphatic leucine and valine residues in the S_2 pocket as argued earlier. However, it is surprising that SmCL2 did not cleave these substrates also, at least to a lesser degree than SmCL1. Since the S_1 pocket of SmCL2 is virtually identical to SmCL1, it should accept Tyr in the P_1 position and its cleavage of Boc-Val-Leu-Lys-NHMec suggest that it can accept Leu in the P_2 position. It is possible that SmCL2 will not tolerate succinyl groups at P_3, although this is unlikely as the S_3 pockets of both enzymes are also very similar. The reason why SmCL2 does not cleave these substrates as opposed to SmCL1 is not clear, however, ongoing collaborative studies with Dr. Ross Brinkworth (University of Queensland, Australia) using computer modelling of the active sites of these proteinases may answer this enigma. These studies may show that the conformation of the active sites are influenced by other areas of the proteins keeping in mind that SmCL1 and SmCL2 are only 43% similar in amino acid sequence.

Tos-Gly-Pro-Arg-NHMec was moderately cleaved by SmCL1, but not at all by SmCL2. This substrate is cleaved with high efficiency by *Fasciola hepatica* cathepsin L2, but not by *F. hepatica* cathepsin L1 (Dowd *et al.*, 1994). It was used to distinguish between the two forms of cathepsin L in this parasite and the authors postulated that the *F. hepatica* cathepsin L1 may have a more 'open' S_2 pocket than *F. hepatica* cathepsin L2, thus allowing a bulky proline P_2 residue to bind in the S_2 pocket. Recent studies have

demonstrated that the S_2 pocket of SmCL1 is more open than that of SmCL2 (Ross Brinkworth, personal communication) and this may accommodate a bulky side chain such as that of proline. However, the moderate activity shown by SmCL1 for Tos-Gly-Pro-Arg-NHMec would indicate that the proline side chain may not fit into the S_2 pocket very well, and cleavage of this substrate may be more a reflection of the affinity of the S_3 site for glycine and the S_1 sites for arginine, rather than any specific affinity of the S_2 subsite for proline. This idea is supported by the fact that the other substrates examined which contained proline in the P_2 position (Tos-Gly-Pro-Lys-NHMec, Tos-Val-Pro-Arg-NHMec) were not cleaved at all by SmCL1.

The other substrates examined, Z-Phe-Ser-Arg-NHMec and Tos-Gly-Lys-Arg-NHMec and Tos-Leu-Gly-Arg-NHMec, were not cleaved by ySmCL1 or ySmCL2. This was expected for the substrates Z-Phe-Ser-Arg-NHMec and Tos-Gly-Lys-Arg-NHMec because a polar serine residue and a positively charged lysine residue in the P_2 position would not bind into the very hydrophobic S_2 pocket. It may be surprising that Tos-Leu-Gly-Arg-NHMec was also not cleaved by either enzyme, as the hydrophobic glycine residue in the P_2 position might have been expected to bind in the S_2 pocket. However, molecular modelling studies reveal that this substrate simply would not have enough sidechain contacts (glycine only has a single hydrogen atom sidechain) to bind into the active site with sufficient affinity to be cleaved (Ross Brinkworth, personal communication).

In summary, it would appear that SmCL1 favours substrates containing the following general motif : aliphatic(P_3)-aliphatic(P_2)-positive or polar hydrophobic(P_1), whereas SmCL2 may favour aliphatic(P_3)-aromatic(P_2)-positive(P_1). However, further

kinetic analysis with a larger panel of substrates coupled with further computer modelling of the active sites is required to fully understand the substrate specificity's, the makeup and structure of the active sites of these enzymes and to identify substrates which may have optimal binding energy.

The substrate specificities observed for ySmCL1 and ySmCL2 display distinct similarities with activities observed by Dalton *et al.* (1996a) in crude extracts and ES products of adult schistosomes. High levels of Z-Phe-Arg-NHMec cleavage were observed in soluble extracts of adult *S. mansoni*, both at pH 4.5 and pH 6.5 and this cleavage was attributed to be due to cathepsin L activity. This would seem to be the case as ySmCL1, and particularly ySmCL2, efficiently cleave Z-Phe-Arg-NHMec (Fig. 4.2). Indeed, a pH profile of Z-Phe-Arg-NHMec-cleaving activity in *S. mansoni* extracts shows the presence of what appears to be two distinct peaks of activity (Dalton *et al.*, 1996a). pH profiles of ySmCL1 and ySmCL2 against Z-Phe-Arg-NHMec show that ySmCL1 has a broad pH profile, peaking at pH 6.5, whereas ySmCL2 had a much narrower range of activity, peaking at pH 5.5. The two peaks of Z-Phe-Arg-NHMec-cleaving activity observed in *S. mansoni* extracts may, therefore, correspond to a combination of SmCL1 and SmCL2, with SmCL1 cleavage occurring over a broad pH range and SmCL2 cleavage occurring below pH 6.0. Cathepsin B in schistosome extracts may also contribute to the cleavage of Z-Phe-Arg-NHMec at pH 6.5 (Dalton *et al.*, 1996a), although it has been shown that recombinant *S. mansoni* cathepsin B does not cleave Z-Phe-Arg-NHMec very efficiently (Lipps *et al.*, 1996).

Cleavage of Boc-Val-Leu-Lys-NHMec was also observed in extracts of *S. mansoni*, although levels of cleavage of this substrate were over 50% lower than levels

observed for Z-Phe-Arg-NHMec (Dalton *et al.*, 1996a). Boc-Val-Leu-Lys-NHMec is a substrate which has been used to detect chymotrypsin-like activity. However, in the studies performed by Dalton *et al.* (1996a), activity against other chymotrypsin substrates was not observed. Furthermore, the Boc-Val-Leu-Lys-NHMec-cleaving activity was enhanced in reducing conditions characteristic of cleavage by a cysteine proteinase and these results led the authors to conclude that this cleavage to be due to a cathepsin L-like proteinase(s). This would, again, appear to be the case as both ySmCL1 and ySmCL2 efficiently cleave this substrate. Levels of cleavage of Boc-Val-Leu-Lys-NHMec were higher at pH 6.5 than at pH 4.5 in *S. mansoni* extracts and it is likely that this cleavage is due to SmCL1 as ySmCL2 is virtually inactive above this pH.

Examination of the pH profiles of the recombinant proteinases may also provide some clues into the possible functions of SmCL1 and SmCL2. ySmCL1 exhibited a broad bell-shaped activity profile against Z-Phe-Arg-NHMec, with an optimum activity at pH 6.5. The enzyme retained over 50% of its activity at pH 4.5 and at pH 8.0. *Fasciola hepatica* cathepsin L1 is also stable over a broad pH range, in contrast to human cathepsin L which is completely inactivated above pH 7.0 (Dowd *et al.*, submitted 1998). The instability of the mammalian cathepsin L at neutral pH has been proposed as one of the mechanisms by which mammalian cells are protected against accidental proteolysis by enzyme which may leak from the lysosomes (Mason *et al.*, 1985). The apparent stability of *F. hepatica* cathepsin L1 over a wide pH range is thought to be essential in order for this enzyme to perform its diverse biological functions, which include immune evasion, tissue penetration and blood feeding. Likewise, the stability of ySmCL1 over a broad pH range may indicate that SmCL1 may have a number of diverse functions or that the

enzyme may have to tolerate different physico-chemical environments. By contrast, ySmCL2 was not active above pH 6.5, and was more active in acidic conditions, characteristics similar to mammalian cathepsins indicating that SmCL2 may have a more defined function.

Summing up, it would appear that SmCL1 is active over a broader pH range and has the ability to cleave a broader range of substrates in comparison to SmCL2. Such observations may be manifested in the biological functions of these enzymes and is also further evidence that the two enzymes perform quite different functions.

5.5: Expression of SmCL1 and SmCL2 in schistosomes

It is thought that from a limited set of ancestral enzymes a complex series of proteinases evolved by a process of gene duplication and divergence. Subsequent to this divergence, the enzymes would have acquired a higher degree of specificity by tailoring their catalysis to a discrete suite of peptide bonds located at specific sites in protein substrates. At the same time, the expression and distribution of these new proteinases could also be restricted, generating proteinases for specific purposes and located in specific tissues (Neurath, 1984; Creighton and Darby, 1989). By determining the stage(s) or site(s) of expression of these enzymes, we may get further insights into the function of the proteinase. The expression of SmCL1 and SmCL2 in the life-cycle stages of *S. mansoni* was examined by RT-PCR and western blotting using anti-sera prepared against the recombinant proteinases expressed in *E. coli*.

RT-PCR analysis revealed that mRNA encoding for SmCL1 and SmCL2 was present in both male and female adult worms. This was expected, as the cDNAs encoding

for both of these proteinases were originally isolated from adult cDNA libraries (Smith *et al.*, 1994; Michel *et al.*, 1995). mRNA encoding for SmCL1 was also detected in cercariae, but not in eggs or miracidia, while SmCL2 was not detected in any of the larval stages. Z-Phe-Arg-NHMec-cleaving activity has been previously detected in extracts of schistosome eggs and miracidiae (Day *et al.*, 1995; Yoshino *et al.*, 1993), however, it would appear that this cleavage is not due to SmCL1 or SmCL2. Since most cDNA libraries are constructed from the mRNA isolated from adult schistosomes, it is possible that there are other as yet undiscovered cathepsin L-like cysteine proteinases expressed in these other stages of the schistosomes life cycle.

Z-Phe-Arg-NHMec-cleaving activity, with a pH optimum of 6.2, has been detected in extracts of *S. mansoni* cercariae (Dalton *et al.*, 1995a). Boc-Val-Leu-Lys-NHMec cleavage was also observed in cercarial extracts, although at much lower levels than Z-Phe-Arg-NHMec cleavage. mRNA encoding for SmCL1 was detected in cercariae and a band of 43 kDa, which seems to correspond to pro-SmCL1, was reactive on immunoblots with anti-bSmCL1 serum. This would suggest that SmCL1 is not responsible for the Z-Phe-Arg-NHMec-cleaving activities observed in cercarial extracts. However, it is possible that pro-SmCL1 may have been processed to an active, mature form during the preparation of the schistosome extract or the during the thirty mins of the enzyme assay. Additionally, or alternatively, cercariae may express an as yet undiscovered cathepsin L-like proteinase other than SmCL1 or SmCL2.

Indeed, the presence of the pro-SmCL1 in cercariae instead of the processed, mature form of SmCL1 would indicate that this enzyme may not be active at all in this stage. It is possible that SmCL1 is being stored in cercariae as the inactive pro-form in

order for its processing and activation at a later stage in development. SmCL1 has been implicated in the degradation of haemoglobin by schistosomes (Dalton *et al.*, 1996a and this thesis) and it known that the nascent digestive tract of the schistosome becomes active within hours following cercarial transformation into schistosomula after tissue penetration of the primary host occurs. Furthermore, *in vitro*, erythrocytes are ingested by schistosomula within days of transformation (Basch, 1981). Therefore, the reason for the expression of the inactive pro-form of SmCL1 in cercariae may be to ensure that the parasite has the digestive machinery in place in preparation for feeding once the skin is penetrated. Cathepsin L-like activity has been detected in developing schistosomula (Dalton *et al.*, 1995a), and this may be due to SmCL1.

The predicted size of mature SmCL1 from its deduced primary amino acid sequence is 24.1 kDa (Smith *et al.*, 1994), however, the bands detected on immunoblots ranged from 33 to 43 kDa. As discussed earlier, the recombinant proteinase is glycosylated by yeast which would suggest that it may also be glycosylated in schistosomes and this may explain why the molecule migrates higher than its predicted size on gels. The 33 kDa band would, therefore, appear to represent the mature SmCL1 while the 43 kDa protein would correspond to pro-SmCL1. A number of immunoreactive proteins were observed on blots at sizes ranging from 33 to 43 kDa may represent differentially processed zymogens of the pro-SmCL1 molecule or degradative products of the molecule resulting from enzymatic cleavage by another proteinases(s) present in the extract. SmCL2 was observed as an apparent mature, processed form which migrated at 32 kDa and a pro-enzyme at 39 kDa. As discussed earlier, while these sizes are higher than the predicted sizes deduced from the primary amino acid sequence of SmCL2 (Michel

et. al., 1995), it appears that this proteinase is not glycosylated, but rather, runs anomolously on gels.

Immunoblot analysis also showed that SmCL1 was present at approximately double the amount in female adult worms in comparison to males. It is estimated that fully mature male and female adult *S. mansoni* ingest 39,000 and 330,000 host red blood cells per hour respectively (Lawrence, 1973). Therefore, it would be expected that digestive enzymes involved in haemoglobin digestion would be overexpressed in female worms. However, such enzymes would be expected to be present in 8-10 times the amounts in females as opposed to males. It is possible that in males SmCL1 may be involved in other functions besides haemoglobin digestion. Immunolocalisation experiments performed in this study have shown that the enzyme is also present in the tegument of both male and female worms as well as in the gastrodermal cells lining the gut (Fig. 2.9); Furthermore, the labelling in the tegument was much more intense in males than in females. The function of SmCL1 in the tegument is not yet understood but it may play some role in surface membrane turnover as discussed later.

SmCL2 was present at approximately five times the amount in females in comparison to males. This was expected as the cDNA encoding for SmCL2 was originally isolated by subtractive hybridisation from a female-enriched cDNA library (Michel *et al.*, 1995). Unlike SmCL1, SmCL2 is not thought to be involved in haemoglobin degradation but its localisation to regions related to the reproductive system of females (Michel *et al.*, 1995) may explain why the enzyme is preferentially expressed in females. SmCL2 is thought to aid the passage of oocytes, vitelline cells and eggs through the uterus by altering the viscosity of the fluid in the ducts of the female reroductive system (Michel *et*

al., 1995). Such a function would require high levels of expression of SmCL2 in females, as the females produce huge numbers of eggs which need to be transported from the uterus to the female genital porus and eventually, to the exterior of the worm.

Both SmCL1 and SmCL2 were also shown to be present in the excretory/secretory (ES) products of adult worms, suggesting that the enzymes are secreted by these worms. Only the mature enzymes were present in the ES products, suggesting that the proteinases are processed as they are trafficked through the secretory pathway, releasing the mature, active zymogen. High levels of cathepsin L-like activity has been found previously in the ES products of *S. mansoni* (Dalton *et al.*, 1996a), and this is likely due to the presence of SmCL1 and SmCL2. Michel *et al.* (1992) postulated that SmCL2 may be secreted due to its apparent presence in the ducts of the female, and, to a lesser extent, the male reproductive systems. The signal obtained on immunoblots suggests that there are low levels of SmCL2 present in the ES products in comparison to the very high levels found in extracts of female worms. It is possible that the presence of SmCL2 in the ES products is due to leakage of the enzyme from the genital porus, rather than any active secretion of the enzyme into the medium by schistosome cells.

In this study, SmCL1 has been localised to the gastrodermal cells lining the gut lumen of adult schistosomes (Fig. 2.9) and the presence of the mature proteinase in the ES products strongly suggests that the enzyme is actively secreted into the gut lumen of the schistosome. As haemoglobin digestion occurs extracellularly in the gut lumen (Lawrence *et al.*, 1973), this data strongly suggests that SmCL1 plays a central role in haemoglobin catabolism by schistosomes (discussed in section 5.6).

5.6: SmCL1 plays a role in degradation of host haemoglobin

For a long time, the cathepsin B of schistosomes was thought to be central in the degradation of haemoglobin, and indeed, the schistosome cathepsin B was considered to be the 'haemoglobinase' (Linquist *et al.*, 1986; McKerrow and Doenhoff, 1988; Gotz and Klinkert, 1993). However, Dalton and co-workers demonstrated the presence of cathepsin L-like activity in extracts and ES products of *S. mansoni*, and showed that the specific activity of this enzyme(s) in these preparations was much higher than the cathepsin B-like proteinase(s) (Smith *et al.*, 1994; Dalton *et al.*, 1996a). This cathepsin L-like activity in adult worm extracts was ostensibly overlooked in previous studies of haemoglobin digestion because only the cathepsin B peptide substrate Z-Arg-Arg-NHMec was employed in the analysis (Linquist *et al.*, 1986; McKerrow and Doenhoff, 1988; Gotz and Klinkert, 1993). Furthermore, recent studies by Lipps *et al.* (1996) revealed that active, recombinant *S. mansoni* cathepsin B expressed in yeast does not show a marked preference towards haemoglobin as substrate, which led these authors to conclude that cathepsin B does not merit the term haemoglobinase. It is possible that in the earlier reports, cathepsin L-like activities was unknowingly co-purified with the cathepsin B-like activity since these enzymes are very similar in molecular size. Consequently, the cathepsin L-like activities may have contributed to the potent haemoglobinolytic activity observed in these preparations.

In this study, the haemoglobinolytic potential of ySmCL1 was examined by incubating human haemoglobin with the purified proteinase at various pH values and analysing the reaction products on SDS-PAGE gels. ySmCL1 exhibited a marked

preference for haemoglobin, cleaving this substrate at more than one site based on the smeared appearance of the digested products on gels. In contrast to peptide substrates and gelatin where ySmCL1 showed a pH optimum for activity of 6.5, the proteinase most efficiently cleaved haemoglobin at pH 4.5. Indeed, haemoglobin was not digested at pH 5.0 or above in these assays. This indicates that the denaturation of the haemoglobin by the acidic pH may be required before it can be digested by SmCL1. Since the physico-chemical environment of the schistosome gut appears to be acidic (Brindley *et al.*, 1997), haemoglobin may be denatured in such conditions, thereby exposing the peptide bonds and allowing proteinases to efficiently digest the protein.

Like schistosomes, malaria parasites also obtain their amino acid nutrients through the degradation of host haemoglobin. Degradation by malaria parasites takes place in a specialised food vacuole that has a pH optimum of approximately 5.0. Three endoproteinases present in this vacuole have been implicated to be involved in haemoglobin digestion in *Plasmodium falciparum*, the most common strain of malaria. Falcipain, a cathepsin L-like cysteine proteinase (Rosenthal *et al.*, 1988; Rosenthal, 1995; Salas *et al.*, 1995), and the aspartic proteinase plasmepsins I and II (Goldberg *et al.*, 1991; Francis *et al.* 1994) are thought to act in an ordered pathway of haemoglobin catabolism. The exact order of this pathway is, as yet, uncertain, however, both plasmepsin I (Goldberg *et al.*, 1991) and falcipain (Salas *et al.*, 1995) have been shown to readily cleave native haemoglobin and both have been implicated to be involved in the initial cleavages of native haemoglobin. This initial cleavage is thought to cause the molecule to unravel and denature, exposing the molecule to further proteolysis by the other digestive enzymes. The proteinases implicated in haemoglobin digestion in schistosomes (Fig. 1.1)

may operate in an ordered fashion similarly to that described for malaria, although no evidence of such a pathway has yet emerged. From this study it would appear that SmCL1 cannot cleave native haemoglobin and is unlikely to be involved in initial cleavages of haemoglobin in such an ordered pathway. Another proteinase, such as cathepsin D, which has been implicated to be involved in haemoglobin digestion, may perform these initial cleavages. However, it would seem unlikely that protozoan and helminth parasites, which are not very highly evolved organisms, would have developed such complex pathways for degradation of host tissues. The unravelling and denaturation of haemoglobin may be achieved by simpler means such as exposure to the low pH environment of the gut, thereby allowing the proteinases to randomly degrade the molecule.

Many of the various proteinases implicated to be involved in the degradation of haemoglobin by schistosomes, including cathepsin B (Ruppel *et al.*, 1985), cathepsin D (Bogitsh and Kirschner, 1986, 1987), cathepsin C (Bogitsh and Dresden, 1983) and schistosome legumain (Zhong *et al.*, 1995), have been localised to the epithelium of the schistosome gut and/or the gastrodermal cells surrounding the gut. As haemoglobin degradation is an extracellular process in schistosomes (Lawrence, 1973), secretion of digestive proteinases into the gut epithelium is essential if they are to be involved in this process. In this study, immunolocalisation experiments were performed to determine the site of expression and/or activity of SmCL1 within the adult schistosome. SmCL1 was shown to be expressed in the gastrodermal cells, which is consistent with the results that demonstrated the presence of SmCL1 in the ES products. Collectively, this data suggests that the enzyme is secreted into the gut epithelium of adult schistosomes.

It is noteworthy that SmCL1 is also located in the tegument of the adult worms. Other enzymes such as schistosome legumain (Zhong *et al.*, 1995) and cathepsin D (Bogitsh and Kirschner, 1986, 1987) that are associated with the gut have also been located in the tegument. These enzymes may function in intracellular turnover or in membrane biogenesis, in the tegument of the worm. Indeed, it is especially interesting that schistosome legumain is expressed in the same tissues as SmCL1 and cathepsin D, a finding which further supports its proposed role in the processing and activation of these enzymes.

To summarise, almost forty years after pioneering studies performed by Timms and Bueding (1959), which suggested that the catabolism of host haemoglobin by schistosomes is achieved by 'the action of one or several proteolytic enzymes present in the alimentary canal of the worms', the mechanism and components involved in this process are still not fully understood. Many proteinases are now implicated in this process (Fig 1.1; Dalton *et al.*, 1995a; Brindley *et al.*, 1997). The present results demonstrating the presence of SmCL1 in the gastrodermal cells lining the schistosome gut and in the ES products, and its ability to digest human haemoglobin at acidic pH indicate that SmCL1 plays some role in the progressive proteolysis of haemoglobin within the schistosome gut. This study, and ongoing studies on other putative haemoglobinases, would seem to support the early hypotheses of Timms and Bueding (1959).

5.7: Conclusions

SmCL1 and SmCL2 were functionally expressed in *Saccharomyces cerevisiae* and the secreted recombinant enzymes purified from the culture medium. The expression of the enzymes in their active form allowed for their biochemical characterisation. Both enzymes displayed biochemical characteristics typical of cathepsin L-like proteinases. However, recombinant SmCL1 and SmCL2 displayed distinct differences in their specificities for synthetic peptide substrates. SmCL1 favours substrates containing the following general motif : aliphatic (P_3)-aliphatic (P_2)-positive or polar hydrophobic (P_1), whereas SmCL2 may favour aliphatic (P_3)-aromatic (P_2)-positive (P_1). The enzymes also exhibited differences in their pH profiles of activity against Z-Phe-Arg-NHMec. SmCL1 maintained high levels of activity over a wide pH range, peaking at pH 6.5. In contrast, SmCL2 was only active between pH 2.7 and 6.5.

These results, coupled with studies on the expression of SmCL1 and SmCL2 in schistosomes would seem to suggest that these enzymes play quite different roles. Data presented in this study would seem to confirm the role of SmCL1 in the degradation of host haemoglobin by these parasites. Previous studies performed by Michel *et al.* (1995) have implicated SmCL2 to play some role in reproduction, possibly in formation of egg-shell precursor proteins or passage of eggs through the reproductive ducts.

Inhibition studies performed by Wasilewski *et al.* (1996) demonstrated that cysteine proteinase inhibitors blocked schistosome haemoglobin degradation *in vitro* and decreased worm burden and egg production in mouse models. These inhibitors were capable of inhibiting cathepsin L-like as well as cathepsin B-like proteinases. It is possible that the arrest in haemoglobin degradation observed by these authors was due to inhibiton

of SmCL1, as well as cathepsin B which the authors originally proposed. As SmCL2 has been implicated in the production of eggs in female worms (Michel *et al.*, 1995), the decrease in egg production observed by Wasilewski *et al.* (1996) may be due to inhibition of this enzyme.

If indeed SmCL1 and SmCL2 have roles to play in nutrition and reproduction in schistosomes, the development of inhibitors targeting these enzymes could represent a very promising approach to new chemotherapy that may arrest both infection and egg-induced pathology associated with schistosomiasis.

5.7: Future prospects

While this and previous studies have gone some way towards characterising SmCL1 and SmCL2, further studies are required to confirm their proposed biological roles. Ultrastructural immunolocalisation experiments are currently being performed in collaboration with Dr. Burton Bogitsch, Vanderbilt University, Nashville, Tennessee which may reveal the sub-cellular location of these proteinases and thereby give further insight into their functions. In particular, the role of SmCL1 in the tegument is unclear. If SmCL1 contributes to membrane turnover in the tegument, analysis of the ability of the recombinant proteinase to cleave structural proteins in the tegument, in addition to subcellular localisation studies, would go some way to answering this question.

Ongoing studies being performed in collaboration with Dr. Ross Brinkworth, Centre for Drug Design, University of Queensland, Australia, aim to identify specific substrates for SmCL1 and SmCL2 by homology modelling and mapping the active sites of the proteinases. This analysis may identify substrates which are cleaved preferentially by

one or both enzymes compared to their mammalian homologues. Molecular modelling may also allow determination of how the pro-region interacts with the active sites of SmCL1 and SmCL2, or more specifically, what residues within the pro-regions bind into the active site pockets to cause inhibition of activity. This data may provide valuable clues for the rational design of peptide inhibitors which may have potential chemotherapeutic value.

Ultimately, however, the design of inhibitors which specifically bind into the active sites of these molecules may depend on determination of the crystal structure of the proteinases. This study describing the functional expression and purification of recombinant SmCL1 and SmCL2 means that the possibility of crystallising these enzymes now exists. In order to perform such studies, the *S. cerevisiae* expression system and the purification scheme would need to be upscaled and optimised so as to produce enough protein to initiate crystallisation studies. Alternatively, other expression and purification systems could be examined.

As well as developing effective drugs, the successful eradication of schistosomiaisis would also involve an immunisation program designed to enhance host resistance to reinfection. No effective vaccine is as yet available, however, both SmCL1 and SmCL2 represent molecular vaccine candidates as these enzymes would appear to be centrally involved in functions which are vital for the survival of the parasite. In particular, if SmCL2 has a central role to play in egg formation in female adult worms, immunisation with this molecule could prove very effective. Blocking egg formation would cause a break in the life cycle and, therefore, stop the transmission of the disease. Furthermore, the hosts inflammatory response to the eggs causes the tissue pathology associated with

schistosomiasis (Mahmoud and Wahals, 1990) and immunisation with SmCL2 may go some way to reducing this pathology. Studies by Dalton *et al.* (1996a) demonstrated the induction of protective immunity in cattle against infection with *F. hepatica* by vaccination with cathepsin L proteinases. Future work would involve vaccination of mice with the recombinant SmCL1 and SmCL2 molecules and subsequent challenge with *S. mansoni* to assess their efficacy as vaccine candidates against schistosomiasis.

Chapter 6

Bibliography

Ammerer, G. (1983) Expression of genes in yeast using the ADC1 promoter. *Methods in Enzymology,* 101, 192-201.

Andersen, K., Tom, T. and Strand, M. (1991) Characterization of cDNA clones encoding a novel calcium-activated neutral proteinase from *Schistosoma mansoni*. *Journal of Biological Chemistry,* 266, 15085-15090.

Anderson, R.M. and May, K.S. (1982) Population biology of parasitic diseases : Part 1. *Nature,* 280, 361.

Asch, H.L. and Dresden, M.L. (1979) Acidic thiol proteinase activity of *Schistosoma mansoni* egg extracts. *Journal of Parasitology,* 65, 543-549.

Auriault, C., Ouaissi, M. A., Torpier, G., Eisen, H. and Capron, A. (1981) Proteolytic cleavage of IgG bound to the Fc receptor of *Schistosoma mansoni* schistosomula. *Parasite Immunology,* 3, 33-44.

Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirschke, H., Knight, C. G., Tamai, M. & Hanada, K. (1982) L-*trans*-epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. *Biochemical Journal,* 201, 189-198.

Barrett, A.J. and Kirschke, H. (1981) Cathepsin B, Cathepsin H and Cathepsin L. *Methods in Enzymology,* 80, 535-561.

Basch, P.J. (1981) Cultivation of *Schistosoma mansoni in vitro*. Establishment of culture from cercariae and development until pairing. *Journal of Parasitology,* 67, 179-185.

Becker, M.M., Harrop, S.A., Dalton, J.P., Kalinna, B.H., McManus, D.P., and Brindley, P.J. (1995) Cloning and characterization of the *Schistosoma japonicum* aspartic proteinase involved in hemoglobin degradation. *Journal of Biological Chemistry,* 270, 24496-24501.

Bogitsh, B.J. (1978) *Schistosoma mansoni*: uptake of exogenous hemoproteins by schistosomules grown *in vitro*. *Experimental Parasitology,* 45, 247-254.

Bogitsh, B.J. (1983) Peptidase activity in the egg-shell-enclosed embryo of *Schistosoma mansoni* using fluorescent histochemisrty. *Transactions of the American Microscopical Society* 102, 169-172.

Bogitsh, B.J. (1989) Observations on digestion in schistosomes or 'Blood and Guts'. *Transactions of the American Microscopical Society,* 108:1-5.

Bogitsh, B.J. and Dresden, M.D. (1983) *Schistosoma japonicum*: Biochemistry and cytochemistry of dipeptidyl aminopeptidase-II-like activity in adults. *Experimantal Parasitology,* 60, 163-170.

Bogitsh, B.J. and Kirschner, K.F. (1986) *Schistosoma japonicum*: Ultrastructural localization of a hemoglobinase using mercury labeled pepstatin. *Experimental Parasitology,* 62, 211-215

Bogitsh, B.J. and Kirschner, K.F. (1987) *Schistosoma japonicum*: Immunocytochemistry of adults using heterologous antiserum to bovine cathepsin D. *Experimental Parasitology,* 64, 213-218.

Bogitsh, B.J., Kirschner, K.F. and Rotmans, J.P. (1992) *Schistosoma japonicum*: Immunoinhibitory studies on hemoglobin using heterologous antiserum to bovine cathepsin D. *Journal of Parasitology,* 78, 454-459.

Bond, J.S. and Butler, P.E. (1987) Intracellular proteases. *Annual Reviews in Biochemistry,* 56, 333-364.

Bossard, M.J., Tomaszek, T.A., Thompson, S.K., Amegadzie, B.Y., Hanning, C.R., Jones, C., Kurdyla, J.T., McNulty, D.E., Drake, F.H., Gowen, M., Levy, M.A. (1996) Proteolytic activity of human osteoblast cathepsin K. Expression, purification, activation and substrate identification. *Journal of Biological Chemistry,* 271, 12517-12524.

Brindley, P.J. (1994a) Relationships between chemotherapy and immunity in schistosomiasis. *Advances in Parasitology,* 34, 134-161.

Brindley, P.J. (1994b) Drug resistance to schistosomicides and other anthelmintics of medical significance. *Acta Tropica,* 56, 213-231.

Brindley, P.J., Kalinna, B.M., Dalton, J.P., Day, S.R., Wong, J.Y.M., Smythe, M.L. and McManus, D.P. (1997) Proteolytic degradation of host hemoglobin by schistosomes. *Molecular and Biochemical Parasitology,* 89, 1-9.

Bromme, D., Bonneau, P.R., Lachance, P., Wiederanders, B., Kirschke, H., Peters, C., Thomas, D.Y., Storer, A.C. and Vernet, T. (1993) Functional expression of human cathepsin S in *Saccharomyces cerevisiae*; purification and characterisation of the recombinant enzyme. *Journal of Biological Chemistry,* 268, 4832-4838.

Bromme, D., Okamoto, K., Wang, B.B. and Biroc, S. (1996) Human cathepsin O2, a matrix-degrading cysteine protease expression in osteoclasts. Functional expression of human cathepsin O2 in *Spodoptera Frugiperda* and characterisation of the enzyme. *Journal of Biological Chemistry,* 271(4), 2126-2132.

Bundy, D.A.P., (1990) New initiatives in the control of helminths. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 84, 467-468.

Bundy, D.A.P., Hall, A., Medley, G.F. and Savioli. (1992) Evaluating measures to control intestinal parasitic infections. *World Health Statistics Quarterly,* 45, 168-179.

Butler, R., Michel, A., Kunz, W. and Klinkert, M-Q. (1995) Sequence of *Schistosoma mansoni* cathepsin C and its structural comparison with papain and cathepsins B and L of the parasite. *Protein and Peptide Letters,* 2, 313-320.

Butterworth, A.E. (1989) Control of schistosomiasis in man. In "The Biology of Parasitism", (Eds. P.T. Englund and A. Sher), p 43-59. Alan R. Liss, Inc., New York.

Carmona, C., Dowd, A.J., Smith, A.M. and Dalton, J.P. (1993) Cathepsin L proteinase secreted by *Fasciola hepatica in vitro* prevents antibody-mediated eosinophil attachment to newly excysted juveniles. *Molecular and Biochemical Parasitology,* 62, 9-18.

Carter, B.L.A., Irani, M., McKay, V.L., Seale, R.L., Sledziewski, A.V. and Smith, R.A. (1987) Expression and secretion of foreign genes in yeast. In "DNA Cloning", (Ed. D.M. Glover), Vol 3, IRL Press, Oxford.

Chappell, C.L. and Dresden, M.H. (1986) *Schistosoma mansoni*: Protease activity of "hemoglobinase" from the digestive tract of adult worms. *Experimental Parasitology,* 61, 160-167.

Chen, J-M., Dando, P.M., Rawlings, N.D., Brown, M.A., Young, N.E., Stevens, R.A., Hewitt, E., Watts, C. and Barrett, A.J. (1997) Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase. *Journal of Biological Chemistry,* 272, 8090-8096.

Coulombe, R., Grochulski, P., Sivaraman, J., Menard, R., Mort J.S. and Cygler, M. (1996) Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *EMBO Journal,* 15 (20), 5492-5503.

Coulombe, R., Li., Y., Takebe, S., Menard, R., Mason, P., Mort, J.S. and Cygler, M. (1996) Crystallisation and preliminary x-ray diffraction of human procathepsin L. *Proteins: Structure, Function, and Genetics,* 25, 398-400.

Creighton, T.E. and Darby, N.J. (1989) Functional evolutionary divergence of proteolytic enzymes and their inhibitors. *Trends in Biochemical Science,* 14, 319-324.

Croall, D.E. and Demartino, G.N. (1991) Calcium-activated neutral protease (calpain) system: structure, function, and regulation. *Physiological Reviews,* 71, 813- 847.

Curtis, R.H., Fallon, P.G. and Doenhoff, M.J. (1996) Sm480: a high molecular weight *Schistosoma mansoni* antigen associated with protective immunity. *Parasite Immunology,* 18, 149-157.

Dalton, J.P. and Brindley, P.J. (1996) Schistosome asparaginyl endopeptidase Sm 32 in hemoglobin digestion. *Parasitology Today,* 12, 125.

Dalton, J.P. and Brindley, P.J. (1997) Proteases of Trematodes. In "Advances in Trematode Biology". (Eds. B. Fried and T.K. Graczyk), Boca Raton (FL), CRC Press.

Dalton, J.P and Heffernan, M. (1989) Thiol proteases released *in vitro* by *Fasciola hepatica*. *Molecular and Biochemical Parasitology,* 35, 161-166.

Dalton, J.P., Smith, A.M., Clough, K.A. and Brindley, P.J. (1995a) Digestion of haemoglobin by schistosomes: 35 years on. *Parasitology Today,* 12, 299-303.

Dalton, J.P., Hola-Jamriska, L. and Brindley, P.J. (1995b) Asparaginyl endopeptidase activity in adult *Schistosoma mansoni*. *Parasitology,* 111, 575-580.

Dalton, J.P., Clough, K.A., Jones, M.K. and Brindley, P.J. (1996a) Characterization of the cathepsin-like cysteine proteinases of *Schistosoma mansoni*. *Infection and Immunity,* 64 (4), 1328-1334.

Dalton, J.P., McGonigle, S., Rolph, T.P. and Andrews, S.J. (1996b) Induction of protective immunity in cattle against infection with *Fasciola hepatica* by vaccination with cathepsin L proteinases and with hemoglobin. *Infection and Immunity,* 64, 5066-5074.

Dalton, J.P., Clough, K.A.,Jones, M.K. and Brindley, P.J. (1997a) The cysteine proteinases of *Schistosoma mansoni* cercariae. *Parasitology,* 114, 105-112.

Dalton, J.P., Day, S.R., Drew, A.C. & Brindley, P.J. (1997b) A method for the isolation of schistosome eggs and miracidia free of contaminating host tissues. *Parasitology,* 115, 29-32.

Darani, H.Y., Curtis, R.H.C., McNeice, C., Price, H.P., Sayers, J.R. and Doenhoff, M.J. (1997) *Schistosoma mansoni*: anomalous immunogenic properties of a 27 kDa larval serine protease associated with protective immunity. *Parasitology,* 115, 237-247.

Davis, A.H., Nanduri, J. and Watson, D.C. (1987) Cloning and gene expression of *Schistosoma mansoni* protease. *Journal of Biological Chemistry,* 262, 12851-12855.

Day, S.R., Dalton, J.P., Clough, K.A., Leonardo, L., Tiu, W.U., and Brindley, P.J. (1995) Cloning and characterization of cathepsin L proteinases of *Schistosoma japonicum*. *Biochemical and Biophysical Research Communications,* 217, 1-9.

Doenhoff, M.J. (1998) A vaccine for schistosomiasis: alternative approaches. *Parasitology Today,* 14, 105-109.

Dolinar, M., Maganja, D.B. and Turk, V. (1995) Expression of full-length human pro-cathepsin L cDNA in *Escherichia coli* and refolding of the expression product. *Biol. Chem. Hoppe Seyler*, 376, 385-388.

Douglass, J., Civelli, O. and Herbert, E. (1984) Polyprotein gene expression: generation of diversity of neuroendocrine peptides. *Annual Reviews in Biochemistry,* 53, 665-715.

Dowd, A.J., Smith, A.M., McGongle, S., Dalton, J.P. (1994) Purification and characterisation of a second cathepsin L proteinase secreted by the parasitic trematode *Fasciola hepatica*. *European Journal of Biochemistry,* 223, 91-98.

Dowd, A.J., McGongle, S. and Dalton, J.P. (1995) *Fasciola hepatica* cathepsin L proteinase cleaves fibrinogen and produces a novel type of fibrin clot. *European Journal of Biochemistry,* 232, 241-246.

Dowd, A.J., Tort, J., Roche, L. and Dalton, J.P. (1997) Isolation of a cDNA encoding *Fasciola hepatica* cathepsin L2 and functional expression in *Saccharomyces cerevisiae*. *Molecular and Biochemical Parasitology,* 88, 163-174.

Drenth, J., Jansonius, J.N., Koekoek, R., Wolthers, B.G. (1971) The structure of papain. *Advances in Protein Chemistry,* 25, 79-115.

Dresden, M.H. and Deelder, A.M. (1979) *Schistosoma mansoni*: Thiol protease properties of adult worm "hemoglobinase". *Experimental Parasitology,* 48, 190-197.

Francis, S.E., Gluzman, I.Y., Oksman, A., Knickerbocker, A., Mueller, R., Bryant, M.L., Sherman, D.R., Russell, D.G. and Goldberg, D.E. (1994) Molecular characterization and inhibition of a *Plasmodium falciparum* aspartic hemoglobinase. *EMBO Journal*, 13, 306-317.

Garoff, H. (1985) Using recombinant DNA techniques to study protein targeting in the eucaryotic cell. *Annual Reviews in Cell Biology,* 1, 403-445.

Goldberg, D.E., Slater, A.F.G., Beavis, R., Chait, B., Cerami, A. and Henderson, G.B. (1991) Haemoglobin degradation in the human malaria pathogen Plasmodium falciparum: a catabolic pathway initiated by a specific aspartic protease. *Journal of Experimental Medicine,* 87, 2931-2935.

Goll, D.E., Thompson, V.F., Taylor, R.G. and Zalewska, T. (1992) Is calpain activity regulated by membranes and autolysis or by calcium and calpastatin? *Bioessays,* 14, 549-556.

Gotz, B. and Klinkert, M-Q. (1993) Expression and partial characterization of a cathepsin B-like enzyme (Sm31) and a proposed "haemoglobinase"(Sm32) from *Schistosoma mansoni*. *Biochemical Journal,* 290, 801-806.

Griffiths, G. and Simons, K. (1986) The trans-Golgi network: sorting at the exit site of the golgi complex. *Science,* 234, 438-443.

Groves, M.R., Taylor, M.A.J., Scott, M., Cummings, N.J., Pickersgill, R.W. and Jenkins, J.A. (1996) The prosequence of procaricain forms an a-helical domain that prevents access to the substrate binding cleft. *Structure,* 15 (4), 1193-1203.

Halton, D.W. (1997) Nutritional adaptations to parasitism within the platyhelminths. *International Journal Parasitology,* 27, 693-704.

Hasnain, S., Hirama, T., Huber, C.P., Mason, P. and Mort, J.S. (1993) Characterization of cathepsin B specificity by site-directed mutagenesis: importance of Glu^{245} in the S_2-P_2 specificity for arginine and its role in transition state stabilization. *Journal of Biological Chemistry,* 268, 235-240.

Hirschberg, C.B and Snider, M.D. (1987) Topology of glycosylation in the rough endoplasmic reticulum and the Golgi apparatus. *Annual Reviews in Biochemistry,* 56, 63-87

Hooker, C.W. and Brindley, P.J. (1996) Cloning and characterisation of strain-specifiic transcripts encoding triosephosphate isomerase, a candidate vaccine antigen from *Schistosoma japonicum*. *Molecular and Biochemical Parasitology,* 82, 265-269.

Hota-Mitchell, S., Siddiqui, A.A., Dekabin, G.A., Smith, J., Tognon, C. and Podesta, R. (1997) Protection against *Schistosoma mansoni* infection with a recombinant baculovirus-expressed subunit of calpain. *Vaccine,* 15, 1631-1640.

Ischii, S-I. (1994) Legumain: Asparaginyl endopeptidase. *Methods in Enzymology,* 244, 604-615

Jaffe, J.J., Meymarian, E. and Doremus, H.M. (1971) Antischistosomal action of tubercidin administered after absorption into red blood cells. *Nature*, 230, 408-409.

Janknecht, R., de Martynoff, G., Lou, J., Hipskind, R.A., Nordheim, A., and Stunnenburg, H.G. (1991) Rapid and efficieint putification of native histidine-tagged protein expressed by recombinant vaccinia virus. *Proceedings of the National Academy of Science, USA*, 88, 8972-8976.

Jankovic, D., Åslund, L., Oswald, I.P., Caspar, P., Champion, C., Pearce, E., Colligan, J. E., Strand, M., Sher, A. and James, S.L. (1996) Calpain is the target antigen of a Th1 clone that transfers protective immunity against *Schistosoma mansoni*. *Journal of Immunology,* 157, 806-814.

Jensen, R.A (1976) Enzyme recruitment in evolution of new function. *Annual Reviews in Microbiology,* 30, 409-425.

Johnston, D.A.J. (1997) The WHO/UNDP/Word Bank *Schistosoma* Genome initiative: Current status. *Parasitology Today,* 13, 45-46.

Kane, S.E. (1993) Mouse procathepsin L lacking a functional glycosylation site is properly folded, stable and secreted by NIH 3T3 cells. *Journal of Biological Chemistry,* 268 (15), 11456-11462.

Karcz, S.R., Podesta, R.B., Siddiqui, A.A., Dekaban, G.A, Strejan, G.H. and Clarke, M.W. (1991) Molecular cloning and sequence analysis of a calcium-activated neutral protease (calpain) from *Schistosoma mansoni*. *Molecular and Biochemical Parasitology,* 49, 333-336.

Karrer, K.M., Peiffer, S.L. and DiTomas, M.E. (1993) Two distinct gene subfamilies within the family of cysteine protease genes. *Proceedings of the National Academy of Science, USA,* 90, 3063-3067.

Kasschau, M.R., Robinson, D.C., and Dresden, M.H. (1986) *Schistosoma mansoni*: activation of hemolytic activity in homogenates from live adult worms. *Experimental Parasitology,* 62, 442-449.

Kelly, R.B. (1985) Pathways of protein secretion in eukaryotes. *Science,* 230, 25-31.

Khembhavi, A.A., Buttle, D.J., Knight, C.G. and Barrett, A.J. (1993) The two cysteine endopeptidases of legume seeds: purification and characterization by use of specific fluorometric assays. *Archieves in Biochemistry and Biophysics,* 303, 208-213.

Kirschke, H. and Barrett, A.J. (1987) Chemistry of lysosomal proteases. In 'Lysosomes : their role in protein breakdown'. (Eds. H. Glaumann and F.J. Ballard), Academic Press, London, UK, pp. 467-473.

Klinkert, M-Q., Cioli, D., Shaw, E., Turk, V., Bode, W. and Butler, R. (1994) Sequence and structure similarities of cathepsin B from the parasite *Schistosoma mansoni* and human liver. *FEBS Letters,* 351, 397-400.

Klinkert, M-Q., Felleisen, R., Link, G., Ruppel, A. and Beck, E. (1989) Primary structures of Sm31/32 diagnostic proteins of *Schistosoma mansoni* and their identification as proteases. *Molecular and Biochemical Parasitology,* 33, 113-122.

Klinkert, M-Q., Ruppel, A. and Beck, E. (1987) Cloning of diagnostic 31/32 kilodalton antigens of *Schistosoma mansoni*. *Molecular and Biochemical Parasitology,* 25, 247-255.

Kopitar, G., Dolinar, M., Strukeli, B., Pungercar, J. and Turk, V. (1996) Folding and activation of human procathepsin S from inclusion bodies produced in *Escherichia coli*. *European Journal of Biochemistry,* 236, 558-562.

Kuhelj, R., Dolinar, M., Pungercar, J., and Turk, V. (1995) The preparation of catalytically active human cathepsin B from its precursor expressed in *Escherichia coli* in the form of inclusion bodies. *European Journal of Biochemical,* 229, 533-539.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature*, 227, 680-685.

Lang, L., Reitman, M., Tang., J., Roberts, R.M. and Kornfeld, S. (1984) Lysosomal enzyme phosphorylation. *Journal of Biological Chemistry,* 259, 14663-14617.

Lawrence, J.D. (1973) The ingestion of red blood cells by *Schistosoma mansoni*. *Journal of Parasitology,* 59, 60-63.

Leatherbarrow, R. J. (1987) Enzfitter, Elsevier Biosoft, Cambridge.

Lipps, G., Fullkrug, R. and Beck, E. (1996) Cathepsin B of *Schistosoma mansoni*. Purification and activation of the recombinant proenzyme secreted by *Saccharomyces cerevisiae*. *Journal of Biological Chemistry,* 271, 1717-1725.

Lindquist, R.N., Senft, A.W., Petitt, M. and McKerrow, J.H. (1986) *Schistosoma mansoni*: purification and characterization of the major acidic proteinase from adult worms. *Experimental Parasitology,* 61, 160-167.

Lodish, H.F. (1988) Transport of secretory and membrane glycoproteins from the rough endoplasmic reticulum to the Golgi. *Journal of Biological Chemistry,* 263, 2107-2110.

Mahmoud, A.A.F. and Wahals, M.F.A. (1990) Schistosomiasis. In "Tropical and Geographic Medicine", 2nd edition, (Eds. K.S. Warren and A.A.F. Mahmoud), p 438-473. McGraw-Hill, New York.

Marikovsky, M., Arnon, R. and Fishelson, Z. (1988a) Proteases secreted by transforming schistosomula of *Schistosoma mansoni* promote resistance to killing by complement. *Journal of Immunology,* 141, 273- 278.

Marikovsky, M., Fishelson, Z. and Arnon, R. (1988b) Purification and characterization of proteases secreted by transforming schistosomula of *Schistosoma mansoni*. *Molecular and Biochemical Parasitology,* 30, 45-54.

Marikovsky, M., Parizade, M., Arnon, R. and Fishelson, Z. (1990a) Complement regulation on the surface of cultured schistosomula and adult worms of *Schistosoma mansoni*. *European Journal of Immunology,* 20, 221- 227.

Marikovsky, M., Arnon, R. and Fishelson, Z. (1990b) *Schistosoma mansoni*: localization of the 28 kDa secreted protease in cercaria. *Parasite Immunology,* 12, 389-401.

Mason, R.W., Green, G.D. and Barrett, A.J. (1985) Human liver cathepsin L. *Biochemical Journal,* 226, 233-241.

Mason, R.W., Wilcox, D., Wilkstrom, P., and Shaw, E.N. (1989) The identification of active forms of cysteine proteinases in Kirsten-virus-transformed mouse fibroblasts by use of a specific radiolabelled inhibitor. *Biochemical Journal,* 257, 125-129.

McGuire, M.J., Lipsky, P.E. and Thiele, D.L. (1997) Cloning and characterization of the cDNA encoding mouse dipeptidyl peptidase I (Cathepsin C). *Biochimica et Biophysica Acta,* 1351, 267-273.

McKerrow, J.H. and Doenhoff, M.J. (1988) Schistosome proteases. *Parasitology Today,* 4, 334-340.

McKerrow, J. H., Newport, G. and Fishelson, Z. (1991) Recent insights into the structure and function of the larval proteinase involved in host infection by a multicellular parasite. *Proceedings of the Society of Experimental Biology and Medicine,* 192, 119-124.

Merckelbach, A., Hasse, S., Dell, R. and Ruppel, A. (1994) cDNA sequences of *Schistosoma japonicum* coding for two cathepsin B-like proteins and Sj32. *Tropical Medicine and Parasitology,* 45, 193-198.

Michel, A., Ghoneim, H., Resto, M., Klinkert, M-Q and Kunz, W. (1995) Sequence, characterization and localization of a cysteine proteinase cathepsin L in *Schistosoma mansoni*. *Molecular and Biochemical Parasitology,* 73, 7-18.

Minard, P., Murrell, K.D., and Stirewalt, M.A. (1977) Proteolytic, antigenic and immunogenic properties of *Schistsoma mansoni* cercarial secretion material. *American Journal of Tropical Medicine and Hygiene,* 26, 491-499.

Moir, D.T. and Davidow, L.S. (1991) Production of proteins by secretion from yeast. *Methods in Enzymology,* 194, 491-507.

Neurath, H. (1984) Evolution of proteolytic enzymes. *Science,* 224, 350-357.

Neurath, H. (1990) The diversity of proteolytic enzymes. In "Proteolytic Enzymes : A Practical Approach", (Eds. Beynon, R.L., and Bond, J.S.), IRL Press, Oxford.

Nishimura, Y. and Kato, K. (1988) Identification of latent procathepsin H in microsomal lumen: characterization of proteolytic processing and enzyme activation. *Archieves in Biochemistry and Biophysics,* 260, 712-718.

North, M.J., Nicol, K., Sands, T.W. and Cotter, D.A. (1996) Acid-activatable cysteine proteinases in the cellular slime mould *Dictyostelium discoidium*. *Journal of Biological Chemistry,* 271, 14462-14467.

Peters, W., and Gilles, H.M. (1995) Schistosomiasis. In "Color Atlas of Tropical Medicine and Parasitology", Fourth Editon. pp 96-109. Mosby-Wolfe Publishers, London, UK.

Pfeffer, S.R. (1988) Mannose 6-Phosphate receptors and their role in targeting of proteins to lysosomes. *Journal of Membrane Biology,* 103, 7-16.

Pierrot, C., Godin, C., Liu, J.L., Capron, A. and Khalife, J. (1996) *Schistosoma mansoni* elastase: an immune target regulated during the parasite life-cycle. *Parasitology,* 113, 519-526.

Rapoport, T.A. (1985) Extensions of the signal hypothesis-sequential insertion model versus aliphatic tunnel hypothesis. *FEBS Letters,* 187, 1-10.

Roche, L., Dowd, A. J., Tort, J., McGonigle, S., McSweeney, A., Curley, P., Ryan, T., and Dalton, J.P. (1997) Functional expression of *Fasciola hepatica* cathepsin L1 in *Saccharomyces cerevisiae*. *European Journal of Biochemistry,* 245, 373-380.

Rosenthal, P.J. (1995) Effects of proteinase inhibitors on globin hydrolysis by cultured malaria parasites. *Experimental Parasitology,* 88, 272-281.

Rosenthal, P.J., McKerrow, J.H., Aikawa, M., Nagasawa, H. and Leech, J.H. (1988) A malarial cysteine proteinase is necessary for haemoglobin degradation by *Plasmodium falciparum*. *Journal of Clinical Investigations,* 82, 1560-1566.

Rothman, J.E. (1987) Protein sorting by selective retention in the endoplasmic reticulum and the Golgi stack. *Cell,* 50, 521-522.

Rowan, A.D., Mason, P., Mach, L. and Mort, J.S. (1992) Rat procathepsin B. Proteolytic processing to the mature form in vitro. *Journal of Biological Chemistry,* 267, 15993-15999.

Ruppel, A., Breternitz, U. and Burger, R. (1987) Diagnostic Mr 31 000 *Schistosoma mansoni* proteins: requirement of infection, but not immunization, and use of the "miniblot" technique for the production of monoclonal antibodies. *Journal of Helminthology,* 61, 95-101.

Ruppel, A., Diesfeld, H.J. and Rother, U. (1985) Immunoblot analysis of *Schistosoma mansoni* antigens with sera of schistosomiasis patients: Diagnostic potential of an adult schistosome polypeptide. *Clinical and Experimental Immunology,* 62, 499-506.

Salas, F., Fichmann, J., Lee, G.K., Scott, M.D., and Rosenthal, P.J. (1995) Functional expression of falcipain, a *Plasmodium falciparum* cysteine proteinase, supports its role as a malarial hemoglobinase. *Infection and Immunity,* 63, 2120-2125.

Salvesan, G. and Nagase, H. (1990) Inhibiton of proteolytic enzymes. In "Proteolytic Enzymes : A Practical Approach", (Eds. Beynon, R.L., and Bond, J.S.), IRL Press, Oxford.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). In "Molecular Cloning : A laboratory manual", 2nd. edition. Cold Spring Harbor Laboratory Press, NY.

Schekman, R. and Novick, P. (1982). The secretory process and yeast cell-surface assembly. In "The molecular biology of the yeast *Saccahromyces*" , Eds. Strathern, J.N., Jones, E.W. and Broach, J.R. Cold Spring Harbor Laboratory, Cold Spring Harbour, New York.

Siddiqui, A.A., Zhou, Y., Podesta, R.B., Karcz, S.R., Tognon, C.E., Strejan, G.H., Dekaban, G.A. and Clarke, M. W. (1993) Characterization of Ca^{++}-dependent neutral protease (calpain) from human blood flukes, *Schistosoma mansoni*. *Biochimica et Biophysica Acta,* 1181, 37- 44.

Smith, A.M, Dowd, A.J., Heffernan, M., Robertson, C.D. and Dalton, J. P. (1993) *Fasciola hepatica*: a secreted cathepsin L-like proteinase cleaves host immunoglobulin. *International Journal of Parasitology,* 23, 977-983

Smith, A.M., Dalton, J.P., Clough, K.A., Kilbane, C.L., Harrop, S.A., Hole, N. and Brindley, P.J. (1994) Adult *Schistosoma mansoni* express cathepsin L proteinase activity. *Molecular and Biochemical Parasitology,* 67, 11-19.

Smith, S.M. and Gottesman, M.M. (1989) Activity and deletion analysis of recombinant human cathepsin L expressed in *Escherichia coli*. *Journal of Biological Chemistry,* 264(34), 20487-20495.

Stanley, P. (1987) Glycosylation mutants and the function of mammalian carbohydrates. *Trends in Genetics,* 3, 77-81.

Stirewalt, M.A. (1974) *Schistosoma mansoni*: cercaria to schistosomule. *Advances in Parasitology,* 12, 115- 182.

Sung, C.K. and Dresden, M.L. (1986) Cysteinyl proteinases of *Schistosoma mansoni* eggs: purification and partial characterization. *Journal of Parasitology,* 72, 891-900.

Tao, K., Stearns, N.A., Dong, J., Wu, Q. and Sahagian, G. (1994) The proregion of cathepsin L is required for proper folding, stability and ER exit. *Archives of Biochemistry and Biophysics,* 311, 19-27.

Tarentino, A.L., Plummer, T.H. and Maley, F. (1974) The release of intact oligosaccharides from specific glycoproteins by endo-N-acetyl-glucosaminidase H. *Journal of Biological Chemistry,* 249, 818-824.

Taylor, M.A.J., Pratt, K.A., Revell, D.F., Baker, K.C., Sumner, I.G. and Goodenough, P.W. (1992) Active papain renatured and processed from insoluble recombinant propapain expressed in *Escherichia coli. Protein Engineering*, 5, 455-459.

Timms, A. R. and Bueding, E. (1959) Studies of a proteolytic enzyme from *Schistosoma mansoni*. *British Journal of Pharmacology,* 14, 68-73.

Valasco, G., Ferrando, A.A., Puente, X.S., Sanchez, L.M. and Lopez-Otin, C. (1994) Human cathepsin O. Molecular cloning from breast carcinoma, production of active enzyme in *Escherichia coli*, and expression analysis in human tissues. *Journal of Biological Chemistry,* 269, 27136-27142.

Vernet , T., Chatellier, J., Tessier, D.C. and Thomas, D.Y. (1993) Expression of functional papain precursor in *Saccharomyces cerevisiae*: rapid screening of mutants. *Protein Engineering,* 6, 213-219.

Verwaerde, C., Auriault, Neyrinck, J.L. and Capron, A. (1988) Properties of serine proteases of *Schistosoma mansoni* schistosomula involved in the regulation of IgE synthesis. *Scandinavian Journal of Immunology,* 27, 17-24.

Wasilewski, M.M., Lim, K.C., Phillips, J. and McKerrow, J.H. (1996) Cysteine protease inhibitors block schistosome hemoglobin degradation *in vitro* and decrease worm burden and egg production *in vivo*. *Molecular and Biochemical Parasitology,* 81, 179-189.

Wittlin, S., Rosel, J. and Stover, D.R. (1998) One step purification of cathepsin D by affinity chromatography using immobilized propeptide sequences. *European Journal of Biochemistry,* 252, 530-536.

Wong, J.Y.M, Harrop, S.A., Day, S.R. and Brindley, P.J. (1997) Schistosomes express two forms of cathepsin D. *Biochimica et Biophysica Acta,* 1338, 156-160.

World Health Organization (1996) WHO Fact Sheet N112. World Health Organization, Geneva, Switzerland.

Xu, Y-Z. and Dresden, M.H. (1986) Leucine aminopeptidase and hatching of *Schistosoma mansoni* eggs. *Journal of Parasitology,* 72, 507-511

Yagel, S., Warner, A.H., Nellans, H.N., Lala, P.K., Waghorne, C. and Denhardt, D.T. (1989) Suppression by cathepsin L inhibitors of the invasion of amnion membranes by murine cancer cells. *Cancer Research,* 49, 3553-3557.

Yamakami, K. and Hamajima, F. (1990) A neutral thiol protease secreted from newly excysted metacercariae of trematode parasite *Paragonimus westermani*: purification and characterisation. *Comparative Biochemistry and Physiology,* 95B, 755-758.

Yamakami, K., Hamajima, F., Akao, S. and Tadakuma, T. (1995) Purification and characterisation of acid cysteine protease from metacercariae of the mammalian trematode parasite *Paragonimus westermani. European Journal of Biochemistry,* 233, 490-497.

Yamamoto, M., Yamakami, K., and Hamajima, F. (1994) Cloning of a cDNA encoding a neutral thiol protease *Paragonimus westermani* metacercariae. *Molecular and Biochemical Parasitology,* 64, 345-348.

Yoshino, T.P., Lodes, M.J., Rege, A.A., and Chappell, C.L. (1993) Proteinase activity in miracidia, transformation excretory-secretory products, and primary sporocysts of *Schistosoma mansoni*. *Journal of Parasitology,* 79, 23-31.

Zhong, C., Skelly, P.J., Leaffer, D., Cohn, R.G., Caulfield, J.P. and Shoemaker, C.B. (1995) Immunolocalization of a *Schistosoma mansoni* facilitated glucose transporter to the basal, but not the apical, membranes of the surface syncytium. *Parasitology,* 110, 383-394

Zussman, R.A., Bauman, P.M., Petruska, J.C. (1970) The role of ingested hemoglobin in the nutrition of *Schistosoma mansoni*. *Journal of Parasitology,* 56, 75-79.

Appendix

Publications

Brady, C.P., Brindley, P.J., Bogitsh, B.J., Dowd, A.J. and Dalton, J.P. (1999) Expression of cathepsins L1 and L2 of *Schistosoma mansoni* suggests discrete differences in their biological functions. In preparation.

Brady, C.P., Dowd, A.J., Brinkworth, R., Dalton, J.P., and Brindley, P.J. (1999) Exploration of subsite binding specificity of cathepsins L1 and L2 of *Schistosoma mansoni* through kinetic studies of the recombinant enzymes and homology modelling. In preparation.

Brady, C.P., Dowd, A.J., Brindley, P.J., Ryan, T., Day, S.R. and Dalton, J.P. (1999) Recombinant expression and localisation of *Schistosoma mansoni* cathepsin L1 supports its role in the degradation of host hemoglobin. *Infection and Immunity*, in press.

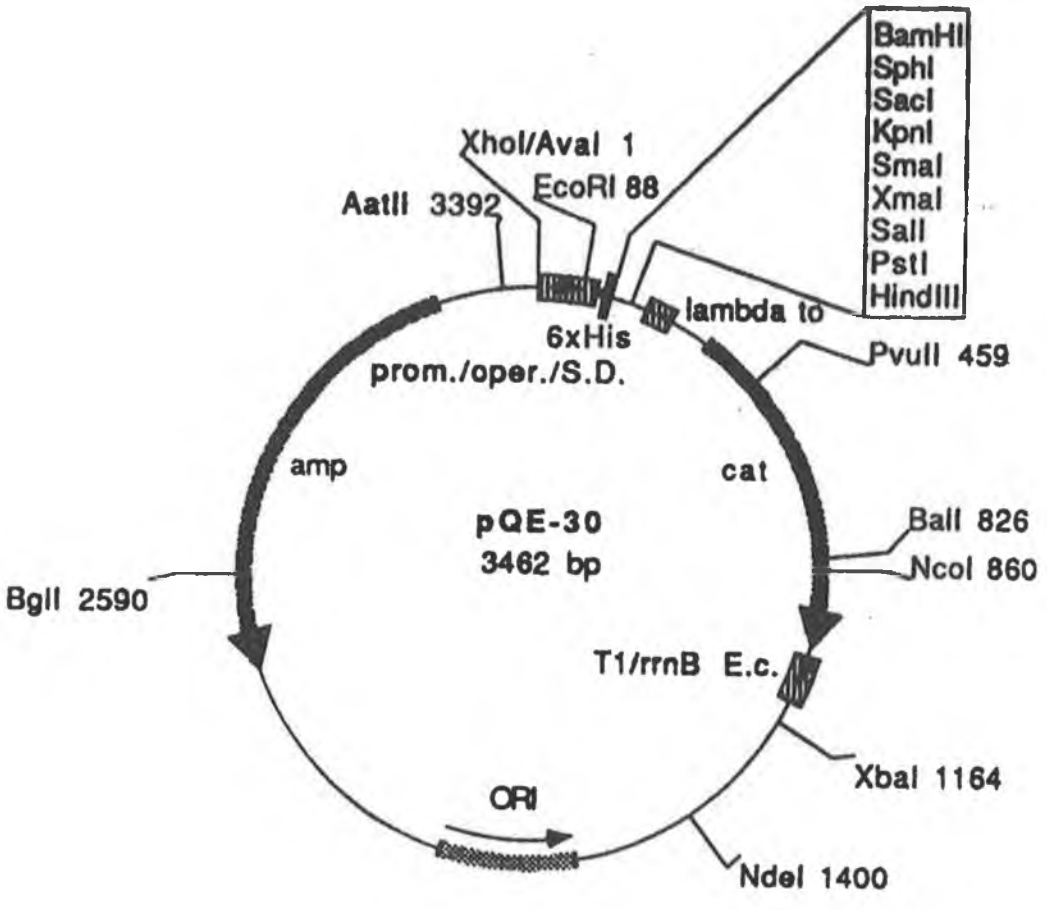
Brady, C.P., Dowd, A.J. and Dalton, J.P. (1999) Cysteine proteases of parasitic trematodes; novel targets for chemotherapy and immunoprophylaxis. *Biochemical Society Transactions*, in press.

Presentations

Brady, C.P., Dowd, A.J., Brindley, P.J., Ryan, T., Day, S.R. and Dalton, J.P. (1998) Recombinant expression and localisation of *Schistosoma mansoni* cathepsin L1 supports its role in the degradation of host hemoglobin. Poster presention to the Biochemical Society, Dublin City University, Ireland.

Brady, C.P., Dowd, A.J., Brinkworth, R., Brindley, P.J. and Dalton, J.P. (1998) Cathepsin L proteinases of *Schistosoma mansoni*; targets for novel drug design and immunoprophylaxis. Presented to Irish Society for Parasitology, University College Dublin, Ireland.

Brady, C.P., Dowd, A.J., Dalton, J.P. and Brindley, P.J. (1998) Functional expression and biochemical characterisation of recombinant cathepsins L1 and L2 of *Schistosoma mansoni*. Presented to the Molecular Parasitology Unit, The Queensland Institute of Medical Research, Brisbane, Australia.


BamHI
SphI
SacI
KpnI
SmaI
XmaI
SalI
PstI
HindIII
XhoI/AvaI 1
EcoRI 88
AatII 3392
lambda to
6xHis
prom./oper./S.D.
PvuII 459
amp
cat
pQE-30
3462 bp
BalI 826
NcoI 860
BglI 2590
T1/rrnB E.c.
XbaI 1164
ORI
NdeI 1400