Isolation and characterisation of the GPI:protein transamidase from *Leishmania mexicana*

James D. Hilley

Wellcome Centre for Molecular Parasitology
Anderson College
University of Glasgow
Glasgow
G11 6NU

This thesis is presented in submission for the degree of Doctor of Philosophy in the Faculty of Veterinary Medicine

Submitted November 1999

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ABSTRACT

Many eukaryotic cell surface proteins are attached to the plasma membrane by glycosylphosphatidylinositol (GPI) anchors. Trypanosomatid parasitic protozoa such as *Leishmania* and *Trypanosoma brucei* make extensive use of this method of protein surface attachment. The perceived importance of GPI-anchored proteins makes the GPI biosynthetic pathway a good target for anti-parasite chemotherapy. The terminal step in the pathway is the addition of complete pre-formed GPI anchors to the carboxyl-termini of proteins. This involves replacing a GPI signal sequence with a GPI anchor in a transamidation reaction. Two of the proteins involved in this step have been identified in yeast and mammals. One of these components, GPI8, has significant homology to a family of plant cysteine proteinases, the legumains, and GPI8 is therefore believed to be the catalytic subunit of the transamidase.

The work described in this thesis focussed on the cloning of the *GPI8* gene from *Leishmania mexicana* and its characterisation. The predicted protein shares 31% identity with yeast and human homologues. The nucleotide sequence of a fragment of the *T. brucei GPI8* gene has also been obtained. Targeted gene replacement of the single copy *L. mexicana GPI8* produced *GPI8* null mutants. The loss of *GPI8* was confirmed by Southern blotting. The phenotype of the *GPI8* null mutants was analysed with the following findings: (1) Mutant promastigotes grow well in culture. (2) GP63 is not detected on the promastigote surface and is greatly reduced in promastigote lysates. (3) Episomal re-expression of *GPI8* restored GP63 to the cell surface. (4) *GPI8* null mutants are able to infect macrophages *in vitro* to approximately wild type levels, and replicate within macrophages. (5) *GPI8* null

mutants are capable of forming lesions in mice. These data show that GPI-anchored proteins of *L. mexicana* are not essential for growth of promastigotes, invasion of macrophages by promastigotes, or infection of mice. It remains to be established if GPI-anchored proteins are required for survival in the sandfly.

Sequence comparison of *L. mexicana*, yeast and human GPI8 proteins identified two potential active site cysteine residues. Mutation in which Cys216 was converted to a glycine led to loss of GPI8 activity, as assessed by GP63 surface expression, indicating that this may be the active site cysteine.

Recombinant GPI8 was produced in *E. coli* and used to inoculate rabbits for the production of anti-serum. These antibodies detected recombinant protein but failed to detect protein in *L. mexicana* cell lysates. Antibodies raised against a peptide of GPI8 also recognised recombinant GPI8, but not GPI8 in *L. mexicana* lysates. A protein of 38-40 kDa, however, was detected in cell lysates of *T. brucei* with the anti-recombinant GPI8 antibodies. This antiserum gave a similar immunofluorescence pattern in *T. brucei* to those of an epitope-tagged ribosomal protein, QM. There is some overlap in fluorescence patterns between tagged QM and the endoplasmic reticulum marker protein BiP. This suggests that GPI8 is located in the ER of *T. brucei*, the same subcellular location that has been identified in yeast. A fusion protein of GPI8 and GFP was expressed in *L. mexicana* to attempt to localise GPI8 in this organism. The fusion protein, however, was unable to restore GPI8 function in the null mutant, therefore data on its localisation cannot be inferred.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF ABBREVIATIONS

CAT	Chloramphenicol acetyltransferase
CIAP	Calf intestinal alkaline phosphatase
CRK	Cdc2-related kinase
DABCO	1,4-diazobicyclo[2.2.2.]octate
DAPI	4,6-diamidino-2-phenylindole
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
DPMS	Dolichol phosphate mannose synthase
ECL	Enhanced chemiluminescence
EDTA	Ethylenediaminetetraacetate
ER	Endoplasmic reticulum
FACS	Fluorescence activated cell sorting
FCS	Foetal calf serum
GIPL	Glycoinositolphospholipid
GPI	Glycosylphosphatidylinositol
HRP	Horseradish peroxidase
LPG	Lipophosphoglycan
MHC	Major histocompatibility complex
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PIG	Phosphatidylinositol glycan
RNA	Ribonucleic acid

RNAse A	Ribonuclease A
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
TBS	Tris-buffered saline
Tris	Tris(hydroxymethyl)amino methane
UV	Ultraviolet
VSG	Variant surface glycoprotein

Measurements

bp	base pairs
kb	kilobase pairs
Da	daltons
kDa	kilodaltons
μg	micrograms
mg	milligrams
ng	nanograms
M	molar
mM	millimolar
μM	micromolar
μl	microlitres
ml	millilitres
V	volts

NOTE ON NOMENCLATURE

Nomenclature for genes and proteins in *Leishmania*, trypanosomes and *Plasmodium* follow the guidelines proposed by Clayton *et al.*, 1998. Genetic nomenclature for the organisms described in this thesis is shown below.

Leishmania, trypanosomes and *Plasmodium*:

Gene: *GPI8* Protein: GPI8

Saccharomyces cerevisiae:

Gene: *GPI8* Protein: Gpi8p Mutant: *gpi8*

Homo sapiens

Gene: *GPI8* Protein: GPI8

Other species:

As for *Leishmania*, trypanosomes and *Plasmodium*.

A number of *Leishmania mexicana* clones are described in this thesis by their trivial names (e.g. P1S1). These clones have also been assigned Wellcome Centre for Molecular Parasitology (WCMP) numbers and these are listed in an Appendix (page 230).

DECLARATION

I declare that the work presented in this thesis is my own except where otherwise stated

James D. Hilley

November 1999

ACKNOWLEDGEMENTS

Firstly, I'd like to thank both of my supervisors, Jeremy Mottram and Graham Coombs, for their help, support and advice for the last three years or so (has it really been that long?). The encouragement offered by Jeremy and Graham, both during the lab years and especially during the writing of this thesis, was immense. I am deeply indebted to Mick North who has been a hugely influential figure, as my lecturer and undergraduate project supervisor at Stirling University and as a friend ever since.

Paul "if you want to know something really obscure about trypanosomatids, ask me" Hassan provided many good humoured discussions as a bench partner (and friend) and never could say no to a pint or two, which helped to contribute to the arguable retention of my sanity. This, of course, brings me on to Drs Darren Brooks, Colin Robertson, Hubert Denise, David Fergusson, Joaquin Muriel, Deepak Sharma, the remainder of Jeremy's and Graham's groups and all at WCMP (both past and present); thanks go to those brave souls not only for their inability to refuse beer/wine, but for their help, advice and friendly discussions/arguments about Apple Macs Vs PCs, Cava being just as good as Champagne, only much cheaper... you name it. Simon Lillico provided a sympathetic ear for my "writing-up whinges" being in the same boat himself. Finally, Affy and Pauline deserve medals for their support (will some biscuits do?).

My wife Abi, has been a tremendous support throughout and without her love and encouragement, I'm not sure where I'd be today. My friends, family and Abi's parents have also been hugely supportive.

I dedicate this thesis to the memory of my father, James Hilley.

CHAPTER 1

GENERAL INTRODUCTION

1.1 The trypanosomatids

Several species of the trypanosomatid protozoa are the causative agent of parasitic diseases of mammals, including humans. These organisms have been the subject of extensive research for many years, not only because of the implications of the range of diseases they cause, both in terms of medical and economic issues, but also because they have some significant differences from other eukaryotic species. The nature of the often very hostile conditions that these species must endure to survive and proliferate has resulted in the evolution of some unusual features that are of considerable interest to eukaryote biology. Many of these features, because they are not present or are used less extensively in higher eukaryotes such as humans, may serve as potential targets for anti-parasitic chemotherapy. Current drugs used for the treatment of Leishmaniasis and human African trypanosomiasis, such as the antimonial-based Pentostam® and the arsenic-based melarsoprol (Mel B®) are far from ideal with the associated side effects and the ever-increasing problem of parasite drug-resistance (Croft *et al.*, 1997). Thus, new targets for anti-parasite chemotherapy are always being sought because the medical and economical implications are the most important reason for study from a humanitarian standpoint.

1.1.1 Parasitic trypanosomatids

This section will introduce two of the most important species of trypanosomatid pathogens.

(I) **An African trypanosome:**

Trypanosoma brucei has a number of subspecies. *Trypanosoma brucei brucei* infects a variety of mammals including cattle, horses and camels causing a serious wasting disease known as nagana. *Trypanosoma brucei gambiense* is the causative agent of the chronic form of the human disease known as African sleeping sickness (African trypanosomiasis) and is distributed throughout Central and West Africa. *Trypanosoma brucei rhodesiense* causes the acute form of this disease and is found mainly in East and Southern Africa. The parasites invade the central nervous system of the host causing symptoms of severe headache, drowsiness and weakness that, if left untreated often results in death. The parasites are spread by the tsetse fly (*Glossina*) and are transmitted by the bite of this insect. Collectively, the African trypanosomes are endemic in 36 countries and as many as 55 million people are at risk of infection (Molyneux, 1997)

(II) ***Leishmania* species:** There are as many as 20 known species of *Leishmania* and these are loosely grouped into old world (Africa, Asia and Europe) and new world (the Americas) species according to their distribution. The spectrum of diseases caused by these organisms are called the leishmaniases and these can be grouped into three main classes:

(A) Cutaneous leishmaniasis: production of skin ulcers on the exposed parts of the body (face, arms and legs). Up to 200 lesions can

form on the host causing permanent scarring and sometimes severe disability. Species: *Leishmania mexicana*, *L. major* and *L. tropica*.

(B) Mucocutaneous leishmaniasis: usually begins with skin ulcers but can lead to extensive damage to the mucous membranes of the nose and throat causing often horrific disfigurement. Species: *L. braziliensis*, *L. peruviana*

(C) Visceral leishmaniasis: also known as kala azar, this is the most serious form of the disease producing symptoms such as fever, significant weight loss, enlargement of the liver and spleen and other secondary effects such as diarrhoea and anaemia. The World Health Organisation (WHO) reports that the fatality rate can be as high as 100% if this disease is not treated (WHO Division of Control of Tropical Diseases). Species: *L. donovani*, *L. infantum*.

Leishmania species are found in 88 countries throughout Africa, Asia, the Americas and Southern European countries. This is the most widespread of the trypanosomatids and as many as 350 million people are at risk of infection with as many as 119 million currently infected (Molyneux, 1997).

1.1.2 Trypanosomatid life cycles

Both *Leishmania* and trypanosomes have a digenetic life cycle; that is, the life cycle requires two hosts: an insect host and a mammalian host and the parasites are transmitted from one host to the next undergoing specific morphological alterations at

each stage. The life cycle of the African trypanosome involves transmission of the mammalian-infective (metacyclic) form of the parasite from the salivary glands of the tsetse fly (*Glossina*). When the infected fly takes a blood meal, the metacyclics are passed directly into the bloodstream of the mammalian host where they differentiate into the proliferative long slender form. As the infection progresses, the parasite becomes a short stumpy form that is non-dividing. These parasites are then consumed by another tsetse fly during a blood meal and, because the short stumpy form is preadapted for survival within the insect, the parasite rapidly becomes the procyclic form, once again capable of cell division. These procyclics migrate forward through the insect midgut to the salivary glands ready for mammalian infection once more.

Leishmania is passed directly into the bloodstream of the mammalian host by the bite of the female sandfly (Figure 1.1). The non-dividing metacyclic parasites are taken up by macrophages via phagocytosis where they differentiate to the non-motile amastigote form, which begin to undergo cell division once more. Amastigotes survive and replicate within the parasitophorous vacuole, essentially a distended lysosome. As the number of amastigotes within a macrophage increases, the macrophage eventually dies releasing the amastigotes, which then go on to infect other macrophages. This proceeds until infected macrophages are taken up in the blood meal of another feeding sandfly. In the gut of the sandfly, the amastigotes emerge from macrophages and differentiate to the flagellated procyclic promastigotes. These parasites migrate forward through the insect gut to reach the mouthparts, where they are passed onto a fresh mammalian host when the sandfly next feeds.

Within the mammalian host *Leishmania* survives intracellularly. By contrast *T. brucei* are free living within the bloodstream of the host. Conditions within the insect host are strikingly different to those of the mammalian host explaining the need for the complex developmental changes observed at each life cycle stage.

1.2 Survival strategies in mammalian host

Leishmania species exist intracellularly in the mammalian host. Uptake by host cells occurs soon after infection thus minimising the time of exposure to the host immune response. By contrast, *T. brucei* is free-living within the bloodstream of the host and is thus subject to antibody-mediated responses. To cope with this, bloodstream form African trypanosomes possess an electron-dense surface coat of the variant surface glycoprotein (VSG), first visualised by electron microscopy by Professor Keith Vickerman. After some time in the bloodstream, antibodies against the VSG are produced and many of the parasites are killed. A small number survive because they have begun to express a different VSG protein to which antibodies are not available and thus escape the antibody response. These parasites proliferate until yet more antibodies are produced when the VSG again is replaced on the surface of a few parasites, and so the process known as antigenic variation continues (Barry, 1997).

Leishmania metacyclic promastigotes are taken up by host macrophages via phagocytosis. Survival of the promastigotes in the bloodstream prior to macrophage entry and macrophage itself is thought to be mediated by the cell surface molecules (Mauel, 1996), which include the major surface metalloproteinase GP63, glycoinositolphospholipids (GIPLs) and lipophosphoglycans (LPGs) and these are

described in more detail in Section 1.4.2. The phagocytotic vesicle containing the promastigote fuses with lysosomes and parasites differentiate to the amastigote forms (Figure 1.1). The amastigotes proliferate despite the harsh conditions within the parasitophorous vacuole (Alexander and Vickerman, 1975) and the reasons for this are still not clearly defined (Mauel, 1996). It could be partly due to the nature of the cell surface molecules, predominantly GIPLs, as mutant *L. amazonensis* deficient in particular amastigote-specific GIPLs exhibit arrested growth following differentiation to amastigotes during *in vitro* macrophage infections (Mensa-Wilmot *et al.*, 1999).

The individual roles of the cell surface molecules of *Leishmania* in parasite invasion and survival in both mammalian and insect hosts are described in more detail in Section 1.4.2.

1.3 The proteinases of trypanosomatids

As previously mentioned, much of the study of parasitic protozoa is designed to identify and characterise potential anti-parasite drug targets. One group of enzymes that have been identified as potential targets are the proteinases (North *et al.*, 1990b; McKerrow, 1989). Generally, proteinases are enzymes capable of breaking down proteins through the hydrolytic cleavage of peptide bonds. Proteolytic enzymes can be divided into two categories: endopeptidases and exopeptidases (North, 1991). The former, cleave peptide bonds within the polypeptide chains of proteins whereas exopeptidases cleave peptide bonds at the ends of polypeptide chains thereby liberating single amino acids. Exopeptidases may be further sub-divided according to whether they cleave at the N or the C terminus of polypeptides. Proteinases are divided into four classes: cysteine, serine and aspartic proteinases, so-called because

of the nature of their active site residue, and metalloproteinases, which require the co-ordination of a metal ion (usually zinc) for activity. Cysteine and metallo- proteinases have received the greatest attention in trypanosomatids to date. In addition to the cysteine proteinase active site residue, cysteine, a histidine residue is also required, making up the so-called catalytic dyad. Serine and aspartic proteinases of protozoa will not be considered further.

A classification scheme for cysteine proteinases has been proposed (Barrett and Rawlings, 1996). The cysteine proteinases are grouped into families (named C1, C2, C3, etc.) based on primary sequence similarities. The families are further separated into clans based on tertiary structure (where known). Thus, to date, 35 families have been identified, most of which are divided into four clans (named CA, CB, CC, and so on) (Barrett and Rawlings, 1996). Where tertiary structure information is unavailable, the order in which the two residues that make up the catalytic dyad appear in the primary sequence, can be utilised to allow clan assignation. The C1 family is the largest with the representative cysteine proteinase called papain and includes the mammalian cathepsins L and B and the major lysosomal proteinases of trypanosomatids (Barrett and Rawlings, 1996). It is these enzymes that have received considerable interest within this and other laboratories, however the identification of cysteine proteinases of other families in trypanosomatids has led to the work presented in this thesis. The asparaginyl endopeptidases, which belong to the unclanned C13 family of cysteine proteinases (Barrett and Rawlings, 1996) include the legumains of plants, implicated in the maturation of seed proteins (Hara-Nishimura *et al.*, 1995) and the "haemoglobinase", SM32, of *Schistosoma mansoni*, which also may be involved in protein maturation rather than direct haemoglobin

degradation (Dalton *et al.*, 1995a). These endopeptidases exhibit a preference for cleavage of peptide bonds carboxyl to asparagine residues, hence the title "asparaginyl endopeptidases". Such processing enzymes have yet to be identified in trypanosomatids but may be required for the maturation one or more of the lysosomal cysteine proteinases described in the next section.

1.3.1 Cysteine proteinases of trypanosomatids

Mammalian cathepsin L and cathepsin B are produced as pre-pro-proteins, where the nascent polypeptide is considerably larger than the mature enzyme (Mason, 1991). The short pre-domain, which is involved in targeting the newly synthesised pre-pro-proteinase to the endoplasmic reticulum (ER), is cleaved by a signal peptidase in the ER. The pro-region, which is immediately carboxyl to the pre- domain, is removed during trafficking (possibly by autoproteolysis) yielding the mature proteinase (Mason, 1991). The pro-region is thought to remain associated with the enzyme during trafficking, behaving as an inhibitor to prevent unwanted proteolysis (Carmona *et al.*, 1996).

The cysteine proteinases identified in *L. mexicana*, which are also synthesised as preproproteins are lysosomal and can be placed in one of three groups depending on their structure and activity (Coombs and Mottram, 1997).

Type I cysteine proteinases are cathepsin L-like and are characterised by long C-terminal extensions of unknown function that are post-translationally removed (Souza *et al.*, 1992; Mottram *et al.*, 1996a). Type I cysteine proteinases of *L. mexicana* are encoded by a tandemly arranged array of *CPB* genes with 19 copies. The various

members of the array encode isoforms of CPB with subtly different substrate preferences and differential expression during *L. mexicana* life cycle (Mottram *et al.*, 1997). Equivalent enzymes from *T. brucei* (Mottram *et al.*, 1989) and the South American trypanosome, *T. cruzi* (Eakin *et al.*, 1992) have also been identified and characterised.

The Type II CP of *L. mexicana* is encoded by a single copy gene, *CPA*, and is also a cathepsin L-like enzyme (Mottram *et al.*, 1992). CPA has a short C terminal extension.

Type III CP is also encoded by a single copy gene, *CPC* (Bart *et al.*, 1995). CPC is more closely related to mammalian cathepsin B than cathepsin L and shows marked differences in substrate preferences to CPB (Bart *et al.*, 1995; Robertson and Coombs, 1993).

Trypanosome homologues of type II have not been identified but type III has been identified in *T. cruzi* (Nobrega *et al.*, 1998).

The genes encoding the leishmanial cysteine proteinases have been knocked out by targeted gene disruption. The *CPA* null mutant has no obvious phenotypical difference to wild type cells (Souza *et al.*, 1994), however, the *CPC* null mutant, whilst still able to infect mice, the survival rate within macrophages is lower (Bart *et al.*, 1997). Disruption of the entire *CPB* gene array has produced interesting results (Mottram *et al.*, 1996a). The *CPB* null mutants were still infective to mice albeit at greatly reduced levels and *in vitro* macrophage infectivity of these mutants is

significantly less than wild type, implicating the Type I Cysteine proteinases as virulence factors (Mottram *et al.*, 1996a).

1.3.2 Metalloproteinases of trypanosomatids

GP63 is the major promastigote surface metalloproteinase of *Leishmania* and is also present in an intracellular form in amastigotes (Frommel *et al.*, 1990). Like CPB, GP63 is encoded by multi copy genes located in a tandem array in *L. major* (Button *et al.*, 1989). The *L. mexicana* promastigote enzyme is attached to the cell surface by a glycosylphosphatidylinositol anchor (see (Medina-Acosta *et al.*, 1989) and Section 1.4). The role of this enzyme is not clear although it has been implicated in the avoidance of complement-mediated lysis by promastigotes first exposed to the mammalian host (Brittingham *et al.*, 1995; Joshi *et al.*, 1998). It has also been suggested to have a role in macrophage entry, via the complement pathway by cleavage of a complement component to facilitate uptake by complement receptors on the macrophage surface (Brittingham *et al.*, 1995). However, deletion of six of the seven genes from *L. major* does not appear to alter the ability of *L. major* promastigotes to enter and survive within macrophages either *in vitro* or in vivo (Joshi *et al.*, 1998). Expression of the seventh gene occurs in metacyclic promastigotes and amastigotes, so perhaps the product of this gene alone is fulfilling the role of GP63 in macrophage entry. The abundance of GP63 on the promastigote surface suggests a role in interaction with the insect host (Coombs and Mottram, 1997) however the *L. major* mutants lacking GP63 encoded by genes 1-6 were capable of infecting and developing within the sandfly (*Phlebotomus argentipes*) (Joshi *et al.*, 1998).

Homologues of GP63 have also been identified in African trypanosomes (El-Sayed and Donelson, 1997). A possible role proposed for a GP63 homologue in African trypanosomes, which do not exist intracellularly in their mammalian host is also in avoidance of complement-mediated lysis. Because, these cells exist in the blood of the host, exposure to the components of the complement pathway will be constant.

1.4 The glycosylphosphatidylinositols (GPIs)

Many cell surface proteins of eukaryotes are attached to the plasma membrane by way of a glycosylphosphatidylinositol (GPI) anchor. This method of attachment is used extensively by many species of parasitic protozoa including the trypanosomatids (McConville and Ferguson, 1993). The major surface metalloproteinase, GP63 of *Leishmania* and the variant surface glycoprotein (VSG) of African trypanosomes are GPI-anchored (described above). Many of the enzymes involved in GPI biosynthesis in mammals and yeast have been identified to date and some potentially exploitable differences have already been identified.

1.4.1 Protein-GPI biosynthesis in mammals, *Saccharomyces cerevisiae* and trypanosomes

The biosynthesis of GPI-protein anchors has received considerable attention over the last decade (Englund, 1993; McConville and Ferguson, 1993; Kinoshita *et al.*, 1997; Stevens, 1995). The anchors are produced in the endoplasmic reticulum (ER), initially on the cytoplasmic face and then on the lumenal face of the ER membrane, in a series of enzyme catalysed reactions, before being transferred *en bloc* to protein precursors (Vidugiriene and Menon, 1993). Although the pathways are somewhat

conserved between protozoa, yeast and humans, fundamental differences have, and continue to be discovered supporting the long-held belief that the GPI biosynthetic pathway will provide excellent anti-fungal and anti-protozoal drug targets (McConville and Ferguson, 1993). In this section, yeast, human and *Trypanosoma brucei* protein GPI anchor biosynthesis is considered. *T. brucei* is included as a representative of the protozoan pathway, owing to the volume of research carried out with this organism, but reference is made to *Leishmania* where relevant. The process is summarised in Figure 1.2.

1.4.1.1 Addition of N-acetylglucosamine to phosphatidylinositol, the first step of GPI anchor biosynthesis

The first step in the biosynthetic pathway involves the addition of N-acetylglucosamine (GlcNAc) from a UDP-GlcNAc donor, to phosphatidylinositol (PI) (Doerring *et al.*, 1989; Hirose *et al.*, 1991). This has been studied in cell-free systems using radiolabelled GlcNAc together with mammalian cell lysates and purified microsomes as well as with yeast membranes (see (Stevens, 1995)). GlcNAc is soluble in aqueous solution until it becomes complexed with PI. Interestingly, this simple step requires four proteins in mammals and at least three in yeast (Watanabe *et al.*, 1998; Tiede *et al.*, 1998; Inoue *et al.*, 1996). The human proteins initially identified are known as PIG-A, PIG-C, PIG-H, so-called because of the name given to classes of human mutant cell lines defective in GPI-anchoring ability and all three classes were found to be defective in GlcNAc transferase activity (Stevens and Raetz, 1991; Sugiyama *et al.*, 1991) and GPI1.

The human *PIG-A* gene was cloned by complementation of the mutant phenotype by a stretch of cDNA including the gene leading to the restoration of GPI-anchored proteins (Miyata *et al.*, 1993). The human disease known as paroxysmal nocturnal haemoglobinuria, which is attributed to defects in GPI-anchor biosynthesis is caused by a mutation in the *PIG-A* gene (Takeda *et al.*, 1993). A homologue of this gene, *GPI3*, has been identified in *S. cerevisiae* by complementation of a temperature-sensitive mutant defective in GlcNAc transferase activity (Vossen *et al.*, 1995; Leidich *et al.*, 1995). These cells were unable to incorporate [^{3}H]inositol into proteins at non-permissive temperatures. The yeast mutant lacked *in vitro* GlcNAc transferase activity (Leidich *et al.*, 1995). Characterisation of *GPI3* revealed significant homology of the predicted product to PIG-A (44% identity). Homology to a bacterial GlcNAc transferase (RFAK) was identified also (Vossen *et al.*, 1995) leading to the proposal that PIG-A is the catalytic subunit of the multi-protein complex. PIG-A is a 484 amino acid (54kDa) ER-membrane protein with a large cytoplasmic domain (Watanabe *et al.*, 1996), a short ER-transmembrane domain and a lumenal domain at the carboxyl-terminus. This domain is required for activity of the protein, probably because of its role in targeting the protein to the rough ER membrane (Watanabe *et al.*, 1996).

The second protein component of the GlcNAc transferase machinery is PIG-C and the yeast homologue is Gpi2p, sharing about 20% identity with one another (Leidich *et al.*, 1995; Inoue *et al.*, 1996). This protein is also an ER-membrane protein, lacking the ER-signal peptide and is extremely hydrophobic, possessing 6 or 7 potential transmembrane domains. One of the features of *in vitro* mammalian biosynthesis is that the source of PI is important for the activity of the GlcNAc

transferase complex (Watanabe *et al.*, 1998). For example, bovine PI was found to be a 100 fold more efficient substrate than soybean PI and the *lyso* form of bovine PI was a poor substrate. This suggests that the GlcNAc transferase complex is capable of recognising PI fatty acyl chains and the hydrophobic nature of PIG-C/Gpi2p means that this protein could be involved in this (Watanabe *et al.*, 1998).

PIG-H is the third component of the mammalian GlcNAc transferase, and interestingly, a yeast homologue for this protein has not been identified, despite the fact that sequencing of the *S. cerevisiae* genome is now complete (see Nature supplement, 29 May 1997). The human gene was cloned by complementation of the mutant class H with a cDNA encoding a 188 amino acid cytosolic protein (Kamitani *et al.*, 1993). This protein was found to interact with PIG-A by the co-purification of tagged PIG-A and PIG-H that were coexpressed (Watanabe *et al.*, 1996). Association of the protein with the cytoplasmic face of the ER was demonstrated by copurification with microsomes and sensitivity to proteinase K treatment. The role of PIG-H is not clear because it has no significant homology to any known protein.

A third GlcNAc-defective yeast mutant, *gpi1*, was identified (Leidich *et al.*, 1994) and complementation of this mutant by a *S. cerevisiae* genomic library led to the isolation of the *GPI1* gene (Leidich and Orlean, 1996). The sequence predicted a 609 amino acid, membrane protein with 5 potential transmembrane domains and, as with Gpi2p and Gpi3p, the protein lacks an ER-signal peptide. Disruption of this gene was found to block [^{3}H]inositol incorporation into proteins at non-permissive temperatures (37°C) but that at permissive temperature, incorporation was approximately 40% that of wild type (Leidich and Orlean, 1996). Overexpression of

GPI2 in the *gpi1* mutant partially rescued the temperature-sensitive phenotype suggesting that excess Gpi2p was, in some way, partially fulfilling the role of Gpi1p (Leidich *et al.*, 1995). More recently, the human and mouse genes have been cloned (Watanabe *et al.*, 1998; Tiede *et al.*, 1998). Both the human and mouse proteins have 581 amino acids and share 24% identity with *S. cerevisiae* Gpi1p (and 85% with one another). Tiede and co-workers (1998) demonstrated that overexpression of the human gene in the yeast *gpi1* mutant allowed the cells to grow at 30°C but not 37°C and marginal GlcNAc transferase activity was detected. By contrast, overexpression of the mouse gene resulted in greater recovery of the temperature-sensitive phenotype, probably because mouse GPI1 has a higher level of sequence identity to yeast Gpi1p than human GPI1. 50% GlcNAc transferase activity was detected when the mouse gene was overexpressed.

Watanabe and co-workers investigated interactions of the human homologue of GPI1 with the other GlcNAc transferase components in a mammalian system (Watanabe *et al.*, 1998). It was earlier demonstrated that tagged versions of PIG-A and PIG-H, when coexpressed interacted with one another, albeit weakly (Watanabe *et al.*, 1996). GST-tagged hGPI1 was coexpressed with FLAG-tagged PIG-A and precipitated with anti-GST antibodies. PIG-A was coprecipitated with hGPI1. This was repeated for PIG-H and PIG-C and in each case, a direct interaction with GPI1 was observed. A direct association of PIG-C with either PIG-A or PIG-H could not be demonstrated. When combinations of GPI1, PIG-A and PIG-H and PIG-C (all GST-tagged) were coexpressed with FLAG-tagged PIG-C, all four proteins were co-precipitated with anti-FLAG antibodies. Further, they precipitated in approximately 1:1 ratios. Cells were transfected with GST-tagged versions of PIG-A, PIG-C, PIG-H and hGPI1

individually and complexes were collected on glutathione beads. All four tagged proteins yielded purified complexes with GlcNAc transferase activity. These results provided evidence that these proteins form the GlcNAc transferase complex. The role of GPI1 in the complex has become slightly clearer with the finding that disruption of mouse *GPI1* caused the PIG-A, PIG-C and PIG-H complex to become virtually undetectable (Hong *et al.*, 1999). PIG-A and PIG-H complexes were readily detected but complexes of PIG-C and PIG-H were not. Therefore, GPI1 appears to have a role in stabilising the complex by keeping PIG-C associated with PIG-A and PIG-H (Watanabe *et al.*, 1998).

It remains a possibility that there are other, as yet, unidentified components of the GlcNAc transferase complex, but the fact that this relatively simple step requires four proteins (three in yeast) seems to be a feature of much of the GPI anchor biosynthetic machinery.

These findings are in agreement with the previously predicted topology of the early stages of GPI-anchor biosynthesis (Vidugiriene and Menon, 1993). When permeabilised cells that had incorporated UDP-[^{3}H]GlcNAc were treated with PI specific-phospholipase C (a bacterial enzyme that cleaves the phosphate-lipid bond between inositol and the lipid component of GPI anchors), the radiolabelled GlcNAc-PI was cleaved. This shows that the newly synthesised GlcNAc-PI is located on the cytoplasmic face of the ER. A similar result was obtained with purified microsomes. It was concluded that the early steps of GPI-anchor biosynthesis occur on the cytoplasmic side of the ER membrane (Vidugiriene and Menon, 1993).

Although none of the genes involved in GPI-anchor biosynthesis have been identified in *Trypanosoma brucei* to date, it was in this species that the structure of a GPI anchor was first elucidated (Ferguson *et al.*, 1985; Ferguson *et al.*, 1988). Furthermore, it was with African trypanosomes that the biosynthetic pathway of GPI anchors was first solved (Mayor *et al.*, 1990a; Mayor *et al.*, 1990b; Masterson *et al.*, 1989). A cell-free system was used with radiolabelled substrates leading to the identification of several of the GPI intermediates. As with mammals and yeast, the first step of African trypanosome GPI biosynthesis is the addition of GlcNAc to PI (Doering *et al.*, 1989). This step has also been observed using a *Leishmania major* cell free system (Smith *et al.*, 1997).

1.4.1.2 The second step of GPI biosynthesis is the rapid deacetylation of GlcNAc-PI

The acetyl group attached to glucosamine is removed after its addition to PI, a step that is common to mammals, yeast and protozoa. Mammalian cell lines deficient in the deacetylation step have been described (Mohney *et al.*, 1994; Stevens *et al.*, 1996). Another mutant cell line (Chinese hamster ovary (CHO) cells), designated class L, was found to accumulate GlcNAc-PI but no GlcN-PI (Nakamura *et al.*, 1997). This phenotype was complemented by expression of a rat cDNA in the mutant CHO cells; [^{3}H]GlcNAc incorporation allowed detection of GlcN-PI synthesis. Furthermore, overexpression of this cDNA led to a higher degree of deacetylation than that observed in wild type cells (Nakamura *et al.*, 1997). The cDNA encodes a predicted 252 amino acid protein, designated PIG-L, with a hydrophobic domain at the N-terminus. This domain is retained by the mature

protein, which is localised to the cytoplasmic side of the ER membrane, as determined by sensitivity of the protein to proteinase K treatment of purified microsomes, indicating that the domain probably serves as an ER transmembrane region (Nakamura *et al.*, 1997). This is, once again, in agreement with the predicted topology of this step in the biosynthetic pathway (Vidugiriene and Menon, 1993). A 304 amino acid *S. cerevisiae* homologue, Gpi12p, has been identified based on sequence similarity, sharing 24% identity to rat PIG-L (Nakamura *et al.*, 1997; Watanabe *et al.*, 1999). Expression of this protein in the class L mutant restored the surface expression of the GPI-anchored protein CD59, demonstrating the inter-species conservation of GPI-anchor biosynthesis. A DNA fragment of the human homologue had been identified by sequencing of expressed sequence tags (ESTs), leading to the production of a full length ORF by PCR (Watanabe *et al.*, 1999). This was also demonstrated to be able to rescue the class L mutant phenotype, restoring CD59 surface expression. The ORF predicted a 252 amino acid protein with 77% identity to rat PIG-L (Watanabe *et al.*, 1999). *In vitro* GlcNAc-PI de-N-acetylase activity was shown for both Gpi12p and hPIG-L, confirming complementation. Disruption of *S. cerevisiae GPI12* is lethal.

Deacetylation of GlcNAc-PI also occurs in *T. brucei* (Doering *et al.*, 1989) and *L. major* (Smith *et al.*, 1997). The *T. brucei* homologue of PIG-L/GPI12 was partially purified by liquid chromatography (Milne *et al.*, 1994). Differences between the substrate specificity of the human and *T. brucei* enzymes have been determined suggesting that this enzyme could be exploited as an anti-trypanosomal drug target (Sharma *et al.*, 1999b).

There is no evidence for a physical interaction between PIG-L and the complex containing PIG-A, PIG-C, PIG-H and GPI1 suggesting that the GlcNAc-PI de-N-acetylase machinery of mammals is distinct from the GlcNAc transferase complex catalysing the first step of GPI biosynthesis (Watanabe *et al.*, 1998). At the time of writing, the PIG-L protein is the only protein identified as being involved in the second step of GPI biosynthesis. If the de-N-acetylase is a complex then the other components have yet to be identified.

1.4.1.3 Inositol acylation is the third step of mammalian and yeast GPI biosynthesis but is not obligatory in trypanosomes

The next step in mammalian and yeast biosynthesis involves the acylation of the GlcN-PI precursor. The addition of palmitic acid (palmitoylation) to the inositol component of the GPI-anchor has been reported to precede the addition of the mannose residues in both yeast and mammals (Costello and Orlean, 1992; Urakaze *et al.*, 1992). Blockage of mannosylation, the fourth step of biosynthesis, leads to a buildup of GlcN-acyl-PI, the inositol-acylated form of GlcN-PI. The mannosylation of non-acylated GlcN-PI could be observed *in vitro* in both yeast and human but at lower efficiency than the acylated precursor (Doerrler *et al.*, 1996). Deacylation of the anchors is observed after addition to protein (Chen *et al.*, 1998a). The GlcN-PI acyl transferase has yet to be identified.

Palmitoylation also occurs in *T. brucei* with the VSG anchors (Krakow *et al.*, 1989; Mayor *et al.*, 1990b) eventually yielding a completed precursor designated glycolipid C. Glycolipid A is the non-acylated form of glycolipid C. The first major divergence

between the biosynthetic pathways of mammals/yeast and trypanosomes is evident at this stage. Where the acylation of GlcN-PI is necessary for biosynthesis to proceed in mammals and yeast, it does not appear to be a requirement in *T. brucei* (Guther *et al.*, 1994). Rather, the acylated and non-acylated forms appear to be in dynamic equilibrium with one another. The acylation/deacylation of the GPI precursor appears also to be stage regulated (Field *et al.*, 1991; Field *et al.*, 1992). The major GPI-linked glycoprotein of insect stage *T. brucei* is procyclic acidic repetitive protein (PARP) and the mature GPI-anchor attached to PARP is resistant to PI-PLC indicating inositol acylation (Field *et al.*, 1991). By contrast the VSG anchor is PI-PLC-sensitive and lacks such acylation of the mature GPI anchor (Mayor *et al.*, 1990a; Mayor *et al.*, 1990b). It has been proposed that inositol acylation and deacylation are important for later steps of GPI biosynthesis, namely ethanolamine addition and remodelling of the GPI-lipids respectively (Guther and Ferguson, 1995). Inhibition of the, as yet, unidentified inositol deacylase with diisopropylfluorophosphate (DFP) leads to a buildup of glycolipid C`, an immature precursor to the completed GPI that has yet to complete fatty acid remodelling (see below). The inositol acyltransferase, which is sensitive to phenylmethylsulphonyl fluoride (PMSF) prevents addition of ethanolamine but does not affect fatty acid remodelling (Guther *et al.*, 1994).

To date, inositol acylated precursors have not been detected in *Leishmania* (Smith *et al.*, 1997).

1.4.1.4 The sequential addition of three mannose residues encompasses the fourth, fifth and sixth steps of GPI anchor biosynthesis

The mannose donor to GPI biosynthesis, dolichol phosphate mannose or Dol-*P*-man is produced by the transfer of mannose from GDP-Man to Dol-*P* by the enzyme dolichol phosphate mannose synthase or DPMS (Orlean, 1990). *S. cerevisiae DPMS1* has been cloned (Orlean *et al.*, 1988) and a human homologue has also been identified (Colussi *et al.*, 1997) although there is a fundamental difference between the homologues; yeast Dpms1p possesses a carboxyl terminal hydrophobic domain that is absent in human DPMS1.

DPMS has also been isolated and characterised from *T. brucei* (Mazhari-Tabrizi *et al.*, 1996) and *L. mexicana* (Ilgoutz *et al.*, 1999a). As with the *S. cerevisiae* protein, both TBDPMS and LMDPMS possess a carboxyl-terminal hydrophobic domain. The *T. brucei* gene was cloned by complementation of a yeast temperature-sensitive mutant, defective in the endogenous DPMS and TBDPMS shares nearly 50% identity with the yeast protein demonstrating the evolutionary conservation of the process (Mazhari-Tabrizi *et al.*, 1996).

The glycosyltransferases responsible for the addition of the first and second mannose residues have not been identified. A candidate protein involved in the transfer of the third mannose residue, however, has been identified. Human *PIG-B* was cloned by complementation of class B mutant (Takahashi *et al.*, 1996). This mutant accumulates a GPI precursor which contains two of the three mannose residues and expression of the PIG-B cDNA allows transfer of the third mannose and the remaining steps in biosynthesis to occur. PIG-B is a 554 amino acid ER protein. Proteinase K treatment revealed that most of the protein (approximately 85%) is contained within the lumen of the ER but that a relatively large amino-terminal

cytoplasmic domain (approximately 60 amino acids) is also present. This domain is not necessary for PIG-B activity. Interestingly, the protein appears to lack any ER-signal peptide and it is not clear how the protein is localised to the ER. It is not known if PIG-B is the GPI mannosyl transferase 3 (GMT3) itself, or if it is a subunit of a GMT3 complex. The yeast homologue has been cloned and was designated *GPI10* (Sutterlin *et al.*, 1998). Gpi10p and PIG-B are functional homologues since *PIG-B* was able to rescue the lethal effect of *GPI10* disruption (Sutterlin *et al.*, 1998). The yeast protein, which is highly hydrophobic has homology to another yeast protein, ALG9 that is required for addition of mannose in α1-2 linkage to N-glycans (Canivenc-Gansel *et al.*, 1998).

The orientation of PIG-B would suggest that at some stage between the deacetylation of GlcNAc-PI and the addition of the second mannose, the GPI precursors are flipped to the lumenal face of the ER membrane in yeast and mammals. This is in contrast with findings in *T. brucei* where it was reported that it is probably a later GPI-intermediate that is translocated across the ER membrane (Vidugiriene and Menon, 1994). Cloning of a *T. brucei PIG-B* homologue would help to confirm or dispute these findings.

Both the trypanosome and human GMT3 appear to be inhibited by mannosamine (2-amino-2-deoxy D-mannose (Lisanti *et al.*, 1991). The mannosamine (ManN) is utilised in GPI biosynthesis to produce ManN-Man-GlcN-PI in trypanosomes (Ralton *et al.*, 1993). This confirms that blockage of the GPI biosynthetic pathway occurs after addition of the first two mannose residues and, since the third mannose residue

is attached via an α1-2 linkage, replacement of the hydroxyl on position 2 of ManN with an amine group means that this bond cannot be formed.

1.4.1.5 The addition of ethanolamine phosphate to Manα1-2Manα1-6Manα1-4GlcN-PI

The final mammalian GPI precursor contains at least one ethanolamine phosphate group (Kamitani *et al.*, 1992). Class F mutant cells are defective in the attachment of ethanolamine to the third mannose, however, ethanolamine is still incorporated into the other mannoses of the GPI precursor (Hirose *et al.*, 1992). One protein has been identified as being involved in ethanolamine phosphate addition, PIG-F (Inoue *et al.*, 1993). The addition of extra ethanolamine groups to the other mannosyl groups is not confined to mammals; the yeast mutant *gpi10-1* demonstrates a buildup of GPI precursor with the first two mannosyl groups added and has an ethanolamine phosphate group attached to the first mannose (Canivenc-Gansel *et al.*, 1998).

In *T. brucei* it has been shown that donation of the terminal ethanolamine phosphate group is made by phosphatidylethanolamine (Menon *et al.*, 1993). Another difference between GPI biosynthesis of mammals/yeast and trypanosomes is that the addition of extra ethanolamine phosphate groups to the other mannose residues does not occur in trypanosomes. The addition of the ethanolamine group appears to be enhanced by the acylation of inositol (Guther and Ferguson, 1995). The GPI precursors have to undergo one more step in biosynthesis at which stage they are ready for transfer to specific nascent polypeptides.

1.4.1.6 Fatty acid remodelling of GPI anchors represents the final step in the biosynthetic pathway

This step in biosynthesis was first observed in African trypanosomes where the existing lipids of the PI component of the VSG anchor are replaced sequentially with myristic acid (Masterson *et al.*, 1990). Myristate is donated by myristoyl-lyso-phosphatidylcholine (M-LPC) via myristoylCoA in an ATP-dependent reaction (Werbovetz and Englund, 1996). The ATP is thought to be required for the hydrolytic cleavage of myristate from M-LPC prior to myristoylation of the GPI anchor. Yeast GPI anchors also undergo such fatty acid remodelling steps (Sipos *et al.*, 1997). Incorporation of ceramide and other long chain fatty acids can occur in the ER and further ceramide remodelling can occur later in the secretory pathway, in the Golgi apparatus. Remodelling within the ER occurs in the absence of protein synthesis where no anchor incorporation into protein takes place (Reggiori *et al.*, 1997).

The GPI anchors can be further modified by the addition of glycan side chains after transfer of the anchor to nascent polypeptide although this is thought to begin in the ER and is continued in the Golgi complex (McConville and Ferguson, 1993).

1.4.2 The cell surface molecules of *Leishmania*

There are three main types of molecule present on the cell surfaces of *Leishmania* species (Figure 1.3):

- Cell surface proteins that are covalently attached at their carboxyl terminus to glycosylphosphatidylinositol (GPI) membrane anchors.

- Lipophosphoglycans (LPGs) that are complex carbohydrates attached to the cell surface via GPI anchors
- Glycoinositolphospholipids (GIPLs) that are essentially free GPI structures.

1.4.2.1 GPI anchored proteins

The major surface metalloproteinase of *Leishmania* promastigotes, GP63, is attached to the cell surface by a GPI anchor (Medina-Acosta *et al.*, 1989; Schneider *et al.*, 1990). This is the most abundant GPI-linked protein of this organism, with around $5x10^5$ copies per cell. Possible roles of this protein are discussed in Section 1.3.2.

A second GPI-linked promastigote surface protein is promastigote surface antigen 2 (PSA2) also referred to as GP46 (Lohman *et al.*, 1990; Murray *et al.*, 1989). Three antigenically similar polypeptides of 96, 80 and 50kDa represent the PSA2 polymorphic family (Handman *et al.*, 1995; Symons *et al.*, 1994). The roles of these proteins are as yet unknown.

Expression of these GPI-linked proteins on the surface of amastigotes is strongly downregulated (Bahr *et al.*, 1993; Handman *et al.*, 1995). No GPI anchored proteins have been positively identified to date in amastigotes.

1.4.2.2 Lipophosphoglycans (LPGs)

Leishmania species possess other types of GPI structures that are not protein-linked (McConville and Ferguson, 1993). Lipophosphoglycans are phosphoglycans attached to a novel GPI anchor that is structurally quite distinct from protein-linked GPIs (Figure 1.3). The anchor contains the common GPI motif **Manα1-4GlcN-PI** but the

second mannose is attached via an α1-3 linkage, rather than the α1-6 bond seen in protein GPIs (Figure 1.3) (McConville and Ferguson, 1993). Instead of the third mannosyl residue, present in the protein GPI anchor, a galactofuranose residue is attached to the second mannose followed by two galactose residues and the phosphoglycan component of the LPG. The phosphoglycan component consists of variable numbers of disaccharide repeats

The LPGs of *Leishmania* are important for the survival of this organism in both the insect and mammalian host (Beverley and Turco, 1998). They are present in high copy number, (approximately $6x10^6$ molecules per cell) in promastigotes but, like GPI-anchored proteins, are downregulated in amastigotes of *L. donovani* (McConville and Blackwell, 1991). This stage-specific expression is also observed in *L. mexicana* and *L. major* (Bahr *et al.*, 1993) and the promastigote LPG glycocalyx protects the parasite from the harsh hydrolytic environment of the sandfly midgut (McConville, 1997).

LPGs have been implicated in survival and development in the sandfly (Pimenta *et al.*, 1992; Pimenta *et al.*, 1994). In *L. donovani* the LPGs in metacyclics are elongated compared with procyclic promastigotes, however this would not prevent binding of the LPG to lectin receptors on the midgut epithelial cells since the sugars on the end of the phosphoglycan units that interact with receptors (capping sugars) are the same (Butcher *et al.*, 1996). Rather, folding and clustering of the ends of the metacyclic phosphoglycan units blocks access of the lectin receptors to the, now buried, capping sugars (Sacks *et al.*, 1991). This LPG elongation is also observed in *L. major* but it is thought to be the nature of side chains along the length of the phosphoglycan repeats

which differ between procyclic promastigotes and metacyclic promastigotes (McConville *et al.*, 1992; McConville, 1997). It is thought that these differences in LPG structure and sandfly epithelial receptor specificity are fundamental in determining which species of sandfly a particular species of *Leishmania* can infect (Pimenta *et al.*, 1994). Thus the geographical distribution of *Leishmania* species is reliant on the range of the insect vector.

LPGs are also thought to be involved in the avoidance of complement mediated lysis by promastigotes within the mammalian host (McConville, 1997). Interaction with components of the complement pathway by metacyclic promastigotes leads to opsonization and preferential uptake by host macrophages via complement receptors on the surface of macrophages (McConville and Ferguson, 1993; Mauel, 1996; Puentes *et al.*, 1990). There is also evidence to suggest that LPGs may interact directly with macrophages to facilitate entry (Kelleher *et al.*, 1992). Upon phagocytosis, the accompanying macrophage respiratory (oxidative) burst is depressed by *Leishmania* promastigotes. This is facilitated by LPG, which is a potent inhibitor of protein kinase C activity (McNeely and Turco, 1987), an enzyme correlating with macrophage oxidative burst (Mauel, 1996).

1.4.2.3 Glycoinositol phospholipids (GIPLs)

GIPLs are free GPI structures and represent the most abundant cell surface molecules of *Leishmania* promastigotes and amastigotes with approximately 10^7 molecules per cell. The survival of amastigotes in the harsh environment of the lysosome-derived parasitophorous vacuole is thought to be due to the network of hydrogen bonds existing between the GIPLs acting as a protective glycocalyx (Winter *et al.*, 1994;

McConville, 1997; McConville and Blackwell, 1991). Recent work by Mensa-Wilmot and co-workers identified two putative, ethanolamine phosphate-containing GIPLs that were significantly less abundant when exogenous GPI-PLC was expressed in *L. amazonensis* (Mensa-Wilmot *et al.*, 1999). Growth of these cells in host macrophages (*in vitro*) was arrested and hamster infections produced smaller lesions and fewer parasites than wild type cells suggesting that free GPIs are important to amastigote viability. This is in agreement with the findings of Ilgoutz, McConville and co-workers (1999a) who demonstrated (albeit indirectly) that GIPLs were essential to survival of *L. mexicana*.

1.4.3 Biosynthesis of protein GPI, LPGs and GIPLs in *Leishmania*

The biosynthesis of GPI anchored proteins in *Leishmania* appears to parallel that of African trypanosomes, considered in previous sections. Many of the early intermediates in protein GPI biosynthesis have been detected in a *L. major* cell free system (Smith *et al.*, 1997).

Because protein GPI anchors, LPGs and GIPLs share the common motif Manα1-4GlcN-PI, it might be predicted that their biosynthesis shares some of the early biosynthetic steps (see Section 1.4.1 for details of these steps for protein GPI biosynthesis). Whilst protein and LPG anchors do appear to share some early steps, the nature of the lipids in PI component of the early GIPL intermediates, differs from those in protein and LPG anchors, indicating that the GIPL PI is derived from a distinct PI pool in *L. mexicana* (Ralton and McConville, 1998). Some GIPLs have an ethanolamine phosphate group added to the second mannosyl, a modification that appears to be unique to GIPL biosynthesis, providing further evidence of the discrete

nature of GIPL biosynthesis from that of protein and LPG GPIs. This ethanolamine transfer is more commonly found in amastigote GIPLs (Winter *et al.*, 1994).

Protein and LPG anchors do seem to share the same common intermediates during the early stages of GPI biosynthesis in *L. mexicana* promastigotes, up to and including the addition of the first mannosyl group to GlcN-PI (Ralton and McConville, 1998). The linkage between the second and the first mannosyl groups of protein GPI anchors is α1-6 compared with α1-3 in LPGs, suggesting that the divergence between the biosynthetic pathways occurs at this stage (Figure 1.3). The remainder of LPG biosynthesis requires a series of glycosyltransferases for the addition of the remaining sugars (galactofuranose, galactose, and glucose phosphate side chain) and probably occurs in the Golgi, rather than the ER (Ha *et al.*, 1996). The fatty acid remodelling of protein GPIs and LPGs is also somewhat different in that the end product of LPG remodelling is an unusual *lyso* species, lacking an acyl chain (as seen in Figure 1.3 and (Ralton and McConville, 1998)). There is also evidence to suggest that the GPI biosynthetic enzymes are located within a subcompartment (the dolichol phosphate mannose synthase tubule) of the ER in *L. mexicana*, associated with DPMS (Ilgoutz *et al.*, 1999b). The latter enzymes in protein GPI anchor biosynthesis (and the type I GIPLs), including the mannosyltransferases and the ethanolamine phosphotransferase as well as the fatty acid remodelling machinery co-localises with DPMS in this tubule, which appears to associate with both the Golgi and the bulk ER. The LPG anchor precursors are thought to complete assembly in the Golgi where the galactofuranosyltransferase is located, as disruption of the DPMS tubule association with the Golgi leads to a

buildup of a LPG precursor (Ilgoutz *et al.*, 1999b). This suggests that there is transfer of material between the DPMS tubule and the Golgi.

1.4.4 The attachment of protein GPI anchors to nascent polypeptides

The final step in protein GPI anchor biosynthesis is the attachment of the complete anchor to proteins that are destined to undergo this post-translational modification. But how does this take place and why do particular proteins undergo anchor attachment while others do not? This section describes the nature of the GPI signal sequence and the evidence to support the idea that the addition of the anchor is mediated by a transamidation mechanism, catalysed by a GPI:protein transamidase.

1.4.4.1 The GPI anchor signal sequence of nascent polypeptides

Proteins that are destined to receive a GPI anchor are synthesised as pre-pro-proteins (Figure 1.4). At the amino-terminus, there is a hydrophobic domain of approximately 30 amino acids that is an ER-signal sequence removed by signal peptidase on entry to the ER. There is another hydrophobic domain at the C-terminus of such proteins. Deletion of this 17 amino acid hydrophobic domain from decay accelerating factor (DAF; a human membrane glycoprotein) leads to its secretion (Caras *et al.*, 1989) indicating that this domain has a role in keeping the protein associated with the ER membrane until processed.

Human placental alkaline phosphatase (PLAP), which is a 530 amino acid protein, is processed to a membrane protein of 484 amino acids containing ethanolamine, glucosamine, mannose, inositol and the lipids palmitate and stearate (Ogata *et al.*, 1988). The phosphatidylinositol anchor is attached to an aspartic acid residue at

position 484 indicating that a 29-amino acid peptide is cleaved from the nascent enzyme (Micanovic *et al.*, 1988). A cell-free system involving the use of site specific antibodies that recognised three different epitopes within the nascent PLAP (Bailey *et al.*, 1988) was developed to investigate the processing of the immature protein (Bailey *et al.*, 1989). The assay was based on the addition of PLAP mRNA to rabbit reticulocyte lysate, which contains all the necessary machinery for mRNA translation (except endogenous methionine) plus [^{35}S]methionine, in the presence and absence of microsomes. Proteinase treatment of the microsomes demonstrated that PLAP was sequestered into the microsome lumen. The use of the site-specific antibodies demonstrated that nascent PLAP is processed at both the amino and carboxyl termini and that this processing was almost certainly specific, rather than caused by random proteolysis events (Bailey *et al.*, 1989).

It was evident that there was more to the GPI signal than the hydrophobic domain and the region 10-20 amino acids N-terminal to the hydrophobic domain is also important (Caras *et al.*, 1989). A 28 amino acid oligopeptide is removed from DAF, very similar in size to the removed PLAP signal, confirming the importance of the 10-20 amino acid domain N-terminal to the hydrophobic domain, which contains the cleavage/attachment site (Moran *et al.*, 1991). The cleavage/attachment site was designated ω (omega). Transfer of the 37 C-terminal amino acids of DAF (which included the hydrophobic domain and the 20 amino acids N-terminal of it) to the carboxyl end of proteins such as the herpes glycoprotein gD-1 (Caras *et al.*, 1987) or the normally secreted human growth hormone (Caras *et al.*, 1989) resulted in both proteins being directed to the plasma membrane. Anchor attachment was confirmed in both cases by release of the proteins by PI-PLC and incorporation of

[^{3}H]ethanolamine. The residue of anchor attachment to DAF was identified as Ser319 and this residue was replaced in the hGH-DAF fusion protein by all of the remaining possible amino acids by site directed mutagenesis. This demonstrated that particular amino acids could serve as acceptors for GPI anchors, namely alanine, aspartate, asparagine, glycine and serine; generally amino acids with small side chains, although cysteine can receive a GPI anchor when Asp484 of PLAP is replaced (Micanovic *et al.*, 1990). Replacement with other amino acids, including cysteine, blocked GPI anchor addition and cell surface transport of hGH-DAF.

The ω site was investigated by the conversion of a non-functional GPI signal, which was added to hGH-DAF, to a functional one (Moran and Caras, 1991a). The region N-terminal to the hydrophobic domain of the DAF signal sequence is serine/threonine-rich. Replacement of this region in the hGH-DAF fusion protein, with a serine/threonine-rich domain (involved in O-glycosylation) from an unrelated protein was insufficient to direct GPI anchor attachment (Moran and Caras, 1991a). Mutation of a valine residue at the equivalent position to Ser319 of the "natural" hGH-DAF GPI signal is insufficient to facilitate GPI anchor attachment, however, if the residue immediately carboxyl to this is also mutated, to glycine, then GPI anchoring is restored. The second residue was mutated to a bulky amino acid, phenylalanine, and GPI anchoring was, once again, defective. Furthermore, the position of the cleavage/attachment site, relative to the hydrophobic domain is also important (Moran and Caras, 1991a). The serine/glycine residues were placed at different distances from the hydrophobic domain, 2, 4, 6, 9, 12 and 16 residues N-terminal to the start of the domain, but processing was only observed when the Ser/Gly was 12 amino acids away. This was confirmed by the absence of

[^{3}H]ethanolamine incorporation in all but the 12 residue-spacer protein. The data suggested that the structure of the hydrophobic and hydrophilic spacer domains of the signal were unimportant. Rather, they are there simply to provide hydrophobic and hydrophilic regions and the only region where sequence is important is the cleavage attachment site. Further evidence for this idea was provided by the production of functional, synthetic GPI signals by the fusion of sequence elements not normally involved in GPI anchor signalling (Moran and Caras, 1991b). In other words, hydrophilic spacers and hydrophobic regions not found in GPI signal sequences made up a functional GPI signal. Furthermore, as well as the residue to which the anchor is attached (ω), two residues carboxyl to this (ω + 1 and ω + 2) also exert an influence (Kodukula *et al.*, 1993).

A more detailed analysis of the ω site revealed that there are variations between species at this portion of the GPI signal sequence (Moran and Caras, 1994). When the VSG of *T. brucei* was expressed in mammalian (COS) cells, only very low levels of GPI-anchored VSG were detected on the cell surface. By contrast, fusion of the DAF GPI signal to VSG had the effect of high level surface expression of GPI-anchored VSG-DAF. The VSG GPI signal was fused to hGH (hGH-VSG) and was compared with hGH-DAF. Almost 20 fold lower GPI anchor attachment was observed with hGH-VSG than with hGH-DAF as determined by incorporation of [^{3}H]ethanolamine. These data indicated that VSG was capable of receiving a GPI anchor (it was correctly folded), but that the lack of anchoring could be attributed to incompatibility of the natural VSG GPI anchor attachment site (Moran and Caras, 1994). To further investigate this, the DAF ω site and spacer region was fused to the VSG hydrophobic domain and added to hGH resulting in GPI-anchored surface

expression. The reverse protein (VSG ω site and spacer region, with the DAF hydrophobic domain), was neither surface expressed nor GPI-anchored. Analysis of the ω site residues (ω, ω+1 and ω+2) indicated that the differences could be attributed to the bulkiness of the amino acids. It was proposed that the mammalian enzyme that catalyses the removal of the signal and addition of the GPI anchor could have trouble accommodating the bulky residues of the VSG ω site in its binding pocket. If this is true, then the trypanosome enzyme may have a larger binding pocket than its mammalian counterpart, meaning that this enzyme could be a potential target for anti-parasite chemotherapy (Moran and Caras, 1994).

1.4.4.2 Evidence that a GPI:protein transamidase catalyses GPI signal cleavage and anchor addition

The removal of the GPI signal sequence and its replacement with the preformed GPI-anchor is required for the correct expression of this type of cell surface protein. Removal of the signal allows the newly exposed carboxyl residue to be attached to the amino group of the ethanolamine portion of a GPI anchor via a peptide bond. It has been predicted that, the simplest model of anchor attachment is one that is catalysed by a GPI:protein transamidase (Ferguson and Williams, 1988). This could be brought about by a transamidation mechanism where the peptide bond between the ω residue and the GPI signal is broken and a new peptide bond is formed simultaneously between the protein and the GPI anchor. Alternatively, more than one enzyme could be involved: a proteinase to cleave the GPI signal and a transferase to form a new peptide bond between the protein and the anchor (Ferguson and Williams,

1988). One important consideration is that transamidation would not require energy whereas proteolytic cleavage followed by transferase activity would.

The method described above for the demonstration of GPI anchor attachment to PLAP was modified to include a simpler substrate based on PLAP (Kodukula *et al.*, 1991) for the study of mammalian transamidase activity. The engineered form of PLAP, preprominiPLAP contains both the ER signal peptide and the GPI signal but lacks glycosylation sites, and the active site. The smaller protein size means that small variations in protein size, such as those observed during processing, will be detected. PreprominiPLAP mRNA was added to microsomes and processing was observed in this cotranslational assay. This modified protein was processed to prominiPLAP and then to miniPLAP in the cell-free system, where microsomes from CHO cells were added. Processing of a number of truncated forms of preprominiPLAP, where parts of the GPI signal were omitted revealed that the completely processed miniPLAP had lower mobility on SDS-PAGE than the truncated forms (Kodukula *et al.*, 1992). Triton X-114 extraction causes GPI-anchored proteins to remain in the detergent phase unless the GPI is removed with PI-PLC, whereupon the proteins associate with the aqueous phase. Mini-PLAP was investigated by this technique and was observed to enter the aqueous phase when treated with PI-PLC suggesting that miniPLAP had received a GPI anchor. A cell-free system for analysis of GPI anchor attachment in *T. brucei* was also developed at around this time (Mayor *et al.*, 1991). This involved analysis of GPI anchor attachment to endogenous VSG and provided direct evidence for GPI attachment, allowing the detection of radiolabelled-GPI incorporation into VSG. Furthermore, the process did not require ATP or GTP, which provides evidence that GPI addition

proceeds via a transamidation mechanism. This finding was apparently contradictory with the processing of preprominiPLAP in a further modified version of the cell free assay (Amthauer *et al.*, 1992). The lag between the incorporation and processing of newly synthesised preprominiPLAP to prominiPLAP and the GPI anchor addition that was observed (Kodukula *et al.*, 1992) was taken advantage of to allow simplification of the assay (Amthauer *et al.*, 1992). Experiments indicated that GTP and ATP enhanced processing to miniPLAP although the authors point out that the ATP and GTP are more likely to be required for other steps such as chaperone-mediated folding of the protein or its translocation to the transamidase site (Amthauer *et al.*, 1992). Further evidence for this came from the demonstration of the interaction of the molecular chaperone BiP with three GPI anchored proteins, including PLAP and the requirement for ATP (Amthauer *et al.*, 1993). Soluble components of the ER lumen, of which BiP is an example, were shown to be required for the processing of prominiPLAP to miniPLAP (Vidugiriene and Menon, 1995). It was demonstrated that if removal of soluble components took place before entry of preprominiPLAP to the ER, processing to prominiPLAP occurred but further processing did not. By contrast, removal of soluble components after processing to prominiPLAP did not alter processing to miniPLAP suggesting the possible involvement of molecular chaperones such as BiP.

An interesting phenomenon that arose from the use of this preprominiPLAP system was that not all of the processed miniPLAP contained a GPI anchor (Maxwell *et al.*, 1995a). It could be argued that this actually provides evidence for the presence of a separate proteolytic enzyme agreeing with the alternative two or more enzyme hypothesis previously mentioned (Ferguson and Williams, 1988). If this were the

case then membranes from GPI-deficient mutants would still be expected to produce free miniPLAP, which does not happen (Maxwell *et al.*, 1995a). It was predicted previously, however, that the GPI:protein transamidase could potentially release free protein as well as GPI-anchored protein in the absence of suitable endogenous GPI precursors (Bailey *et al.*, 1989). The ethanolamine component acts as a nucleophile, attacking an activated carbonyl on the protein at the ω site. For the production of free miniPLAP, the enzyme could utilise water as a replacement nucleophile for the GPI anchor such that it is essentially functioning as a proteinase. Thus, when other more powerful nucleophiles such as hydrazine and hydroxylamine were included in the cell free assay an increase in the amount of free miniPLAP was observed and demonstrated to be a direct conversion from prominiPLAP to free miniPLAP (Maxwell *et al.*, 1995b). It was not shown, however, that the miniPLAP detected was actually the hydrazide or hydroxamate of free miniPLAP as the amounts produced were too low for detection with the current assay systems. The fact that free miniPLAP levels were increased in the presence of these nucleophiles, in a heat-sensitive fashion, suggests that the reaction is catalysed by a transamidase. Furthermore, the ω site amino acid requirement for processing with hydrazine parallels that required for GPI anchor addition (Ramalingam *et al.*, 1996). This study determined that production of free miniPLAP using microsomes from GPI defective cell lines with hydrazine, was maximum at 30°C and pH 7.0-8.0.

Collectively, these data strongly suggest that GPI anchor addition is catalysed by a GPI:protein transamidase but all of the evidence is indirect. Identification of proteins that form part of the GPI:protein transamidase would provide more direct evidence for the transamidation mechanism.

1.4.4.3 Identification of components of the putative GPI:protein transamidase

As with many of the proteins involved in GPI anchor biosynthesis, two components of the transamidase were identified by complementation of mutant cells defective in GPI-anchored proteins.

A yeast mutant *end2*, originally thought to be defective in endocytosis, was found to be defective in the maturation of Gas1p, a major yeast GPI-linked, cell wall protein (Hamburger *et al.*, 1995). GPI anchor attachment was also severely affected and both of these phenotypes were reversed by expression of the gene, which was renamed *GAA1*. Gaa1p is an ER transmembrane glycoprotein of 68kDa (prior to glycosylation) with a large lumenal domain and as many as 5 or 6 transmembrane domains. Gaa1p shows no significant sequence homology to proteins of known function. The *gaa1* mutant synthesises a complete GPI anchor indicating that the mutation lies after biosynthesis but before attachment to nascent polypeptides carrying a GPI signal sequence. Mutation of the ω site of Gas1p to various amino acids had the effect of reducing the percentage of mature Gas1p, however overproduction of Gaa1p could partially overcome the reduced GPI-anchoring of the Gas1p mutants (Hamburger *et al.*, 1995). These findings are consistent with Gaa1p being either the GPI:protein transamidase or an essential component of a transamidase complex.

A second putative component of the transamidase complex was identified in yeast (Benghezal *et al.*, 1996). A yeast mutant (*gpi8*), like *gaa1*, demonstrates a buildup of a complete GPI anchor precursor. Expression of the *GPI8* gene from a plasmid

prevented this buildup suggesting complementation. *GPI8* predicts a 411 amino acid protein of 47kDa. The nascent protein contains two hydrophobic domains, at either terminus; the N-terminal domain is predicted to be an ER-signal peptide. There are three potential N-glycosylation sites which are utilised. Subcellular fractionation confirmed that GPI8 co-fractionates with a known ER protein (Wbp1p) and proteinase protection assay of yeast microsomes confirmed that GPI8 has a large ER lumenal domain, a single transmembrane domain and a short cytoplasmic tail (Benghezal *et al.*, 1996). This structure is reminiscent of GAA1 but unlike GAA1, GPI8 was found to have significant homology to a family of plant cysteine endopeptidases known as the legumains, which belong to the C13 family on the cysteine proteinase classification scheme (Barrett and Rawlings, 1996). A partial sequence for the human homologue of *GPI8* was identified (Benghezal *et al.*, 1996).

A previously identified human mutant cell line, class K (Mohney *et al.*, 1994) was found to accumulate complete GPI anchors but microsomes from this cell line, when used with the preprominiPLAP cell free assay, failed to produce GPI-anchored miniPLAP (Chen *et al.*, 1996). Furthermore, inclusion of the nucleophiles hydrazine and hydroxylamine failed to produce free miniPLAP indicating that the mutation probably lay in the GPI:protein transamidase complex. It was later found that the defect was in the human homologue of *GPI8* and the remainder of the human gene was cloned and characterised (Yu *et al.*, 1997). Expression of this gene in class K cells abolished buildup of complete GPI anchors. Microsomes from these cells were able to produce free miniPLAP with hydrazine as well as GPI-miniPLAP indicating that transamidase activity had been restored.

The observed similarity with proteolytic enzymes suggests that GPI8 is the transamidase itself. A modified cell-free assay was recently designed to analyse GPI anchor attachment in African trypanosomes and this provided the first direct evidence that the nucleophile hydrazine was indeed covalently incorporated at the C-terminal end of a VSG in place of a GPI anchor (Sharma *et al.*, 1999a). This was done by using biotinylated hydrazine and demonstrating that VSG released from membranes with hydrazine was precipitated with streptavidin. Furthermore the VSG-hydrazide was not susceptible to proteolytic cleavage by carboxypeptidases (exopeptidases that cleave at C-termini of proteins) unlike truncated VSG indicating that the carboxyl terminus of released VSG was blocked. This study revealed that the transamidation reaction was inhibited with sulphydryl alkylating reagents such as iodoacetamide suggesting that a thiol group is important for the activity of the transamidase (Sharma *et al.*, 1999a). Since GPI8 has homology to the legumains, which are cysteine proteinases, this data provides further evidence for the notion that the GPI:protein transamidase is a novel cysteine proteinase.

1.5 Aims of this study

The identification of components of the putative GPI:protein transamidase has paved the way for isolation and characterisation of homologous genes from parasitic protozoa. The extensive use of GPI anchors for membrane attachment of cell surface proteins in parasitic protozoa makes them model systems for analysing the GPI:protein transamidase.

The aims of this work were to:

- Isolate and sequence *GPI8* from *L. mexicana*
- Delete the gene to determine the role of GPI8
- Investigate the role of GPI anchored proteins in the life-cycle of *Leishmania*
- Determine if *GPI8* is present in other species of parasitic protozoa such as *Plasmodium* and *T. brucei*
- Assess if the GPI:protein transamidase is a potential anti-trypanosomatid drug target.

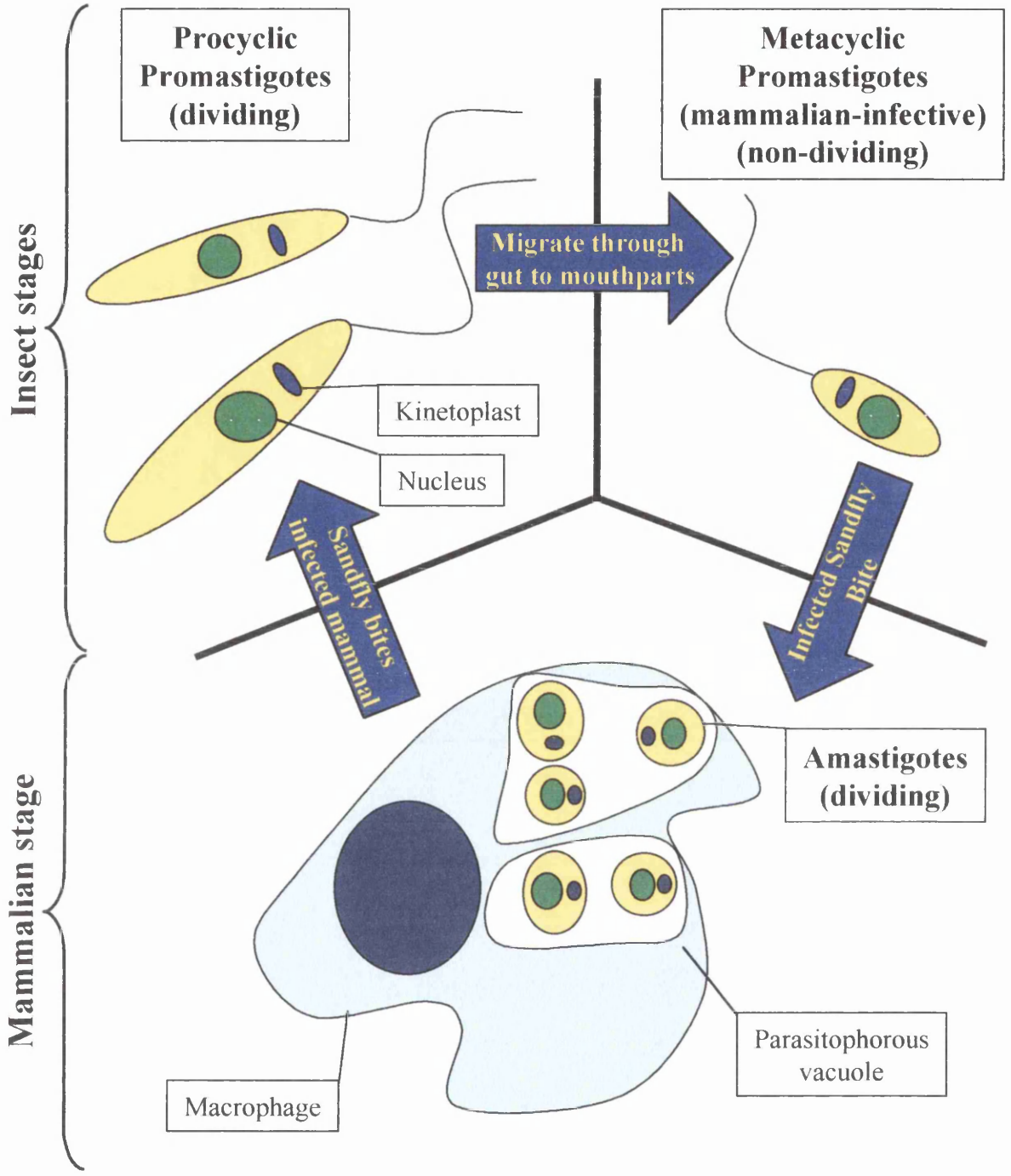
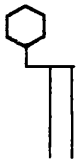


Figure 1.1: The life cycle of *Leishmania*

Figure 1.2: Biosynthesis of protein GPI anchors in mammals occurs on the ER membrane. [1] Addition of GlcNAc to PI is catalysed by a multiprotein complex and is rapidly deacetylated [2]. The anchor is then inositol acylated (palmitoylated) before addition of mannose residues; at some point the GPI precursor is flipped from the cytoplasmic face of the ER membrane to the lumen [3-5]. The third mannose residue is added by PIG-B [6] and ethanolamine is added, presumably by PIG-F (not shown) before fatty acid remodelling [7-8]. The complete anchor is transferred to nascent polypeptides carrying a GPI signal sequence on the carboxyl-terminal [9]. Adapted from Ferguson (1999).

KEY


Phosphatidylinositol (PI)

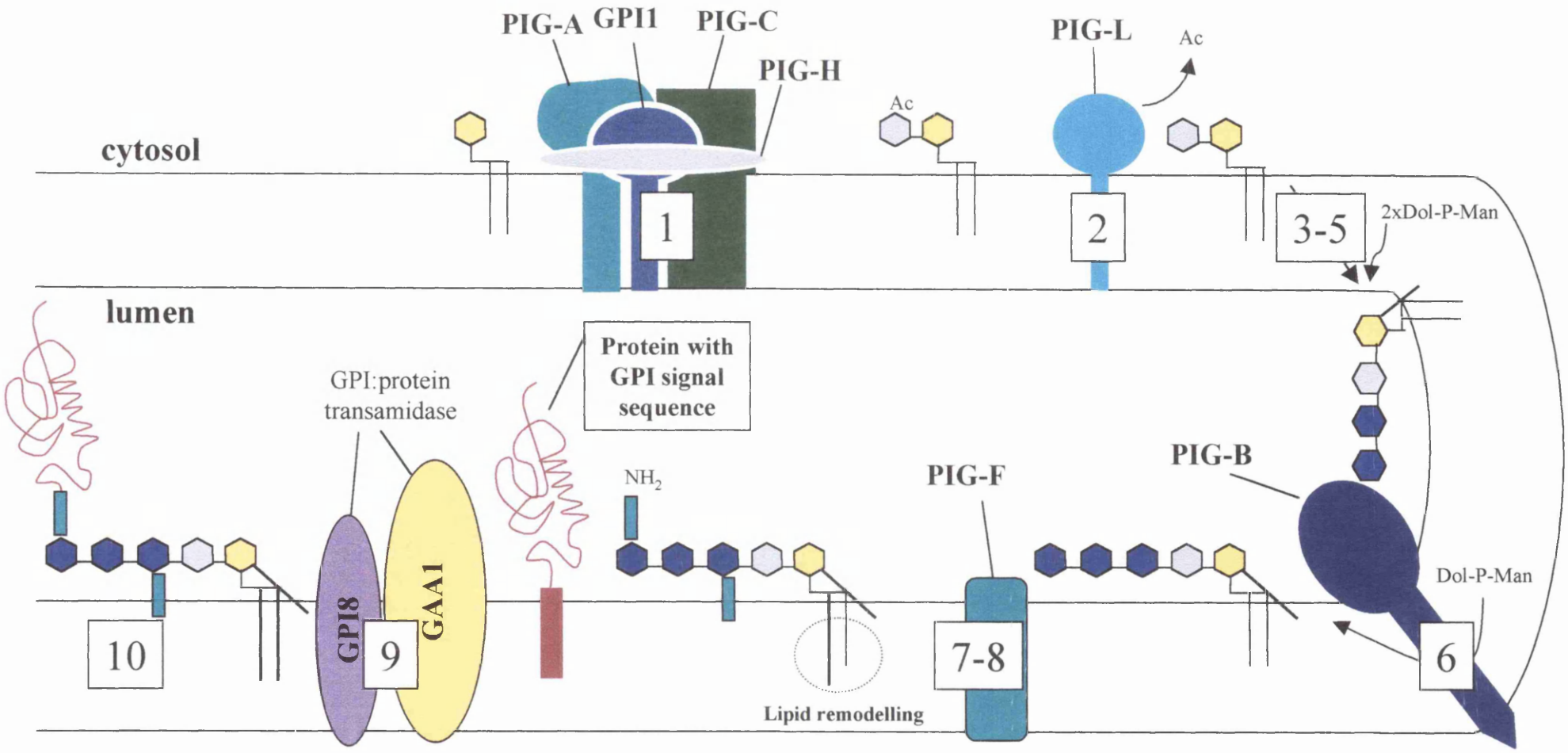

Palmitoyl group

N-acetylglucosamine


glucosamine


Mannose


Ethanolamine phosphate


PIG-A
GPI1
PIG-C
PIG-H
PIG-L
Ac
Ac
cytosol
1
2
3-5
2xDol-P-Man
lumen
Protein with GPI signal sequence
GPI:protein transamidase
NH2
PIG-F
PIG-B
GPI8
GAA1
10
9
7-8
6
Dol-P-Man
Lipid remodelling

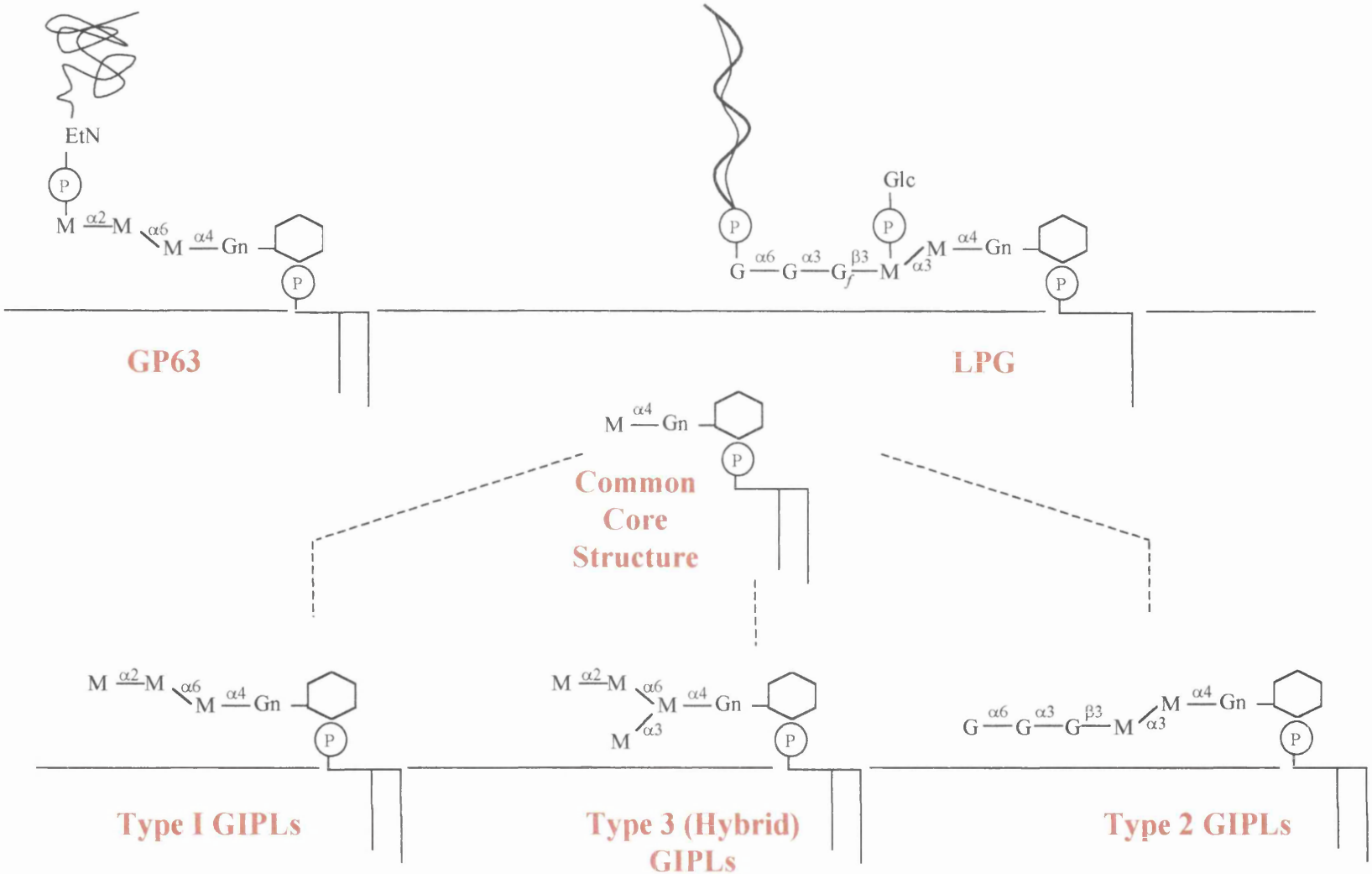


Figure 1.3: Schematic structures of GPI molecules in *Leishmania*. The common core component, Manα1-4GlcN-PI is also shown (centre). Adapted from McConville (1997).

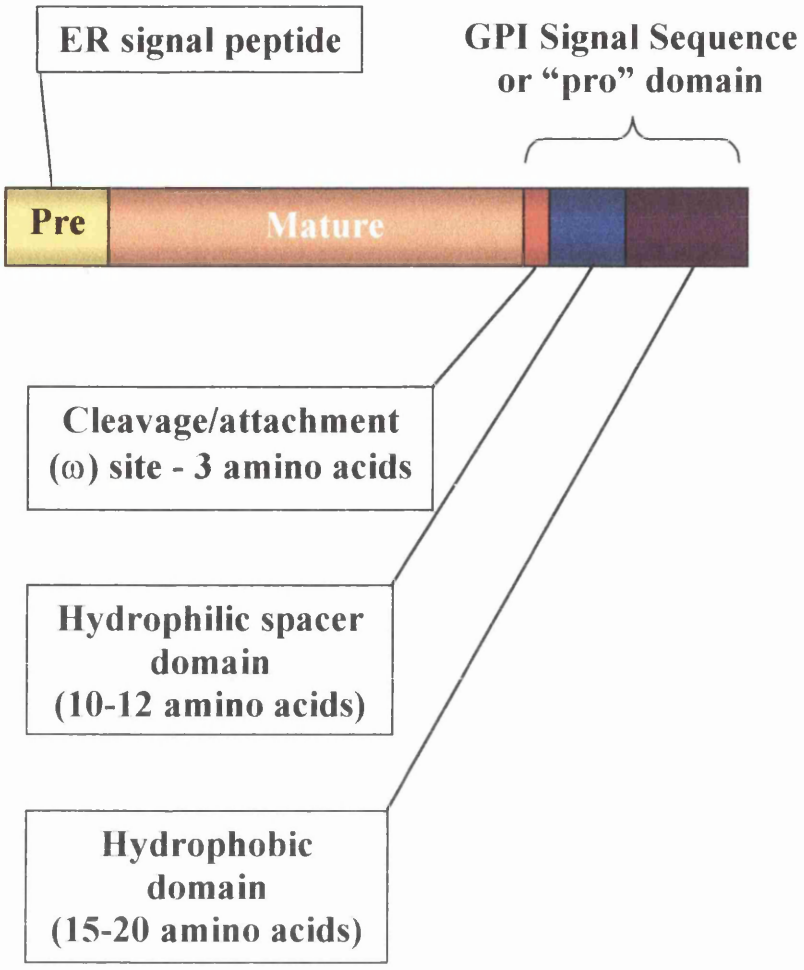


Figure 1.4: General structure of a nascent polypeptide prior to GPI anchor addition. Nascent preproproteins consist of three main domains: an N-terminal ER signal sequence (yellow), a mature domain (orange) and a GPI signal domain. Both of the terminal signal sequences are removed and the C-terminal domain is replaced with a GPI anchor. The three components of the GPI signal are indicated.

CHAPTER 2

MATERIALS AND METHODS

2.1 Molecular methods

2.1.1 Bacterial strains

Several strains of bacteria were used throughout this study and these are described in Table 2.1.

Table 2.1: Bacterial strains

Strain	Uses	Reference/manufacturer
XL1-Blue MRF`	Supercompetent cells used for transformations	Stratagene
M15[pRRP4]	Used for protein overexpression (Kanamycin resistant)	Qiagen
AB1899 F`	Parent strain for AB1899 F` GroEL/ES	A gift of Dr. John Hyde (UMIST)
AB1899 F` GroEL/ES	Used for co-expression of protein with chaperonins (Kanamycin resistant)	A gift of Dr. John Hyde (UMIST)
LE392	Used for uptake of λ phage in genomic library screening	(Sambrook *et al.*, 1989)

2.1.2 Long term storage of bacteria (glycerol stocks)

Cells were plated on LB-agar plates with appropriate antibiotics and incubated overnight at 37°C. A single colony was used to inoculate 5ml of LB broth and cells were grown overnight at 37°C. 0.5ml of the overnight culture was mixed with an equal volume of 2% peptone/40% glycerol and cells were stored at –70°C.

2.1.3 Bacterial culture methods

5-10ml of LB broth containing appropriate antibiotics (see Table 2.2) in a sterile glass test tube was inoculated with a single colony using a sterilised toothpick. Cultures were grown at 37°C in a rotary incubator (220rpm). Growth times were dependent on requirements. For example stationary phase cultures were required for plasmid preparations and cells were grown overnight. For small-scale plasmid preparations, the overnight cultures were used directly and for medium-scale plasmid preps (midi preps), 1ml of an overnight culture was used to inoculate 25ml of fresh LB broth. When cells were used for overexpression of exogenous protein, growth times were altered accordingly (see relevant results chapters).

2.1.4 Preparation of electrocompetent cells

A single colony of bacterial strain from an LB-agar plate was used to inoculate 10ml of LB broth containing appropriate antibiotics. The entire culture was added to 400ml of fresh LB-broth and incubated at 37°C (220rpm) in a rotary incubator until the OD (600nm) had reached 0.6. OD was measured in disposable plastic cuvettes (Sarstedt). Cells were harvested (4000 x g, 15 minutes, 4°C) and washed two times, each with 400ml of ice-cold sterile water. The cells were resuspended in 200ml of ice-cold sterile water and harvested and the pellet was washed with 10ml of ice-cold 10% glycerol (sterile). The cells were resuspended in 2ml of ice-cold, sterile 10% glycerol and divided into 40µl aliquots that were snap-frozen in a dry ice/ethanol bath. The electrocompetent cells were stored at –70°C.

2.1.5 Plasmid purification

Several different kits were used to purify plasmids depending on requirements. If plasmids were to be sequenced, Qiagen Tip-20 mini-preps were used according to manufacturer's instructions, as particularly high quality DNA was generally obtained with this kit. The Wizard™ Minipreps kit (Promega) was used if plasmids were to be used in restriction digests, perhaps for subsequent subcloning, as the method is quick and efficient producing good plasmid yields. Plasmid quality and yields were assessed by spectrophotometry. An absorbance reading of 1 at a wavelength of 260nm is equivalent to 40μg/ml of double stranded DNA (Sambrook *et al.*, 1989).

2.1.6 Polymerase chain reaction (PCR)

PCR was used in the amplification of DNA fragments for subsequent cloning and analysis or for the diagnosis of ligation by whole-cell PCR. A 20μl reaction contained 1.8μl of 11.1x PCR buffer 50-100ng of template DNA, 200ng of each oligonucleotide primer (sense and antisense) and 2U of Taq polymerase (Applied Biosystems). Taq polymerase lacks proofreading activity so this enzyme was replaced with VENT polymerase where high fidelity PCR was required. Reaction conditions varied and are described where appropriate. For cloning into T-vectors such as pGEM-T (Promega) and pTAG (R&D Systems), Taq-amplified fragments could be ligated directly, but adenine overhanging residues had to be added to VENT-amplified fragments by incubation for an additional 10 minutes at 72°C with 2U of Taq polymerase.

When screening transformed colonies for plasmids containing the insert of interest, a cocktail stick was inserted into a colony and cells were transferred directly into a 20μl

PCR reaction without template DNA, and containing 5% (v/v) dimethylsulphoxide (DMSO).

2.1.7 Restriction digests

Restriction enzymes were obtained from New England Biolabs (NEB) and used according to manufacturer's instructions with the buffers supplied. A typical plasmid digest was set up as follows: 0.5-1.0μg of plasmid DNA was digested in a 20μl reaction (containing appropriate enzyme buffer) for 1-3 hours at 37°C. Reactions were halted by the addition of a one tenth volume of DNA loading buffer at 1x (v/v) final concentration (DNA loading buffer was prepared as a 10 fold concentrated stock) followed by electrophoresis. Occasionally reactions would be extracted with phenol/chloroform and/or ethanol precipitated to separate restriction enzymes from amplified DNA (see Sambrook *et al.*, 1989).

2.1.8 DNA gel electrophoresis

Typically gels containing 0.7-1.2% agarose in TBE buffer were used for analysis of DNA fragments, but gels of up to 1.8% were used to separate smaller fragments, up to approximately 400bp. Samples were mixed with 10x DNA loading buffer to a final concentration of 1x loading buffer and were electrophoresed at 50-110V until the dye in the loading buffer had migrated approximately two thirds of the length of the gel. Generally, ethidium bromide was mixed with the molten gel prior to setting at a final concentration of approximately 0.3μg/ml. Where uniform staining of a gel was required, the gel was soaked for 20-30 minutes in TBE buffer containing 0.3μg/ml ethidium bromide before washing in distilled water. Gels were viewed under UV

illumination and images obtained using The Imager™ gel documentation system (Appligene).

2.1.9 Purification of DNA from agarose gels

Where purification of DNA fragments was required, agarose gels were prepared with TAE buffer rather than TBE. Following electrophoresis, the gel was visualised under low intensity UV light to minimise the risk of mutagenesis allowing identification of the band of interest. This band was cut from the gel using a scalpel blade. For small scale purification (approximately less than 1μg), the Qiaquick Gel Extraction Kit (Qiagen) was used. For larger scale preps, the gel fragment was cut into small fragments and place in Spin-X tubes (Costar). Gel fragments were incubated at –20°C for 15 minutes followed by 37°C for the same duration. Tubes were centrifuged at approximately 13000 x g in a microfuge for 15 minutes and the collected DNA solution was ethanol precipitated and resuspended in either sterile water or TE at the required concentration.

2.1.10 Phosphatase treatment of digested plasmids

Where plasmids were digested with a single restriction enzyme prior to ligation, it was treated with calf intestinal alkaline phosphatase (CIAP) to remove the 5` terminal phosphate. This minimised the chances of plasmid re-ligation rather than ligation with the added DNA fragment of interest.

The procedure used was as follows:

2μg of plasmid was digested in a 20μl reaction. 2U of CIAP (NEB) was added to the digest and incubated at 37°C for 15 minutes then 56°C for 15 minutes. An additional 1U of CIAP was added and incubated at 37°C and 56°C as before. Reaction was

diluted to 200µl with TE buffer and extracted once with phenol/chloroform (1:1) then chloroform/isoamyl alcohol (24:1). DNA was ethanol precipitated and resuspended in TE to appropriate concentration.

2.1.11 DNA ligation

Ligation of DNA fragments with commercially available T vectors was carried out according to manufacturer's protocol.

Ligation into other vectors was carried out as follows: Plasmid and insert were mixed in ratios ranging from 1:3 to 3:1 in a 10µl reaction containing 1x T4 DNA ligase buffer and 200U of T4 DNA Ligase (NEB). Reactions were incubated overnight at 16°C.

2.1.12 Transformation of competent bacteria

Commercially available XL1-Blue MRF' supercompetent cells (Stratagene) were transformed by heat shock according to manufacturers protocol.

Transformation of electrocompetent cells was performed as follows:

1µl of an overnight ligation mix was added to 40µl of electrocompetent cells kept on ice. Cells were mixed and placed in a pre-chilled 0.2cm electroporation cuvette (Biorad) and were electroporated using the Genepulser II apparatus (Biorad) under the following conditions: 25µF capacitance, 200 ohms resistance and 2.5kV. Cells were transferred to 1ml of prewarmed (37°C) SOC medium and allowed to recover for 1 hour at 37°C in a rotary incubator (200rpm). 50-200µl of transformed cells was plated

onto LB-agar plates containing appropriate antibiotics (generally 100μg/ml ampicillin). Plates were incubated overnight at 37°C.

2.1.13 DNA sequencing

DNA sequencing was carried out using the ABI-Prism™ cycle sequencing kit (Perkin-Elmer) according to manufacturer's instructions. Template DNA was added at 300-500ng in a 20μl reaction with 3.2 pmoles of an appropriate oligonucleotide primer. Samples were analysed by the University of Glasgow Molecular Biology Support Unit (MBSU) or by Oswell Sequencing, University of Southampton using ABI 373 or 377 automated DNA sequencers.

2.2 *Leishmania mexicana* and trypanosome methods

2.2.1 Cell lines

Wild type *L. mexicana* cell line used throughout this research was *Leishmania mexicana mexicana* (MNYC/BZ/62/M379). All newly produced cell lines were derived from this parent line. *Trypanosoma brucei* were grown and supplied by Dr. Pat Blundell (427 bloodstream forms) and Dr. Jaap van Hellemond (EATRO795 procyclic form).

2.2.2 Tissue culture

L. mexicana promastigotes were grown at 25°C in HOMEM medium (Berens *et al.*, 1976) supplemented with 10% heat inactivated foetal calf serum (FCS) (Labtech). Cells were inoculated into fresh medium at approximately 1 x 10^6 cells/ml and were passaged into fresh medium when cultures reached late-log or early stationary phase (approximately 1-2 x 10^7 cells/ml).

For maintenance of transgenic cell lines, appropriate antibiotics were added to growth medium at the following concentrations:

- Puromycin (Calbiochem)– 10μg/ml
- Nourseothricin (Hans Knoll Inst., Germany) – 25μg/ml
- Geneticin (G418, GibcoBRL) – 25-50μg/ml

Cells were diluted and fixed by adding 5μl of cells to 45μl of PBS and 50μl of 4% formaldehyde in PBS. Cells were loaded onto an improved Neubauer haemocytometer (Weber Scientific) and counted under microscopy.

Cells were generally harvested by centrifugation at 2000xg for 5 minutes followed by two washes in PBS both followed by centrifugation as before. Cell pellets were used immediately or stored at –70°C until required.

HOMEM plates were made by a 1:1 dilution of 2 x HOMEM with 2% agar containing appropriate antibiotics. Plates were allowed to dry in an air-flow hood for 15 minutes and were pre-warmed to 25°C for 30-60 minutes before cells were spread over the solid medium. Liquid was allowed to absorb into the plate for 5 minutes before the plates were sealed with parafilm and incubated at 25°C.

2.2.3 Stabilate preparation for long term storage of *L. mexicana* cell lines

0.5ml of a log phase culture was diluted with an equal volume of fresh medium containing 10% DMSO in a Cryotube™ vial (Nunc). Samples were stored overnight at –70°C and then stored under liquid nitrogen until required.

2.2.4 Transfection of *L. mexicana*

Transfections were carried out using the procedure of Coburn and co-workers (Coburn *et al.*, 1991). Cells were harvested and washed twice in electroporation buffer (EPB) before being resuspended at 1 x 10^8 cells/ml in EPB and kept on ice. 4 x 10^7 cells (400µl) were transferred to a pre-chilled 0.2cm electroporation cuvette (Biorad) and were electroporated with Genepulser II apparatus at 0.45kV (2.25kV/cm) and 500µF capacitance. Cells were allowed to recover for 10 minutes on ice before being transferred to fresh medium and incubated at 25°C overnight. Cells were then plated on HOMEM-agar plates containing appropriate antibiotics at the concentrations described in Section 2.2.2. Plates were allowed to absorb the transferred medium for 5 minutes before being sealed with Parafilm for growth at 25°C. Colonies were transferred to liquid medium plus antibiotics using a sterile, disposable Pasteur pipette and grown at 25°C.

2.3 DNA manipulation

2.3.1 Whole cell PCR

10µl of late-log/early stationary phase promastigotes was added to 100µl of sterile distilled water. 2µl of this hypotonic lysate was used as DNA template in PCR with appropriate oligonucleotide primers and Taq polymerase (see Section 2.1.6).

2.3.2 Isolation of genomic DNA from *L. mexicana*

A mini-prep procedure based on a recently published method (Medina-Acosta and Cross, 1993) was utilised throughout this project. Up to 2 x 10^8 cells were harvested

and resuspended in 400µl of TELT buffer and incubated at room temperature for 5 minutes. The TELT lysate was extracted carefully with an equal volume of water-saturated phenol/chloroform twice (followed, each time by centrifugation, 13000 x g, 5 minutes) and once with chloroform before ethanol precipitation with 2 x 70% ethanol washes. The precipitated DNA was allowed to dry for no more than five minutes and resuspended in TE containing 20µg/ml RNAse A. The DNA was incubated at 37°C for an hour and DNA quality was analysed by agarose gel electrophoresis for quality and spectrophotometry for concentration.

2.3.3 Southern blotting of DNA fragments

5-10µg of genomic DNA was digested overnight and fragments were electrophoresed (0.7% agarose TBE gel). The gel was incubated with 0.3µg/ml ethidium bromide and analysed under UV to allow a scale image of the gel to be obtained (a ruler was placed beside the gel to allow measurement of resulting autoradiograms to be equated with the position of the DNA markers). The gel was washed in distilled water for 10 minutes before incubation for 20 minutes in depurination solution. The gel was rinsed in distilled water and incubated for 30 minutes in denaturation solution. After another wash in distilled water, the gel was incubated for a further 30 minutes in neutralisation solution and washed for 10 minutes in 20 x SSC. DNA was transferred to Hybond-N nylon membrane (Amersham) overnight, by capillary transfer (Sambrook *et al.*, 1989).

2.3.4 Southern hybridisation with nucleic acid probes

The transferred DNA fragments on Hybond-N were covalently attached to the membrane by UV crosslinking (1200J UV Crosslinker, Stratagene) and baking at 80°C

for 10 minutes. Membrane was added to 5ml of Church-Gilbert solution and incubated at 65°C (for *L. mexicana* blots) or 55°C (for *T. brucei* blots). 30ng of the nucleic acid probe was radiolabelled using the Prime-It II Random Primer kit (Stratagene) according to manufacturer's instructions. Probes were purified on NucTrap® columns (Stratagene) and boiled for 5 minutes before incubation on ice for 2 minutes. Probe was added directly to the Church-Gilbert solution/Southern membrane and incubated overnight at the appropriate hybridisation temperature (65°C for *L. mexicana* blots or 55°C for *T. brucei* blots). Blots were washed with 3 x 15 minutes washes with 2 x SSC/0.1% SDS prewarmed to the appropriate temperature followed by 3 x 15 minute washes with similarly warmed 0.1 x SSC/0.1% SDS. Membranes were sealed in polythene and exposed to autoradiography film (Reflection™ film (NEN™ Life Sciences). Autorads were developed using a Kodak X-omat automated developer.

2.3.5 *L. mexicana* genomic library screen

2.3.5.1 Preparation of bacterial suspension

LE392 bacterial cells were prepared for library screening. A single LE392 colony was transferred from an LB-agar plate to 50ml LB-broth supplemented with 0.2% maltose and grown overnight. Cells were harvested (4000 x g, 10 minutes at room temperature) and resuspended in 20ml of 10mM $MgSO_4$.

2.3.5.2 Titration of the genomic library

A previously described λ Dash II *L. mexicana* genomic library (Mottram *et al.*, 1997) was titred: a ten fold and a hundred fold dilution of the library (each in a total volume of 100µl) was mixed with 100µl of LE392 bacterial suspension and allowed to adsorb at 37°C for 20 minutes. 90mm plates containing BBL bottom agar were prepared and

dried previously. BBL top agarose was melted, aliquoted 2.5ml each into two Bijou bottles and maintained in molten state at 45°C. The cells/phage were mixed with the molten top agarose and quickly swirled over bottom agar to produce a uniform layer, which quickly solidified. Plates were incubated at 37°C overnight and the resulting plaques in the bacterial lawn were counted.

2.3.5.3 Primary screening of the genomic library

Three 150mm Petri dishes were filled with 80ml of bottom agar and dried. 0.3ml of LE392 suspension was mixed with 72µl of the *L. mexicana* genomic library in each of three sterile Eppendorf tubes and were incubated at 37°C for 20 minutes. The cells/phage were added to 3 x 6.5ml of molten top agarose and rapidly added to the three bottom agar plates. Plates were allowed to set and were incubated overnight at 37°C. Hybond-N nylon membrane was cut to fit onto the surface of the plates (six membranes were prepared, two for each plate). Filters were allowed to settle onto the surface of each plate allowing the phage to attach to the filter by brief capillary flow. The first filter was added to each plate for 90 seconds during which time, at least three asymmetric holes were punched through the filter into the agar with a 21 gauge needle. The second filter was added for 2.5 - 3 minutes and the same holes were punched through it. Filters were floated on the surface of a container of denaturation solution for 1 minute and submerged for a further 5 minutes, before being transferred to neutralising solution for 5 minutes. Finally, filters were washed 5 minutes in 2 x SSC, dried for 20 minutes and baked for 10 minutes at 80°C. Filters were incubated in Church-Gilbert solution at 65°C for at least 2 hours before being probed overnight with radiolabelled nucleic acid. Filters were washed for an hour with several changes of 2 x SSC/0.1% SDS and exposed to autoradiography film. Positions of the punched holes

were marked on the autorad allowing the position of hybridising plaques to be pinpointed on the agar plates.

2.3.5.4 Secondary screening of genomic library

Plaques identified as hybridising were removed from the agar plate by removal of a "plug" of agar transferred to an Eppendorf tube containing 500µl SM (phage) buffer and incubated at 37°C for an hour to allow the phage to diffuse out of the agar. A plug is assumed to contain approximately 10^6 phage particles and an estimated 100 phage and 500 phage were each added to 100µl of LE392 suspension. The mixtures were incubated at 37°C for 20 minutes before plating onto 90mm bottom agar plates in 2.5ml of molten top agarose. Plates were incubated at 37°C overnight. Plaque lifts were carried out (membranes and plates were, again, needle-punctured) and the Hybond-N membranes were treated as before. Hybridisation with the nucleic acid probe was repeated and hybridising plaques were identified and plugged into SM buffer.

2.3.5.5 Large scale purification of phage DNA by the plate lysis method

The isolated phages were plated out and ten plaques were transferred to a single Eppendorf tube containing 650µl of SM buffer (incubated at 37°C for an hour). 500µl of phage was mixed with 500µl LE392 suspension, incubated 20 minutes at 37°C and plated out, with molten top agarose onto five 90mm bottom agar plates (200µl of cells and phage mixture per plate). Plates incubated at 37°C overnight produced near-confluent lysis of the LE392 bacterial lawn. Plates were incubated at 4°C for 2 hours to allow hardening of the agar. 3ml of SM buffer was added to each plate and incubated at 4°C for several hours. The top agarose/SM buffer of all five plates was scraped into

a 50ml centrifuge tube using a spatula and the agar was broken up as much as possible and incubated for 20 minutes at room temperature. The agar was pelleted (10000xg, 10 minutes) and the liquid transferred to a 15ml Falcon tube. The Wizard™ Lambda Preps DNA Purification System (Promega) was used to purify the phage DNA according to manufacturer's instructions. Phage DNA quality and quantity were assessed by DNA agarose gel electrophoresis and spectrophotometry.

2.3.6 Screening of a *Trypanosoma brucei brucei* genomic library

The library was supplied and titred by Nick Robinson and Nils Burman of the University of Glasgow and was screened in the same way as the *L. mexicana* genomic library with a few modifications. The probe was a 1.1kb *Xba*I/*Xho*I DNA fragment containing the *LMGPI8* ORF (Figure 3.11A). Because this probe was exogenous, hybridisation was performed at lower stringency 55°C rather than 65°C as was membrane-washing (55°C, 2.5 x SSC/0.1% SDS). Phage DNA was also purified by the Wizard™ Lambda Preps DNA Purification System (Promega).

2.3.7 DNA and protein analysis

All DNA and protein analysis was performed using the Genetics Computer Group (GCG) Molecular Analysis software package (Wisconsin), Vector NTI (v5.0.3) software (Informax Inc.) and the molecular biology and proteomics programmes available at the Expasy website (http://www.expasy.ch/tools).

2.4 Biochemical methods

2.4.1 SDS-PAGE

Generally 10^8 promastigotes were lysed in 75µl of 0.25% Triton X-100 on ice. 25µl of SDS-PAGE loading buffer was added and lysates were boiled for 5 minutes.

Proteins were separated by SDS-PAGE method on 10% acrylamide mini-gels of 0.75mm thickness, which were prepared and run using a Mini-Protean II dual slab system (BioRad) according to manufacturer's instructions. Gels could be stained to detect total protein by 1% Coomassie R250 or they were electroblotted to allow specific-protein analysis.

2.4.2 Western blotting

Following SDS-PAGE, separated proteins were transferred to Hybond-C nitrocellulose membrane (Amersham) by electroblotting, using a BioRad mini transblot cell according to manufacturer's instructions. Kaleidoscope pre-stained molecular weight protein standards (Novex) were included to allow quick assessment of protein transfer to nitrocellulose. Proteins were transferred at 100V for 1 hour (4°C).

Membrane was rinsed in Western blotting transfer buffer to remove fragments of acrylamide gel before being transferred to a 50ml Falcon tube and overnight incubation at 4°C with a protein-rich Blocking solution (three were used, Blocking solution 1, 2 and 3) to block the remainder of the membrane, which was protein-free.

2.4.2.1 Antibody detection of proteins

Primary antibodies were incubated with the membrane in appropriate blocking solution, generally for 60-90 minutes at room temperature. Blots were washed 3 x 10 minutes with TBS-Tween (TBST) and secondary antibody (either anti-mouse IgG or anti rabbit

IgG) conjugated to horseradish peroxidase (HRP) (Promega) was generally used at 1 in 5000 dilution in blocking buffer for 45-60 minutes at room temperature. Blots were washed in TBST as before. Antibody binding was detected by Supersignal™ Enhanced Chemiluminescence (ECL) reagents (Pierce) according to manufacturer's instructions. Fluorescence was detected by autoradiography with Reflection™ autoradiography film (NEN™ Life Sciences). Film was developed by Kodak X-omat automated developer.

2.4.3 Immunofluorescence (*L. mexicana*)

Approximately 30µl of late log-phase/stationary-phase *L. mexicana* promastigotes were diluted to 100µl in PBS. 30µl was dropped on a glass slide and smeared over the complete surface using a second slide. The slides were allowed to dry for 5 minutes before cells were fixed in 2.5% paraformaldehyde/PBS for 15 minutes. Slides were washed in PBS before primary antibody was added (generally at 1 in 100 dilution in PBS). Slides were incubated in a humidified chamber for 30 minutes, before being washed thoroughly with PBS for 10 minutes (with several changes of PBS). Cells were then blocked for 15-30 minutes with 10% HI-FCS in PBS before being washed in PBS as before. A mixture of secondary antibody (between 1 in 100 and 1 in 500) and DAPI (4,6-diamidino-2-phenylindole) (20µg/ml final) in PBS was added to the slides, which were incubated in the dark at RT for 30 minutes. Slides were washed in PBS as before and 25µl of the anti-quenching agent MOWIOL-DABCO was dropped onto each slide under a glass coverslip. Cells were viewed by UV fluorescence microscopy on a Zeiss Axioplan fluorescence microscope with a Hamamatsu Digital Camera using the OpenLab Software Package (Improvision, University of Warwick).

2.5 Buffers and Reagents

Ampicillin: Stock solution is 100mg/ml in distilled water. Storage at –20°C. Used at 100μg/ml.

BBL Bottom agar: 1% (w/v) BBL Trypticase peptone, 100mM NaCl, 1% (w/v) agarose. Autoclave and store at room temperature.

BBL Top agarose: 1% (w/v) BBL Trypticase peptone, 100mM NaCl, 12mM $MgCl_2$, 0.7% (w/v) agarose. Autoclave and store at room temperature.

Blocking solution 1: 5% milk powder in TBS-Tween. Made up fresh.

Blocking solution 2: 5% milk powder and 10% horse serum in TBS-Tween. Made up fresh.

Blocking solution 3: 3% Bovine Serum Albumin (BSA) in TBS-Tween. Made up fresh.

Church-Gilbert's solution: 340mM Na_2HPO_4, 158mM NaH_2PO_4, 240mM SDS, 1mM EDTA. Heated gently to solubilise components. Store at room temperature.

DNA loading buffer (10x): 25% Ficoll, 0.42% bromophenol blue. Store at room temperature.

Depurination solution: 0.125M HCl. Store at room temperature.

Denaturation solution: 1.5M NaCl, 0.5M NaOH. Store at room temperature.

Dithiothreitol (DTT): Made up as a 1M stock solution in water. Store at –20°C.

Electroporation buffer (EPB): 21mM HEPES pH7.5, 137mM NaCl, 5mM KCl, 0.7mM phosphate buffer, 5mM glucose. Store at 4°C.

Ethidium bromide: 10mg/ml stock solution in distilled water. Store at room temperature.

HS Buffer: As LS buffer but minus glycerol and imidazole. Store at room temperature.

Imidazole Buffer: As LS-T plus 250mM imidazole. Store at room temperature.

Kanamycin: Made up as a 25mg/ml stock in water. Store at -20°C. Used at 25μg/ml.

LS Buffer: 50mM MOPS, 100mM NaCl, 1mM EDTA, 1mM EGTA, 10mM NaF, 1mM NaOVa, 10% (v/v) glycerol, 25mM imidazole in distilled water. Store at room temperature.

LS-T Buffer: As LS Buffer plus 1% Triton X-100. Store at room temperature.

Luria-Bertani (LB) agar: LB broth plus 0.8% (w/v) agar. Autoclaved to sterilise and store at room temperature.

Luria-Bertani (LB) broth: 1% bactotryptone (Difco), 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl. Made up in distilled water and autoclaved for sterilisation. Store at room temperature.

MG132: (carbobenzoxyl-leucinyl-leucinyl-leucinal) – a tripeptide inhibitor of the proteasome, stock solution at 10mg/ml in ethanol.

Microsome Buffer 1: 0.1mg/ml leupeptin, 5μg/ml pepstatin A, 0.5mg/ml Pefabloc, 1mM 1,10 phenanthroline in water.

Microsome buffer 2: As microsome buffer 1 plus 100mM HEPES-NaOH pH7.4, 50mM KCl, 10mM $MgCl_2$, 20% (v/v) glycerol

MOWIOL-DABCO: 6g glycerol and 2.4g MOWIOL 4-88 (Calbiochem) added to 6ml distilled water. 12ml 0.2M Tris-HCl, pH8.5 was added and incubated with agitation (shaking platform) for several hours at room temperature. Mixture was incubated at 50°C for 10 minutes then centrifuged (5000 x g, 15 minutes). 1,4-diazobicyclo[2.2.2.]octate or DABCO (Sigma) was added to 0.1% (w/v) final concentration and divided into 1ml aliquots, which were stored at –20°C. Solution was incubated at room temperature for at least an hour before use.

Neutralisation solution: 1.5M NaCl, 0.5M Tris-HCl pH7.0, store at room temperature.

PBS: 10mM phosphate buffer, 2.7mM KCl, 137mM NaCl, pH7.4. Autoclaved when required for tissue culture. Store at room temperature.

PCR mix (11.1x): 45mM Tris-HCl (pH8.8), 11mM ammonium phosphate, 4.5mM $MgCl_2$, 6.7mM β-mercaptoethanol, 4.4mM EDTA (pH8.0), 1mM dATP, 1mM dCTP, 1mM dGTP, 1mM dTTP and 113μg/ml BSA. Store at -20°C.

Periodate Buffer: 10mM sodium metaperiodate, 20mM sodium acetate, pH5.4, 5mM NaN_3, 50mM NaCl. Store at room temperature.

Periodate wash buffer 1: 0.2M glycerol, 150mM NaCl, 20mM Tris-HCl, pH7.2

Periodate wash buffer 2: As wash buffer 1 but minus glycerol

RNAse A: 10mg/ml stock solution in sterile water. Store at -20°C.

SDS-PAGE loading buffer (4x): 200mM Tris-HCl pH6.8, 400μM β-mercaptoethanol, 8% SDS, 40% glycerol plus a few crystals of bromophenol blue.

SM (Phage) buffer: 50mM Na_2HPO_4, 22mM KH_2PO_4, 86mM NaCl, 1mM $MgSO_4$, 0.1mM $CaCl_2$

SOC medium: 2% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.05% (w/v) NaCl, 0.4% (w/v) glucose. The solution is prepared without glucose (SOB medium), autoclaved then sterile glucose solution is added.

SSC (20 x): 3M NaCl, 0.3M tri-sodium citrate pH7.0. Store at room temperature.

TAE (1 x): 40mM Tris-acetate, 1mM EDTA. Store at room temperature.

TBE (10 x): 0.9M Tris-HCl, 0.9M boric acid, 25mM EDTA. Store at room temperature.

TBS-Tween (TBST, 1 x): 20mM Tris, 137mM NaCl pH7.6, 0.1% Tween 20. Store at room temperature.

TE buffer: 10mM Tris-HCl, pH7.4, 1mM EDTA pH8.0. Store at room temperature.

TELT Buffer: 50mM Tris-HCl pH8.0, 62.5mM EDTA pH9.0, 2.5M LiCl, 4% Triton X-100. Store at room temperature.

Western blotting transfer buffer: 5mM Tris, 2mM glycine, 20% methanol in distilled water. Store at 4°C.

CHAPTER 3

CLONING AND CHARACTERISATION OF *GPI8* FROM PARASITIC PROTOZOA

3.1 INTRODUCTION

The biosynthesis of glycosylphosphatidylinositol (GPI) anchors occurs mainly on the cytoplasmic face of the endoplasmic reticulum (ER) membrane in a series of stepwise additions to the anchor structure beginning with phosphatidylinositol (described in Chapter 1). As biosynthesis nears completion, the GPI anchors are flipped to the lumenal face of the ER membrane where they are completed and ready for the terminal step in biosynthesis, attachment to particular cell surface proteins (Vidugiriene and Menon, 1995). Such proteins have C-terminal signal sequences that do not appear to be well conserved evolutionarily, in terms of their primary amino acid sequence, but which have overall conserved structural features (Ferguson and Williams, 1988). GPI-signals consist of three structural components (beginning with the most C-terminal component and moving towards the N-terminus):

C-terminal hydrophobic domain - 15-20 amino acids, to keep the nascent polypeptides associated with the lumenal face of the ER membrane.

Hydrophilic spacer - 9-12 amino acids.

Cleavage/attachment (ω) site - 3 residues termed ω, ω+1 and ω+2. The ω residue receives the anchor, ω+1 and ω+2, which are C-terminal to ω, are removed with the remainder of the GPI signal. The ω site is restricted to a

relatively small number of residues, namely Ser, Asn, Gly, Asp, Ala or Cys (Moran *et al.*, 1991; Nuoffer *et al.*, 1993; Micanovic *et al.*, 1990)

The signal is recognised, removed and replaced with a GPI-anchor by a multi-protein complex known as the GPI:protein transamidase (Ferguson and Williams, 1988; and see Chapter 1). To date, most of the data relating to the transamidase subunits has come from yeast and mammalian systems. Two putative subunits have been identified in *Saccharomyces cerevisiae*:

Gaa1p – The protein is a 68kDa ER protein which has a large lumenal domain and several transmembrane domains.

Gpi8p – A 47kDa ER protein with a large lumenal domain, a single ER-membrane spanning region and a short cytoplasmic tail (Benghezal *et al.*, 1996).

Conditional yeast mutants defective in each of these proteins demonstrated similar phenotypes, a build-up of GPI anchor precursors but no attachment of the GPI-anchor to cell surface proteins at non-permissive temperatures, indicating that the terminal step in biosynthesis was blocked (Hamburger *et al.* 1995; Benghezal *et al.*, 1996). Similar phenotypes have been observed in mutant human cells referred to as class K cells (Chen *et al.*, 1996), leading to the identification of the human GPI8 homologue by complementation of the mutant phenotypes with cDNA libraries (Yu *et al.*, 1997). The human GAA1 homologue was identified in this way also (Hiroi *et al.*, 1998). Gpi8p has significant similarity to jackbean asparaginyl endopeptidase, which has been

shown to have *in vitro* transamidase activity (Abe *et al.*, 1993) suggesting that Gpi8p is the catalytic subunit of the complex. Data from T. Kinoshita and co-workers (Osaka University, Osaka, Japan) presented at the International Meeting on Interactions of GPI-anchors with Biological Membranes, Splügen, Switzerland (1997) indicated that GAA1 seems to be involved in the recognition/binding of the GPI-signal sequences whilst GPI8 catalyses the transamidation reaction.

The GPI:protein transamidase is of considerable interest in the context of parasitic protozoa, which are known to make extensive use of this method of GPI-protein attachment; for example in *Trypanosoma brucei* the variant surface glycoprotein VSG (Ferguson *et al.*, 1988) and procyclin (Clayton and Mowatt, 1989) membrane coats are the major GPI-linked proteins as is the promastigote surface metalloproteinase GP63 in *Leishmania* (Bordier *et al.*, 1986).

The presence of GPI-anchored proteins in *Leishmania* (see (Cross, 1990; Ferguson and Williams, 1988)) implies that a GPI:protein transamidase must be present. The aim of this section was to identify and clone the gene (*GPI8*) encoding the predicted catalytic subunit of the complex, in *L. mexicana*.

Our approach was to use primers for the amplification of gene fragments using degenerate oligonucleotides designed to regions of genes which encode conserved stretches of amino acids between related proteins. This method of "fishing" for the presence of parasite genes rather than proteins as described originally by Sakanari and colleagues (Sakanari *et al.*, 1989) and used successfully by Eakin and co-workers (Eakin *et al.*, 1990) has much potential since the presence of a protein can be inferred

from the presence of the gene. This also serves to allow subsequent characterisation of the primary sequence of the protein as well as further analysis, including production of recombinant protein. One of the ways of maximising the possibility of amplifying the correct gene (rather than false positives) is to carry out amino acid alignments of several related proteins and design primers against highly conserved regions each of 7 or 8 amino acids. These conserved regions often surround active site residues that maintain structural constraints during evolution. Differences in one or more amino acids in these regions, as well as codon degeneracy and codon bias in the species of interest, means that the primers themselves will be degenerate. Thus much care must be taken when designing these primers and strategies must be employed to minimise the degeneracy since the greater the degeneracy the more dilute the homologous primer will be in the degenerate mixture and the less likely a successful PCR will occur. If tables of codon bias are available, as is the case for *L. mexicana* then codon degeneracy can be reduced. *Leishmania* species have a GC-rich genome (Alonso *et al.*, 1992) so several codons are used at very low frequencies (less than 9%) and can be ignored.

The same approach was taken to identify similar genes from *T. brucei* and *Plasmodium falciparum*

3.2 RESULTS

3.2.1 Design and use of primers for PCR with *L. mexicana* template DNA

At the start of this project, three GPI8 amino acid sequences were available, (human, yeast and a putative *Caenorhabditis elegans* homologue, as well as a number of

related proteins, including legumain, a jackbean asparaginyl endopeptidase and *Schistosoma mansoni* and *S. japonicum* haemoglobinases.

An amino acid alignment of *Saccharomyces cerevisiae*, human and a putative *C. elegans* GPI8 identified in the *C. elegans* genome sequencing project with asparaginyl endopeptidases is shown in Figure 3.1. Several highly conserved sequences were identified (Figure 3.1) and used to design degenerate oligonucleotides (Table 3.1). *L. mexicana* has a GC-rich genome and four primers were designed accordingly using *Leishmania* codon bias (Langford *et al.*, 1992) and also supplementary information supplied by Dr. Colin Robertson, Glasgow. The same regions were used to design primers for *P. falciparum* taking into account the AT-richness of the *Plasmodium* genome (Musto *et al.*, 1995) and their use is described in Section 3.2.5. The sequences of all of these primers are shown in Table 3.1.

Table 3.1

Degenerate primers for PCR of *GPI8* from *L. mexicana* and *P. falciparum* and the corresponding amino acid consensus sequences

Primer Name (mer)	Organism	Sequence (5'- 3')	Length
1) OL76	*L. mexicana*	ACSAAYAAYTGGGCNGTBCTYGT	23
OL80	*P. falciparum*	ACWAATAATTGGGCWGTWCTWGT	23
	Amino Acid	**T N N W A V L V**	
2) OL77	*L. mexicana*	CGBTTYTGGTTYAAYTAYCGBCA	23
OL81	*P. falciparum*	ARATTTTGGTTTAATTATAGACA	23
	Amino Acid	**R F W F N Y R**	
3) JATR3	*L. mexicana*	TCNTMVSHVCGRTARTCVACYTCVAC	26
	Amino Acid	**V E V D Y R G/S Y/E E**	
4) OL78	*L. mexicana*	CCVCCRTGVCCSGTVAKRTARCT	23
OL82	*P. falciparum*	CCWCCRTGWCCWGTHAKATAWAT	23
	Amino Acid	**I Y M/L T G H G G**	

Key to nucleic acid symbols : R=A+G, **Y**=C+T, **M**=A+C, **K**=G+T, **S**=G+C, **W**=A+T, **H**=A+T+C, **B**=G+T+C, **N**=A+G+C+T, **V**=G+A+C

3.2.1.1 PCR using *L. mexicana* genomic DNA and cDNA

Nested PCR was used as a means of increasing the specificity of the polymerase chain reaction. In the first round of PCR the degenerate primers OL76 and OL78 were used with genomic DNA, log-phase (promastigote) first strand cDNA, stationary-phase (metacyclic promastigote) first strand cDNA and amastigote first strand cDNA template at low annealing temperature (42°C). The cDNA template was produced from poly A^+ mRNA purified from log-phase axenic promastigotes, stationary phase axenic promastigotes and mouse lesion-derived amastigotes respectively. Reactions were diluted 10 fold with water and 2µl was used in a second round of PCR, this time with primers OL77 (which is internal to OL76) and OL78 under the same conditions as the first round. Based on the yeast and human sequences, the size of the expected PCR product would be approximately 350bp. The reactions were electrophoresed on an agarose gel (Figure 3.2). A clear DNA fragment of the expected size was amplified from all *L. mexicana* DNA samples (Lanes 1-4). The products from all four template reactions were cloned into the pTAG vector utilising the overhanging adenosine nucleotides left at the 3` end of each strand of amplified DNA by Taq polymerase, during PCR (the vector is supplied in a linear form with overhanging 5` thymidine residues). The insert disrupts the plasmid *lacZ* gene allowing positive colonies to be screened by blue/white selection. White colonies were checked for recombinant plasmid by whole cell PCR using primers OL77 and OL78 under the same conditions that generated the original PCR products (see Figure 3.2 legend). Sequencing of these four PCR products using the T7 primer revealed that all had identical sequences of 362bp, other than at the extreme ends where primer degeneracy resulted in some

differences (see Figure 3.3). The presence of equal PCR products using each template suggests that *GPI8* is being expressed during all stages of the life cycle, however this would need to be confirmed by Northern analysis to detect the transcript directly.

A BLAST search was carried out using these sequences translated using all six reading frames (BLASTX) and similarity was found with yeast and human GPI8, as well as the putative *C. elegans* GPI8, legumain and other cysteine proteinases (Figure 3.4).

3.2.1.2 Southern analysis with *L. mexicana* PCR product probe

To confirm that the PCR products were derived from *L. mexicana*, Southern blot analysis was performed. *L. mexicana* gDNA was digested with *Eco*RI, *Hin*cII, *Hin*dIII, *Pst*I or *Xho*I. The digests were electrophoresed and blotted by capillary transfer. The blot was probed with $[\alpha]^{32}$P-CTP labelled gDNA PCR product cut from the pTAG with *Pvu*II yielding a 700bp fragment consisting of the PCR product flanked by vector sequence. After washing under high stringency, the blot was exposed to X-ray film for 5 days. The results are shown in Figure 3.5. A single hybridising band was detected in the *Xho*I, *Hin*dIII and *Eco*RI lanes (Figure 3.5, lanes 1, 3 and 5, respectively) whilst two bands were detected in the *Pst*I and *Hin*cII lanes (lanes 2 and 4). The *Eco*RI digest was slightly smeared which could have been due to more than one band or non-specific hybridisation of the probe. The hybridisation confirmed that the PCR products were derived from *L. mexicana* and indicate that the gene is probably single-copy. There is some discrepancy; the *Pst*I digest shows two bands, which correlates with the fact that the PCR product contains a *Pst*I site. By contrast the *Hin*cII digest gives two bands, a major one at 3kb and a second, fainter

band of 3.8kb, despite the apparent absence of such a site within the sequence. It is possible that the two *Hin*cII sites have arisen through restriction fragment length polymorphism (RFLP) in one of the *Hin*cII sites rather than the second band being due to another copy of the gene since the other digests produce single bands. This was investigated further (see Section 3.2.3).

3.2.1.3 Isolation of full length *GPI8* from an *L. mexicana* genomic library

The focus for subsequent work was to be with the *L. mexicana* putative *GPI8*, thus it was initially necessary to isolate the full-length gene. A *L. mexicana* genomic library consisting of a *Sau*3AI partial genomic digest cloned into λDashII Vector (Stratagene) (Mottram *et al.*, 1997) was used at a titre of 5.5×10^5 plaque forming units (pfu)/μl and was initially screened by plating out with the LE392 strain of *E. coli*. At least 40,000 plaques, which arose through the localised lysis of the bacteria by the infecting λphage, were lifted (by brief capillary flow) and the λ DNA was fixed onto nitro-cellulose filter (Hybond-N) following denaturing and neutralisation. The membranes were screened by hybridisation with the purified *Pvu*II fragment containing the cloned *L. mexicana* gDNA PCR product originally used in the Southern analysis (Section 3.2.1.2). Ten hybridising plaques were identified and picked into 500μl of fresh phage buffer and four of these were used in a second round of screening. Four independent λ clones were isolated after secondary screening and two of these (λ1 and λ2) were grown up to near confluent lysis on 150mm diameter bacterial plates and the λ DNA purified. Both λ DNA preps were digested with a number of restriction enzymes (*Bam*HI, *Eco*RI, *Hin*dIII, *Sal*I and *Xho*I; λ 2 was also digested with *Xba*I) and the

digested DNA was electrophoresed on an agarose gel and blotted. The blot was probed with the gDNA PCR fragment and washed under high stringency before being exposed to X-ray film for about 30 minutes. The resulting autoradiogram is shown in Figure 3.6. Three of the digests yielded bands of convenient sizes for subcloning i.e. less than 5-6kb. Three bands which hybridised to the probe, λ1 *Sal*I (2.2 kb); λ2 *Xho*I (2.4kb) and λ2 *Xba*I (4.5kb) were subcloned into pBluescript SK+ (Stratagene). Following transformation of XL1-blue supercompetent *E. coli*, ampicillin-resistant colonies were screened by blue/white selection and whole-cell PCR using primers OL77 and OL78. Positive colonies were grown up and plasmid mini-preps were carried out; test digests with the corresponding enzymes confirmed that correctly sized inserts were present within the plasmids (data not shown). The 2.2kb *Sal*I fragment, the 2.4kb *Xho*I fragment and the 4.5kb *Xba*I fragment were called pGL185, pGL186 and pGL187, respectively.

3.2.1.4 Sequencing of the *L. mexicana* genomic fragments

The subcloned lambda fragments were sequenced to yield 3.5kb of double stranded sequence covering the open reading frame of the *L. mexicana GPI8*. New primers were designed to overlap the ends of existing sequence (as it was obtained) as well as antisense primers to obtain sequence from the complementary strand. The 5` flank and ORF was obtained from pGL186 whilst the 3` flank was obtained from pGL185 and pGL187 (see Figure 3.7). Sequence was analysed by the Molecular Biology Support Unit (MBSU, University of Glasgow) or Oswell Scientific (University of

Southampton) and between 300 and 600bp of sequence was obtained from each reaction.

The complete double stranded sequence of the coding region of the gene as well as 1.3kb of 5' and 1kb of 3' flanking region was determined (Figure 3.8A). The longest open reading frame was 1047bp (Figure 3.8A, shown in red text) and the corresponding amino acid sequence is given in Figure 3.8B. This sequence contained the predicted amino acid sequence obtained from the cloned PCR products (Section 3.2.1.1).

There were two further open reading frames of the same orientation as the putative *GPI8*. One in the 5` flanking region of 740bp began 775bp and ended 35bp upstream of *GPI8* start codon. The other ORF was in the 3` flanking region (704bp) and began 32bp downstream of the *GPI8* stop codon. Blast searches of the encoded proteins determined that neither of these proteins had significant homology to any proteins in the databases.

The complete 3337bp nucleotide sequence was submitted to the EMBL database with the accession number AJ242865 and the amino acid sequence of LMGPI8 has the accession number CAB55340.

3.2.2 Primary structure of LMGPI8

The longest of the three open reading frames (ORF) was 1050bp and encoded a protein of 349 amino acids with a predicted molecular mass of 38,900Da.

The predicted primary sequence of LMGPI8 was compared with GPI8 from other species (Figure 3.9). BLAST searches at the National Centre for Biotechnology Information internet web site using the BLASTP programme which compares amino acid sequences with those in protein databases such as Swiss-prot, revealed homology of the predicted *L. mexicana* GPI8 with yeast and human GPI8 as well as the putative *C. elegans* GPI8. The plant asparaginyl endopeptidases including the Jack Bean endopeptidase which has been shown to exhibit *in vitro* transamidase activity (Abe *et al.*, 1993) and the schistosome haemoglobinases (Klinkert *et al.*, 1989; Gotz and Klinkert, 1993) were found to have lower homology to LMGPI8 than yeast GPI8. The "GAP" programme in the GCG package was used to calculate percentage identities between LMGPI8 and the related proteins and the data are given in Table 3.2.

Table 3.2

Percentage identities between LMGPI8 and related proteins as calculated by the "GAP" programme (GCG Package, Wisconsin)

	ScGPI8	HsGPI8	CeGPI8	Legumain	Sm31
LMGPI8	**31**	**31**	**31**	**19**	**21**
ScGPI8		**46**	**43**	**19**	**22**
HsGPI8			**46**	**20**	**22**
CeGPI8				**20**	**22**
Legumain					**22**

Lm; *Leishmania mexicana*
Sc; *Saccharomyces cerevisiae*
Hs; *Homo sapiens*
Ce; *Caenorhabditis elegans*
Legumain – a plant asparaginyl endopeptidase
Sm31 – haemoglobinase from *Schistosoma mansoni*

L. mexicana GPI8 shares 31% identity with both human and yeast GPI8, somewhat less than the 46% these two homologues share with one another. This is still

significantly higher than the 19-21% identity shared with the other related enzymes, the legumains and the Schistosome haemoglobinases. The *C. elegans* protein appears to be more similar to the yeast and human proteins, however the protein sequence may not be complete since this is a hypothetical protein derived from the *Caenorhabditis elegans* genome sequencing project (see next section).

3.2.2.1 LMGPI8 has several conserved amino acids of interest

Comparison of the amino acid sequence of LMGPI8 with the yeast and human homologues as well as *Canavalia ensiformis* (jack bean) legumain is shown in Figure 3.9. Cysteine proteinases belonging to the C13 family (Barrett and Rawlings, 1996), possess the classical catalytic dyad residues cysteine and histidine and the positions of conserved Cys and His residues are indicated in Figure 3.9. Two histidine (His^{63} and His^{174}) and two cysteine residues (Cys^{94}, Cys^{216}) were conserved between all of the GPI8 homologues. Whilst both histidine residues were found to be conserved in legumain and GPI8, of the cysteine residues, only Cys216 was conserved making it the most likely candidate for the active site cysteine, assuming that GPI8 is a cysteine proteinase and that legumain has a similar three dimensional structure.

LMGPI8 also possesses three potential glycosylation sites at N^{163}, $N^{268}$and N^{307} (Figure 3.9).

3.2.2.2 Comparison of hydrophobicity profiles of GPI8 homologues

The EXPASY Internet web site (http://www.expasy.ch/tools) provides a programme called "Prot-Scale" which calculates a hydrophobicity plot of any given amino acid sequence using a number of algorithms; the Kyte-Doolittle algorithm was used in this case. This assigns scores to individual amino acids according to degrees of hydrophobicity such that hydrophobic residues have scores greater than zero and hydrophilic residues score less than zero. The data is presented in numerical format, which can be entered into graphical software packages such as Microsoft Excel. Thus an overall hydropathy profile of a polypeptide can be constructed graphically and compared with those of related proteins. Similar profiles of hydrophobicity and hydrophilicity may be indicative of similar folding patterns in protein secondary structure. The hydrophobicity plot of LMGPI8 is compared with yeast and human GPI8 in Figure 3.10. There are notable patterns of similarity between the hydrophobicity profiles of each GPI8, including a strongly hydrophobic region at the N-terminus (corresponding to the endoplasmic reticulum signal peptide (Walter *et al.*, 1984)) and similar peaks of hydrophobicity and hydrophilicity throughout most of the remainder of the protein. For example, the hydrophobic peaks between approximately 140 and 210 amino acids of *L. mexicana* GPI8, indicated by the blue arrows in Figure 3.10A correspond to those between 120 and 190 amino acids of *S. cerevisiae* Gpi8p (Figure 3.10B). This region also has a similar profile in human GPI8, between 140 and 200 amino acids (Figure 3.10C). There is a second strongly hydrophobic region at the C-terminus of both the yeast and human GPI8 proteins corresponding to ER transmembrane domains of these ER-lumen proteins (Benghezal *et al.*, 1996). Interestingly, this is absent in LMGPI8.

3.2.3 Confirmation of *LMGPI8* copy number by Southern analysis

A number of genomic digests using either one or two restriction enzymes were set up to confirm that *GPI8* is a single copy gene. The enzymes chosen and the fragment sizes are given in Table 3.3.

Table 3.3
Restriction digests of genomic DNA in the investigation of *GPI8* copy number

Digest	Fragment size (kb)
*Hin*cII	3.1
*Hin*dIII	>10
*Pst*I	1.5 and 2.4
*Xba*I	5.0
*Xho*I	2.4

The digests were electrophoresed and blotted as before and were probed with a 1.1kb *Xba*I/*Xho*I *GPI8* fragment obtained from pGL187 (Figure 3.11A) which represents the entire coding region of the gene (including 30bp of the 5' flank). The Southern blot was exposed to X-ray film for several days and is given in Figure 3.11B. The presence of a single 3.1kb band in the *Hin*cII digest precisely matches the known DNA sequence (Panel A and 3.11B, lane 1). This indicates that the other bands seen using the cloned PCR product as probe (Section 3.2.1.2) were probably due to non-specific hybridisation of the PCR product or the plasmid flanking regions rather than a true restriction fragment length polymorphism. All of the other digests produced single bands except *Pst*I (Figure 3.11B – lane 3) which produced two as expected, given the presence of a *Pst*I site in the gene coding region (Figure 3.11A). From the DNA sequence, a *Pst*I site is found 1.5kb downstream of the *Pst*I site within the open reading frame, so there must be a third *Pst*I site 2.4kb upstream of the this ORF site to

yield the larger of the two fragments. Taken together, the data indicates that *GPI8* is a single copy gene.

3.2.4 PCR using *Trypanosoma brucei* genomic DNA

PCR was carried out firstly using the *Leishmania* primers in all of the possible combinations with *T. brucei* genomic DNA, supplied by Nick Robinson (Glasgow). A single specific DNA fragment was detected, however this proved to be somewhat larger than expected at 500bp. Nevertheless, it was selected for cloning and sequencing but was found to be a gene fragment encoding the major cathepsin L-like cysteine proteinase from *T. brucei* (Mottram *et al.*, 1989) rather than a *T. brucei GPI8* homologue. The experiment on *T. brucei* DNA was repeated using the *P. falciparum* primers (OL80-82), however no PCR products could be detected. After the sequence of the *L. mexicana GPI8* was obtained it was found that there were several amino acid substitutions within the *L. mexicana* primer regions that differed from those used to design the original degenerate primers. On the assumption that the *T. brucei* gene was likely to be very similar to that of *L. mexicana* two new degenerate primers were designed with no codon bias (reflecting the neutral G/C content of the African trypanosome genome (Musto *et al.*, 1994)) to this revised sequence (corresponding to primers OL77 and OL78 (see Table 3.1)). These are given in Table 3.4.

Table 3.4

Degenerate primers for nested PCR of *T. brucei GPI8* transamidase gene and the corresponding amino acid sequence.

Primer Name	Organism	Sequence (5' - 3')	Length (mer)
1) OL76	*Leishmania* (and *T. brucei)*	ACSAAYAAYTGGGCNGTBCTYGT	23
	Amino Acid	**T N N W A V I V**	

2) OL47	*T. brucei*	MGNTAYYTNTTYAAYTAYMGNCA	23
	Amino Acid	**R/S Y L/F F N Y R/S H**	
3) OL48	*T. brucei*	GCNCCRTGNCCNGCNACRTADAT	23
	Amino Acid	**I Y V A G H G A**	

Key to nucleic acid symbols: R=A+G, **Y**=C+T, **M**=A+C, **K**=G+T, **S**=G+C, **W**=A+T, **H**=A+T+C, **B**=G+T+C, **N**=A+G+C+T, **V**=G+A+C

PCR was carried out as described previously (Figure 3.2 Legend) using primers OL76 and OL48 in the first round and OL47 and OL48 in the second round. This time, however, no PCR products were detected despite repeated attempts at this reaction. It is possible to vary the components of reactions to obtain optimum conditions, however this was not attempted, as an alternative method of isolating the gene, heterologous library screening was attempted.

3.2.4.1 Screening of a *T. brucei brucei* genomic library with the *L. mexicana GPI8*

The close evolutionary distance between African trypanosomes and *Leishmania* means that there would be expected to be a high level of nucleotide sequence identity between *GPI8* homologues, as has been found for many homologue genes between the species. For example, the type I (cathepsin L-like) cysteine proteinases (Souza *et al.*, 1992; Mottram *et al.*, 1989) and the cyclin dependent kinase CRK3 (Grant *et al.*, 1998) share 47% and 78% identity respectively. Thus the *Xba*I/*Xho*I fragment from pGL187 containing *LMGPI8* ORF (Figure 3.11A) was used as a probe to screen a *T. brucei brucei* genomic library (prepared and supplied by Dr. Nils Burman and Nick Robinson of the University of Glasgow) at low stringency (see Materials and Methods). The procedure was carried out as described in Section 3.2.1.3 except that

only one obvious hybridising plaque was obtained in the primary screen. Secondary screening yielded several hybridising plaques of which two were isolated and their DNA purified.

Both phage preps were digested with a number of restriction enzymes (see Figure 3.12 legend), electrophoresed, blotted and probed at low stringency to determine which fragments contained the putative *TBGPI8*. The gels were stained with ethidium bromide and visualised prior to blotting. Both clones produced identical restriction patterns as expected from having come from a single plaque at primary screening (Figure 3.12A). The probe hybridised to fragments of just over 5kb in the *Hin*dIII and *Xho*I digests (Figure 3.12B, Lanes 1 and 4, respectively). The *Sal*I digest (Figure 3.10A, Lane 2) produced a smear of DNA in each case, which may be due to the presence of contaminating deoxyribonucleases (restriction enzymes are stored as communal stocks used by several research groups, so the possibility of contamination is increased). Despite this, a 10kb fragment was labelled on the Southern blot (Figure 3.12B, lane 2).

Due to time constraints, the *TBGPI8* has not yet been subcloned, however the *Hin*dIII and *Xho*I fragments are of a good size for subcloning and will be gel purified and ligated with pBluescript (Stratagene) for further analysis in the near future.

3.2.4.2 PCR with *T. b. brucei* λ DNA containing putative *TBGPI8*

PCR reactions using both primer pairs OL76 and OL48 and OL47 and OL48 were carried out using as template, DNA isolated from λ phage described in the previous

section. A PCR product of the expected size of about 350bp was detected with primers OL76 and OL48 but not with OL47 and OL48 (Figure 3.13A). This fragment was gel-purified and cloned into pGEM-T Vector (Promega), similar to the pTAG vector used with *L. mexicana* PCR products (Section 3.2.1.1). Plasmid (pGL446) was isolated from transformants using Plasmid Mini-Prep Kit (Sigma) and it was sequenced in double stranded form (Figure 3.13B).

The sequence was used in a BLAST search using the BLASTX programme (as described in Section 3.2.1.1) and was found to have greatest similarity to GPI8 proteins of yeast and humans (Figure 3.13C). If the primer regions, which are degenerate and unlikely to be of the exact gene sequence, are ignored, the predicted amino acid sequence is found to share around 70% identity with *L. mexicana* GPI8 over this region (Figure 3.13D) as determined by the "GAP" programme (GCG, Wisconsin). Thus a putative *T. brucei GPI8* has been isolated.

3.2.5 PCR using *Plasmodium falciparum* genomic DNA

To identify a putative *Plasmodium* homologue of *GPI8* PCR was carried out with the *L. mexicana* primers on *P. falciparum* genomic DNA (a gift from Dr. John Hyde, UMIST; all conditions were identical to those described in Section 3.2.1.1.), however no PCR products were obtained. Primers were redesigned (Table 3.1: OL80-82) using *P. falciparum* codon bias (Musto *et al.*, 1995) and these were used in nested PCR as described previously, for *L. mexicana* (Section 3.2.1.1). This time, however, three bands were detected between about 300 and 500bp (Figure 3.14). In order to confirm which of these were genuine products of all primers used (and not just from single primers binding non-specifically at this relatively low annealing temperature),

reactions using all of the combinations of primers both alone and with another primer were also carried out (see Figure 3.14 legend). Only one band (arrowed in yellow) was found to be specific as the other bands were also amplified when only one primer was used. This band was selected for purification, cloning and sequencing.

The gel purified 350bp fragment was cloned into the pGEM-T vector and a number of positive colonies were identified. Recombinant plasmid (pGEM-T/PfPCR, the plasmid was not assigned a pGL number) was isolated by the Wizard Plus Mini-Prep Kit and sequenced. BLAST sequence similarity searches with the nucleotide sequence translated in all six reading frames (BLASTX programme) revealed that the PCR product had high percentage identities with both human and yeast GPI8 (70 and 85 percent respectively) while the nucleotide sequence identity was 62 and 73% identical respectively. The percentage of either A or T in the third position of each codon was more than 75%, which is indicative of *Plasmodium* DNA due to high A/T-rich genome (Musto *et al.*, 1995). To test if this PCR product was of *Plasmodium* origin Southern analysis was carried out using the cloned fragment excised from the plasmid vector, as a probe.

3.2.5.1 *P. falciparum* Southern analysis

P. falciparum genomic DNA was digested with *Eco*RI and *Hin*dIII and these reactions were electrophoresed on a gel and blotted as before. The probe was derived from the *P. falciparum* cloned fragment, which was excised from pGEM-T/PfPCR by digestion with *Sac*I and *Sal*I giving a fragment of about 550bp (including some vector sequence)

and this was labelled with $[\alpha]^{32}$P-CTP using the Prime It II Random Primer Kit (Stratagene). No hybridisation to *Plasmodium* DNA was observed, even when low stringency conditions were used (not shown). A positive control to check the quality of *Plasmodium* DNA, consisting of a 750bp PCR fragment of the *P. falciparum* dihydrofolate reductase gene was labelled and used to probe the same blot. This demonstrates a single hybridising band in each digest lane as expected indicating that the DNA had indeed transferred to the nitro-cellulose (not shown). Thus it is likely that the PCR product was not of *P. falciparum* origin. The high level of sequence identity with the yeast GPI:protein transamidase suggests that there may have been some contaminating yeast DNA in the *Plasmodium* DNA used for PCR reaction. Perhaps the culture was yeast-contaminated prior to DNA purification.

The finding that the cloned PCR product amplified from a sample of *P. falciparum* gDNA did not hybridise to Plasmodium DNA suggested a low-level contamination of the template with DNA from another source. The encoded amino acid sequence of cloned fragments had high homology to both yeast and human proteins, but was sufficiently different to rule out contamination from either of these sources. It should be possible to identify *P. falciparum* GPI8 from sequence data generated by the genome project and the database is screened on a regular basis.

3.3 DISCUSSION

We predicted the presence of *L. mexicana GPI8* and subsequently cloned a gene, which encodes a protein with homology to yeast and human Gpi8p (yeast nomenclature). Fragments of genomic DNA, which included the ORF and 5` and 3`

flanking regions, were characterised by automated sequencing using specifically designed primers such that double stranded sequence was obtained for a 3.5kb region, consisting of *GPI8* and 1.3kb and 900bp of the 5` and 3` flanking regions, respectively.

The initial PCR data where fragments of GPI8 were cloned from cDNA produced from different life cycles, indicated that the gene was likely to be expressed during all three stages of the *Leishmania* life cycle. This remains to be confirmed, either by the direct detection of the protein itself or by Northern analysis to detect the transcript. It is known, however that, in many cases, regulation of gene expression occurs post-transcriptionally in trypanosomatids since many genes are transcribed simultaneously in the form of large, polycistronic messenger RNA molecules which are processed into individual mRNA molecules of varying stability (Coulson *et al.*, 1996; Curotto de Lafaille *et al.*, 1992; Vanhamme and Pays, 1995). Regulation at the post-translational level is apparently not as prevalent but does also occur. Expression of GP63, for example, is regulated post-translationally, where glycosylation of the protein contributes to its intracellular stability and the addition of a GPI-anchor is necessary for the protein to remain associated with the cell ((McGwire and Chang, 1996; Mensa-Wilmot *et al.*, 1994) and Chapter 4). The presence of GPI-anchored proteins in all life cycle stages, for example GP63 in promastigotes but not in amastigotes (Medina-Acosta *et al.*, 1993) and promastigote surface antigen 2 (PSA2), which has recently been detected in amastigotes (Handman *et al.*, 1995) would also mean a requirement for the transamidase in each. Therefore direct detection of the protein by Western blotting using GPI8 specific antibodies would be better evidence for the presence of the GPI:protein transamidase in each stage of the life cycle.

The primary sequence of *L. mexicana* GPI8 shares a significant degree of identity with both yeast and human homologues. The proteins have similar hydropathy profiles to one another, perhaps indicating similarities in protein secondary structure (something that can be clarified, eventually, by X-ray crystallography) although no GPI:protein transamidase structures have been solved to date. The apparent lack of a C-terminal transmembrane domain, which keeps the yeast and human homologues associated with the ER membrane, may suggest differences in the way GPI8 is anchored to the ER membrane. It is possible that GPI8 interacts with another ER integral membrane protein, perhaps with a homologue of GAA1 (Hamburger *et al.*, 1995; Hiroi *et al.*, 1998). If GPI8 does not contain an ER transmembrane domain, then how is the protein retained within the ER? The four amino acid residues at the C-terminus are YDLE and may function as an ER retention motif given that they are similar to the classical mammalian ER-retention signal, KDEL (Munro and Pelham, 1987) and HDEL in yeast (Pelham *et al.*, 1988). The ER molecular chaperone, BiP in *T. brucei* has the tetrapeptide MDDL and deletion of this motif results in secretion of the protein (Bangs *et al.*, 1996). The GPI8 tetrapeptide could be investigated by addition of this sequence to a protein that is normally secreted (or targeted to post-ER compartments such as the lysosome) and immunofluorescence could be used to determine its final location compared with wild type cells. Alternatively, the tetrapeptide could be deleted and the effect on GPI8 could be determined either directly, by immunofluorescence (when antibodies are available), or indirectly, by the assessment of GPI8 activity through analysis of GPI-anchor attachment for proteins such as GP63.

Variation between ER isoenzymes involved in GPI biosynthesis has been reported in other organisms. Two classes of dolichol phosphate mannose synthase (DPMS), have been identified in *T. brucei*, yeast (*Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*) and humans (Colussi *et al.*, 1997). The trypanosome and budding yeast enzymes contain a C-terminal hydrophobic domain, whereas human and fission yeast enzymes do not. Both the human and *Saccharomyces cerevisiae* genes can complement a lethal null mutant in fission yeast, however, indicating that the enzymes are functionally equivalent (Colussi *et al.*, 1997). It is likely that the *L. mexicana* GPI8 will be able to functionally replace the human enzyme; there may be differences associated with the substrate binding pockets as has been demonstrated with *T. brucei* GPI:protein transamidase (Moran and Caras, 1994). VSG, when expressed in mammalian cells was not expressed on the cell surface until the cleavage/attachment site (ω, ω+1, ω+2) of a mammalian protein was used to replace the VSG ω site. It was concluded that the binding pocket of the trypanosome transamidase would have to be larger than that of the human homologue because to accommodate the VSG ω site. We would expect the *T. brucei* GPI:protein transamidase to be similar to the *L. mexicana* enzyme in terms of structure and perhaps also substrate specificity; the former appears to be less stringent than its mammalian homologue.

GPI8 belongs to the C13 family of cysteine proteinases (Barrett and Rawlings, 1996; Riezman and Conzelmann, 1998), which includes a family of plant cysteine proteinases known collectively as the legumains (Ishii, 1994) and the schistosome haemoglobinases (Dalton *et al*, 1995a; Dalton *et al*, 1995b; Gotz and Klinkert 1993). Cysteine proteinases such as papain and legumain require a cysteine and a histidine

residue for activity and primary sequence analysis has revealed four residues that could potentially be part of the GPI8 active site. Cys^{94}, Cys^{216}, His^{63} and His^{174} residues (*L. mexicana* GPI8 numbering) are conserved between *L. mexicana*, yeast and human GPI8. Cys216, is also conserved in legumain from *Canavalia ensiformis*, which has been shown to have *in vitro* transamidase activity (Abe *et al.*, 1993) making it the most likely candidate, if GPI8 is a cysteine proteinase. Activity of the *T. brucei* GPI:protein transamidase is sensitive to sulphydryl alkylating reagents, strongly suggesting that GPI8 is indeed a cysteine proteinase (Sharma *et al.*, 1999a). The protein also contains three potential N-glycosylation sites however, none of these are conserved with the N-glycosylation sites on yeast Gpi8p, which is heavily glycosylated (Benghezal *et al.*, 1996). Further investigation of the active site residues and glycosylation of GPI8 is described in Chapter 5. In addition to these features, there is a 5 amino acid insert (beginning at residue 112, *Leishmania* numbering, Figure 3.9) not present within either *Saccharomyces cerevisiae* Gpi8p or in human GPI8. Interestingly, part of this insertion is absent from the putative TBGPI8 described below (and Figure 3.13C).

The use of *LMGPI8* ORF as a heterologous probe to screen a *T. b. brucei* genomic library has led to the isolation of a single lambda clone which is likely to contain the ORF of *TBGPI8*. PCR using the primers specifically designed for amplification of a *T. brucei* PCR product (OL76 and OL48) with the purified lambda clone yielded a PCR product of the expected size. When cloned, sequenced and translated, this was found to have a high degree of identity (70%) with the *L. mexicana* protein over this region. Interestingly, the outermost primer pair designed for nested PCR produced this fragment, however the innermost primer pair (OL47 and OL48) did not, suggesting

that there is a problem with primer OL47. Analysis of the sequence to which OL47 would anneal indicates that there is no obvious reason that the primer doesn't function since the actual *TBGPI8* sequence is found within the degeneracy of the primers. Perhaps the primer has been synthesised incorrectly. The sequence of the *L. mexicana* equivalent primer (OL77) is similar enough that it may work with *T. brucei* DNA template in combination with the other *T. brucei* primers (OL76 and OL48), however this remains to be tested.

Sequence analysis is highly indicative that the *L. mexicana* gene encodes a functional equivalent of yeast Gpi8p however it remains to be seen if the encoded protein does indeed have a role as a component of the GPI:protein transamidase complex as has been seen with yeast and human homologues. It remains a possibility that LMGPI8 is an asparaginyl endopeptidase involved in some, as yet unknown, biochemical pathway in *Leishmania* perhaps analogous to the seed storage protein-maturation role exhibited by the legumains (Hara-Nishimura *et al.*, 1995; Abe *et al.*, 1993) since it is now known that legumains are not confined to plants (Chen *et al.*, 1997; Chen *et al.*, 1998b). In addition, the similarity to the schistosome haemoglobinases, which are thought to be processing enzymes that bring about the activation of another proteinase (Dalton *et al*, 1995a; Dalton *et al*, 1995b; Gotz and Klinkert 1993) could mean that LMGPI8 has some role in proteinase processing. The remaining chapters will describe experiments carried out to address the question of the role of LMGPI8. The fact that a single copy gene encodes LMGPI8 means that the function of the protein can be assessed by gene knockout experiments.

Figure 3.1: Amino acid sequence alignment of N-terminal regions of *Saccharomyces cerevisiae* and *Homo sapiens* GPI8, *Caenorhabditis elegans* putative GPI8, two Schistosome asparaginyl endopeptidases (haemoglobinases) and legumain. Numbering of the pileup is according to *H. sapiens* GPI8. Note that this protein was predicted from a partial cDNA such that the N-terminus is incorrect. The complete protein sequence is now known (Yu *et al.*, 1997). The conserved regions used to design primers are highlighted as follows: **OL76**, **OL77**, **JATR3** and **OL78**.

Abbreviations:

Sc = *S. cerevisiae* GPI8 (Benghezal *et al.*, 1996)

Hs = *H. sapiens* GPI8 (Benghezal *et al.*, 1996)

and (Yu *et al.*, 1997)

Ce = *C. elegans* putative GPI8

Sj = *Schistosoma japonicum* haemoglobinase (Swiss-Prot accession number P42665)

Sm = *Schistosoma mansoni* haemoglobinase (Gotz and Klinkert, 1993)

Lg = Legumain (*Canavalia ensiformis* asparaginyl endopeptidase (Takeda *et al.*, 1994; Abe *et al.*, 1993)).

```
    1                                                            50
Sc  ~~~~~~~~~M RIAMHLPLLL LYIFLLPLSG ANNTDAAHEV IATNINNWAV
Hs  SLHEADSLSR AATVLATVLL LSFGSVAASH IEDQAEQFFR .SGHFNNWAV
Ce  ~~~~~~MRST AIHNSAQFPP LRFRIFFGFR FERSSGLSEK VTGHFNNWAV
Sj  ~~~~~~~~~M FYSIFFIHIL RIVLVDCN.. ...EYSEENV DDR..HKWAV
Sm  ~~~~~~~~MM LFSLFLISIL HILLVKCQLD TNYEVSDETV SDN..NKWAV
Lg  ~~~~~~~MVM MLVMLSLHGT AARLNRREWD SVIQLPTEPV DDEVGTRWAV

    51                                                          100
Sc  LVSTSRFWFN YRHMANVLSM YRTVKRLGIP DSQIILMLSD DVACNSRNLF
Hs  LVCTSRFWFN YRHVANTLSV YRSVKRLGIP DSHIVLMLAD DMACNPRNPK
Ce  LVCTSKFWFN YRHVSNVLAL YHSIKRLGIP DSNIIMMLAE DVPCNSRNPR
Sj  LVAGSNGFEN YRHQADVCHA YHVLLSKGVK PEHIITFMYD DIAHNKENPF
Sm  LVAGSNGYPN YRHQADVCHA YHVLRSKGIK PEHIITMMYD DIAYNLMNPF
Lg  LVAGSNGYGN YRHQADVCHA YQLLIKGGVK EENIVVFMYD DIAYNAMNPR

    101                                                         150
Sc  PGSVFNNK.. ......DHAI DLYGDSVEVD YRGYEVTVEN FIRLLTDRWT
Hs  PATVFSHK.. ......NMEL NVYGDDVEVD YRSYEVTVEN FLRVLTGRIP
Ce  PGTVYAAR.. .......AGT NLYGSDVEVD YRGEEVTVES FIRVLTGRHH
Sj  PGKIFNDYRH KDYYKG.... ......VVID YKGKKVNPKT FLQVLKGDKR
Sm  LGKLFNDYNH KDWYEG.... ......VVID YRGKKVNSKT FLKVLKGDKS
Lg  PGVIINHPQG PDVYAG.... ......VPKD YTGEDVTPEN LYAVILGDKS

    151                                                         200
Sc  EDHPKSKRLL .TDENSNIFI YMTGHGGDDF LKFQDAEEIA SEDIADAFQQ
Hs  PSTPRSKRLL .SDDRSNILI YMTGHGGNGF LKFQDSEEIT NIELADAFEQ
Ce  PATPRSKRLL .TDHQSNVLI YLTGHGGDSF MKFQDSEELT NVDLAYAIQT
Sj  ....AGGKVL KSGKNDDVFI YFTDHGAPGI LAFPDDD.LH AKPFINTLKY
Sm  ....AGGKVL KSGKNDDVFI YFTDHGAPGL IAFPDDE.LY AKEFMSTLKY
Lg  KVKGGSGKVI NSNPEDRIFI FYSDHGGPGV LGMPNAPFVY AMDFIDVLKK

    201                                                         250
Sc  MYEKKRYNEI FFMIDTCQAN TMYSKFYSPN ILAVGSSEMD ESSYSHHSDV+++
Hs  MWQKRRYNEL LFIIDTCQGA SMYERFYSPN IMALASSQVG EDSLSHQPDP+++
Ce  MFEDNRYHEM LVIADSCRSA SMYEWIDSPN VLSLSSSLTH EESYSYDVDT+++
Sj  LRQHRRYSKL VIYVEACESG SMFAGLLPTD INIYATTAAR PDESSYATFC+++
Sm  LHSHKRYSKL VIYIEANESG SMFQQILPSN LSIYATTAAN PTECSYSTFC+++
Lg  KHASGGYKEM VIYIEACESG SIFEGIMPKD LNIYVTTASN AQENSFGTYC+++
```

Figure 3.2: *Leishmania mexicana* PCR products using nested PCR. Initially, primers OL76 and OL78 (see Table 3.1) were used with Taq polymerase using template DNA derived from *L. mexicana* (genomic DNA and cDNA from each of the three life cycle stages). These reactions were diluted ten fold in sterile water and 1µl of each was used as the template for a second round of PCR, this time using primers OL77 and OL78. Both rounds of PCR were run as follows:

94°C, 4 minutes	x 1
94°C, 30 seconds 42°C, 30 seconds 72°C, 60 seconds	x 25 cycles
72°C, 4 minutes	x 1

The PCR product was expected to be approximately 350bp. Reactions were electrophoresed on 1.2% agarose gel and the gel was stained with ethidium bromide.

Lane 1: Genomic DNA template

Lane 2: Middle log-phase (promastigote) cDNA template

Lane 3: Stationary-phase (metacyclic promastigote) cDNA template

Lane 4: Lesion-derived amastigote cDNA template.

Markers: 1kb ladder (GibcoBRL)

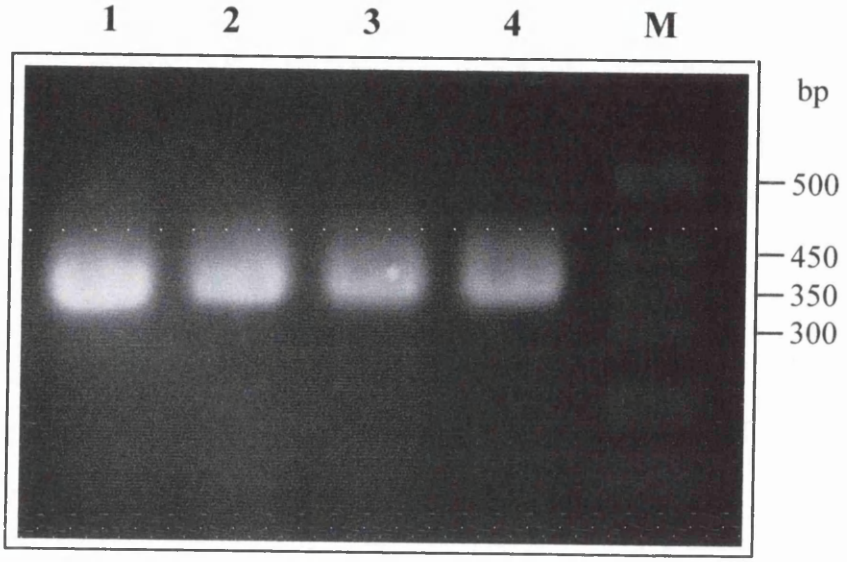

1
2
3
4
M
bp
500
450
350
300

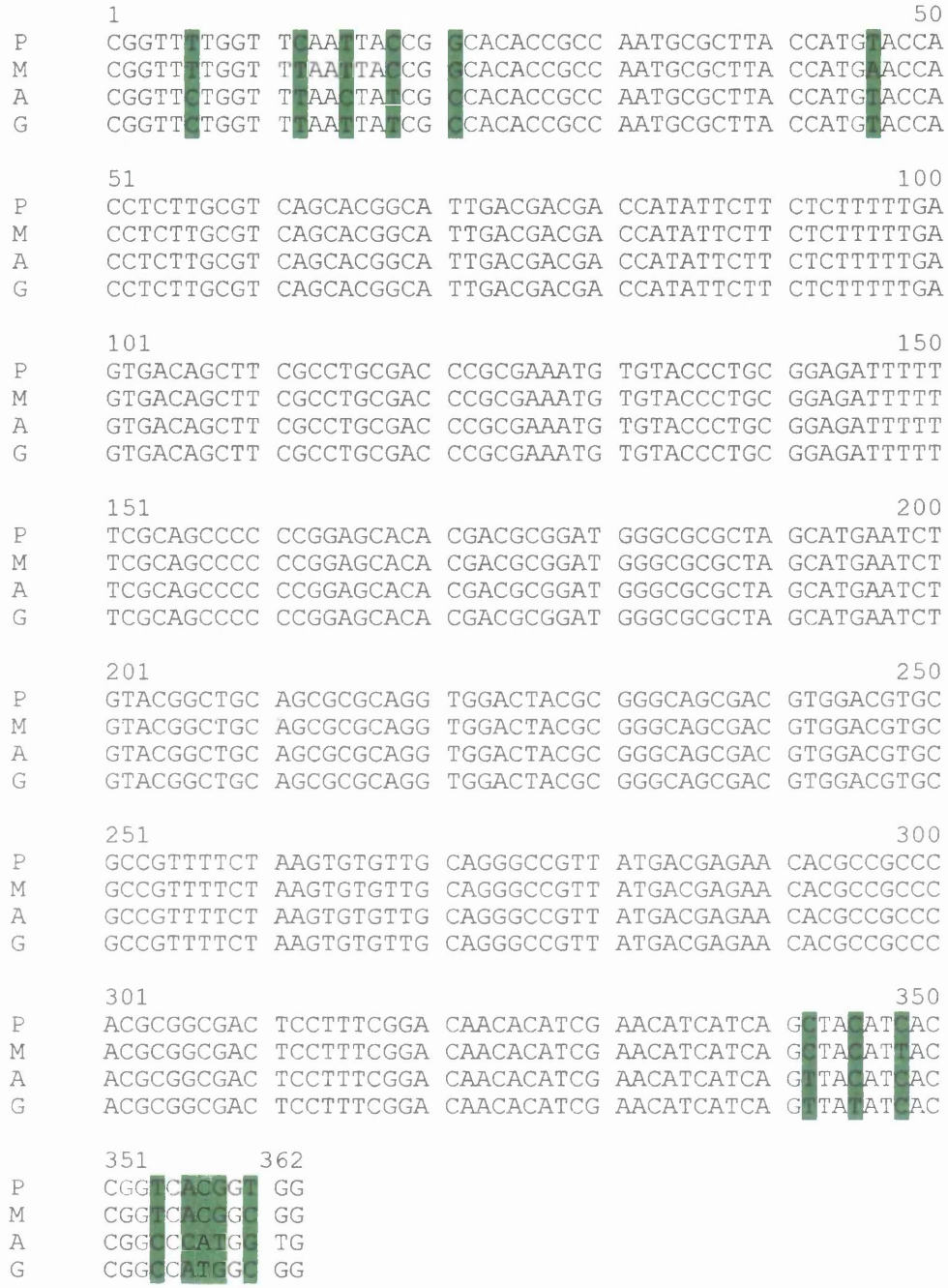
Figure 3.3: Alignment of the nucleotide sequences of four cloned *L. mexicana* PCR products. The PCR products from each of the four reactions were cloned into pTAG and sequenced with the T7 primer using automated sequencing with the AmpliTaq Ready Mix. Sequences were aligned using the Pileup programme from the GCG Software Package. Differences in the sequences are highlighted in green.

P = Promastigote

M = Metacyclic

A = Amastigote

G = Genomic


1 50
P CGGTTTTGGT TCAATTACCG GCACACCGCC AATGCGCTTA CCATGTACCA
M CGGTTTTGGT TTAATTACCG GCACACCGCC AATGCGCTTA CCATGAACCA
A CGGTTCTGGT TTAACTACCG CCACACCGCC AATGCGCTTA CCATGTACCA
G CGGTTCTGGT TTAATTACCG CCACACCGCC AATGCGCTTA CCATGTACCA
51 100
P CCTCTTGCGT CAGCACGGCA TTGACGACGA CCATATTCTT CTCTTTTTGA
M CCTCTTGCGT CAGCACGGCA TTGACGACGA CCATATTCTT CTCTTTTTGA
A CCTCTTGCGT CAGCACGGCA TTGACGACGA CCATATTCTT CTCTTTTTGA
G CCTCTTGCGT CAGCACGGCA TTGACGACGA CCATATTCTT CTCTTTTTGA
101 150
P GTGACAGCTT CGCCTGCGAC CCGCGAAATG TGTACCCTGC GGAGATTTTT
M GTGACAGCTT CGCCTGCGAC CCGCGAAATG TGTACCCTGC GGAGATTTTT
A GTGACAGCTT CGCCTGCGAC CCGCGAAATG TGTACCCTGC GGAGATTTTT
G GTGACAGCTT CGCCTGCGAC CCGCGAAATG TGTACCCTGC GGAGATTTTT
151 200
P TCGCAGCCCC CCGGAGCACA CGACGCGGAT GGGCGCGCTA GCATGAATCT
M TCGCAGCCCC CCGGAGCACA CGACGCGGAT GGGCGCGCTA GCATGAATCT
A TCGCAGCCCC CCGGAGCACA CGACGCGGAT GGGCGCGCTA GCATGAATCT
G TCGCAGCCCC CCGGAGCACA CGACGCGGAT GGGCGCGCTA GCATGAATCT
201 250
P GTACGGCTGC AGCGCGCAGG TGGACTACGC GGGCAGCGAC GTGGACGTGC
M GTACGGCTGC AGCGCGCAGG TGGACTACGC GGGCAGCGAC GTGGACGTGC
A GTACGGCTGC AGCGCGCAGG TGGACTACGC GGGCAGCGAC GTGGACGTGC
G GTACGGCTGC AGCGCGCAGG TGGACTACGC GGGCAGCGAC GTGGACGTGC
251 300
P GCCGTTTTCT AAGTGTGTTG CAGGGCCGTT ATGACGAGAA CACGCCGCCC
M GCCGTTTTCT AAGTGTGTTG CAGGGCCGTT ATGACGAGAA CACGCCGCCC
A GCCGTTTTCT AAGTGTGTTG CAGGGCCGTT ATGACGAGAA CACGCCGCCC
G GCCGTTTTCT AAGTGTGTTG CAGGGCCGTT ATGACGAGAA CACGCCGCCC
301 350
P ACGCGGCGAC TCCTTTCGGA CAACACATCG AACATCATCA GCTACATCAC
M ACGCGGCGAC TCCTTTCGGA CAACACATCG AACATCATCA GCTACATTAC
A ACGCGGCGAC TCCTTTCGGA CAACACATCG AACATCATCA GTTACATCAC
G ACGCGGCGAC TCCTTTCGGA CAACACATCG AACATCATCA GTTATATCAC
351 362
P CGGTCACGGT GG
M CGGTCACGGC GG
A CGGCCCATGG TG
G CGGCCATGGC GG

Sequences producing High-scoring Segment Pairs:		Reading Frame	High Score	Probability P(N)	N
gi|2558891	**(AF022913) GPI transamidase** [Homo...	+1	170	1.4e-34	2
gnl|PID|e259413	**(Y07596) gpi8 [Homo sapiens]**	+1	170	1.4e-34	2
sp|P49048|YJ96_CAEEL	**HYPOTHETICAL 35.7 KD PROTEIN T05E...**	+1	165	1.2e-32	2
gnl|PID|e349322	(Z68751) T05E11.6 [Caenorhabditis...	+1	165	1.5e-32	2
gnl|PID|e257726	(Z75551) T28H10.3 [Caenorhabditis...	+1	134	3.8e-17	2
sp|P49018|GPI8_YEAST	**GPI-ANCHOR TRANSMIDASE** /pir||S597...	+1	168	4.2e-16	1
gnl|PID|e286211	(Y09862) legumain [Homo sapiens]	+1	107	2.0e-10	2
gnl|PID|d1010173	(D55696) cysteine protease [Homo ...	+1	107	2.0e-10	2
sp|P49043|VPE_CITSI	VACUOLAR PROCESSING ENZYME PRECUR...	+1	106	1.8e-09	2
sp|P42665|HGLB_SCHJA	HEMOGLOBINASE PRECURSOR (ANTIGEN ...	+1	104	5.5e-09	2
sp|P09841|HGLB_SCHMA	HEMOGLOBINASE PRECURSOR (ANTIGEN ...	+1	102	5.6e-09	2
gnl|PID|d1019665	(D61395) gamma-VPE [Arabidopsis t...	+1	100	4.1e-07	2
gi|161019	(M21308) hemoglobinase [Schistoso...	+1	92	7.3e-07	2
gnl|PID|e354317	(Z99956) asparagine-specific endo...	+1	98	8.7e-07	2
sp|P49044|VPE_VICSA	VACUOLAR PROCESSING ENZYME PRECUR...	+1	100	1.8e-06	2
sp|P49042|VPE_RICCO	VACUOLAR PROCESSING ENZYME PRECUR...	+1	99	1.8e-06	2
pir||S60050	vacuolar processing enzyme (EC 3....	+1	95	2.7e-06	2
sp|P49046|LEGU_CANEN	LEGUMAIN PRECURSOR (ASPARAGINYL E...	+1	93	5.0e-05	2
sp|P49047|VPEA_ARATH	VACUOLAR PROCESSING ENZYME, ALPHA...	+1	87	5.0e-05	2
gnl|PID|e354316	(Z99957) legumain-like proteinase...	+1	91	5.2e-05	2
gnl|PID|e348871	(Z99174) cysteine proteinase prec...	+1	100	5.4e-05	1
pir||A57293	latent transforming growth factor...	-1	48	0.0011	4

Figure 3.4: BLAST search data showing proteins similar to that encoded by cloned *L. mexicana* PCR products. Data was submitted as nucleic acid sequence to the NCBI BLAST search web site (http://www.ncbi.nlm.nih.gov/BLAST) using the BLASTX programme, which translates nucleotide sequence into all six reading frames and compares the resulting amino acid sequences with those of known proteins. The closest matching proteins are named and the most relevant matches are shown in boldface.

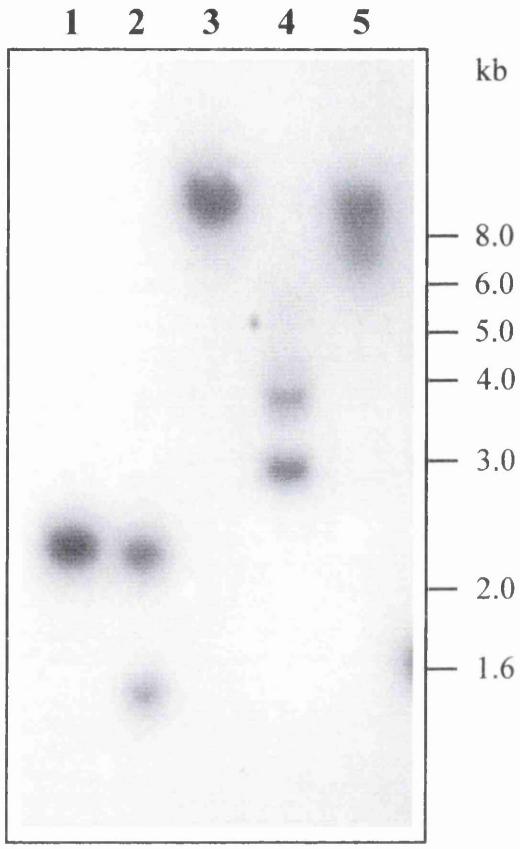


Figure 3.5: Southern blot of *L. mexicana* genomic digests probed with the cloned gDNA PCR product. 5µg of genomic DNA isolated from wild type *L. mexicana* was restriction digested, electrophoresed on 0.7% agarose TBE gel, blotted and probed with the 700bp, plasmid-derived *PvuII* fragment containing the cloned gDNA PCR fragment. The positions of the marker DNA bands (1kb ladder, GibcoBRL) are shown.

Lane 1: *Xho*I

Lane 2: *Pst*I

Lane 3: *Hin*dIII

Lane 4: *Hin*cII

Lane 5: *Eco*RI

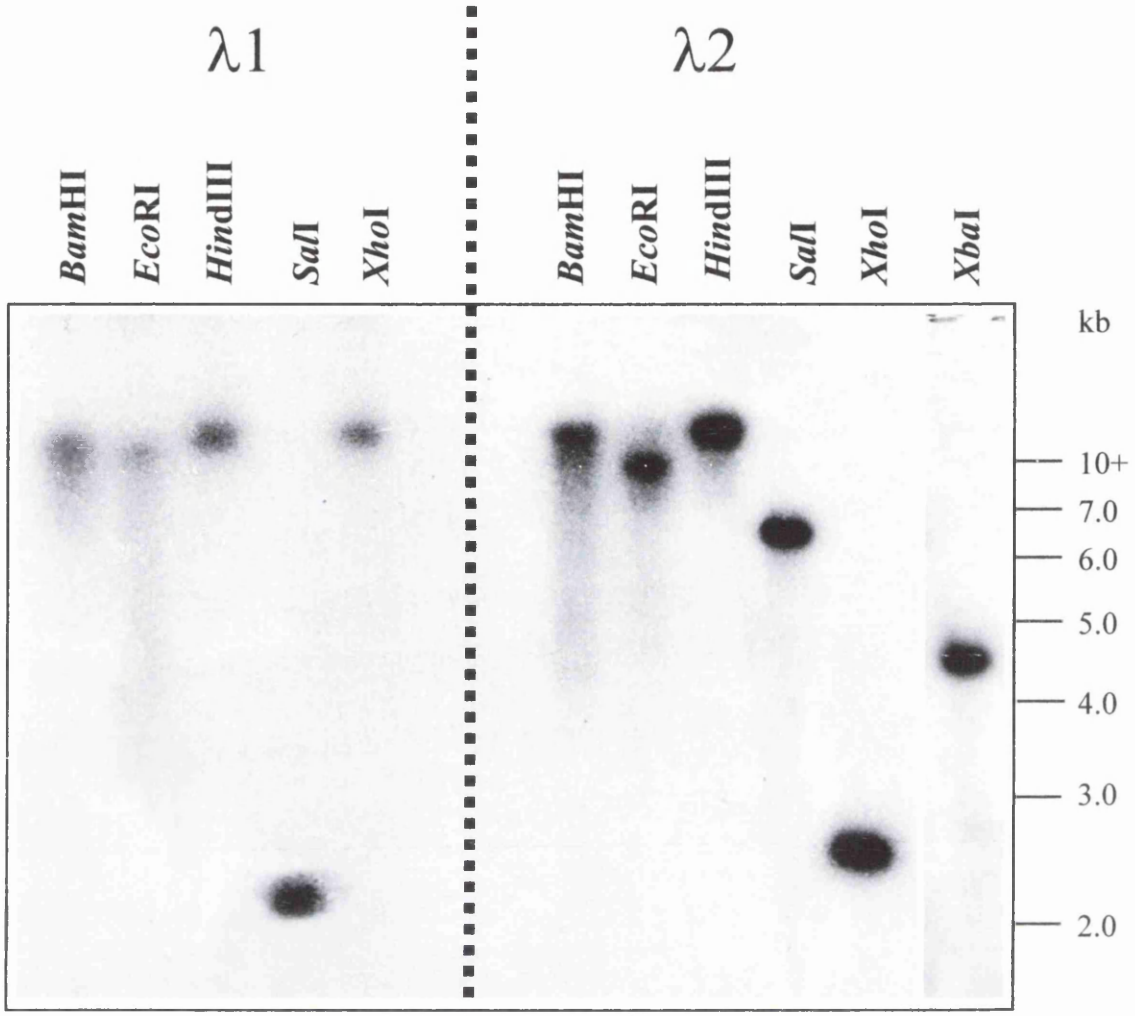
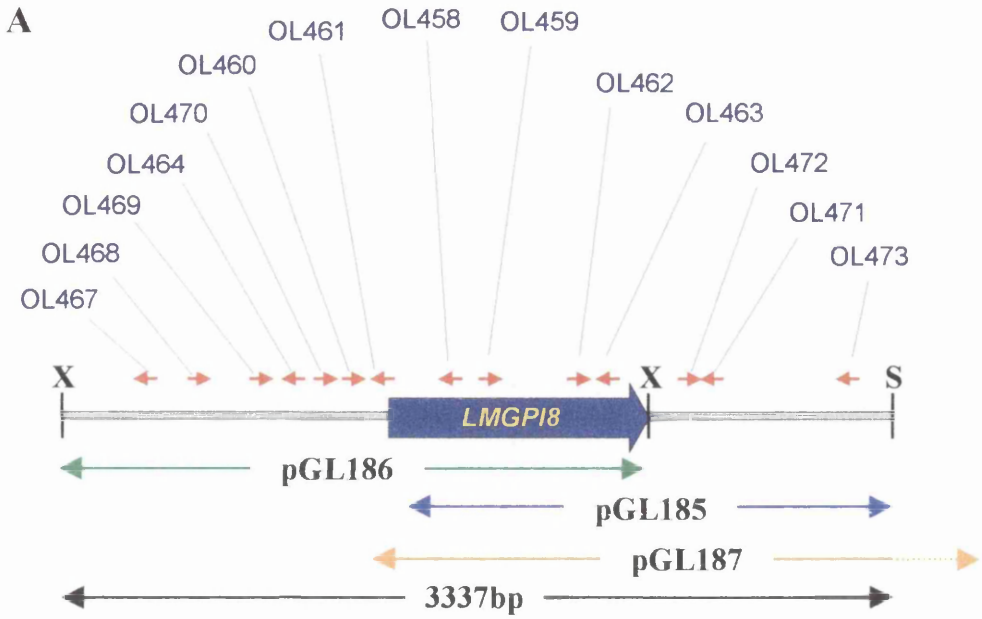


Figure 3.6: Southern blot of lambda clones containing *GPI8*. λ1 and λ2 clones were digested with the following restriction enzymes: *Bam*HI, *Eco*RI, *Hin*dIII, *Sal*I and *Xho*I before the fragments were electrophoresed on a 0.7% agarose TBE gel, blotted and probed with the plasmid-derived *Pvu*II fragment containing the cloned gDNA PCR product (described in Section 3.2.1.2). Markers were 1kb Ladder (GibcoBRL).

Figure 3.7: Strategy for sequencing of the *GPI8* locus. The position of the internal primers (OL458 - 473) used to obtain the sequence contigs, are shown as red arrows **(A)**. Other primers that anneal to plasmids on either side of inserted fragments were used to determine the remainder of the sequence (not shown). The distribution of *LMGPI8* and its flanking regions in pGL185, pGL186 and pGL187 are also indicated (**X** = *Xho*I site; **S** = *Sal*I site). The contig figure **(B)** was produced using the Gelmerge, Gelassemble and Gelview programmes from the GCG package, which takes multiple sequence fragments and determines the positions of overlap, producing a single consensus sequence of the entire nucleotide sequence. This has been overlaid with the schematic of the gene and its 5` and 3` flanking regions from **(A)**.

Large Blue Arrow = *GPI8* ORF


A
OL461
OL458
OL459
OL460
OL462
OL470
OL463
OL464
OL472
OL469
OL471
OL468
OL473
OL467
X
X
S
LMGPI8
pGL186
pGL185
pGL187
3337bp

B

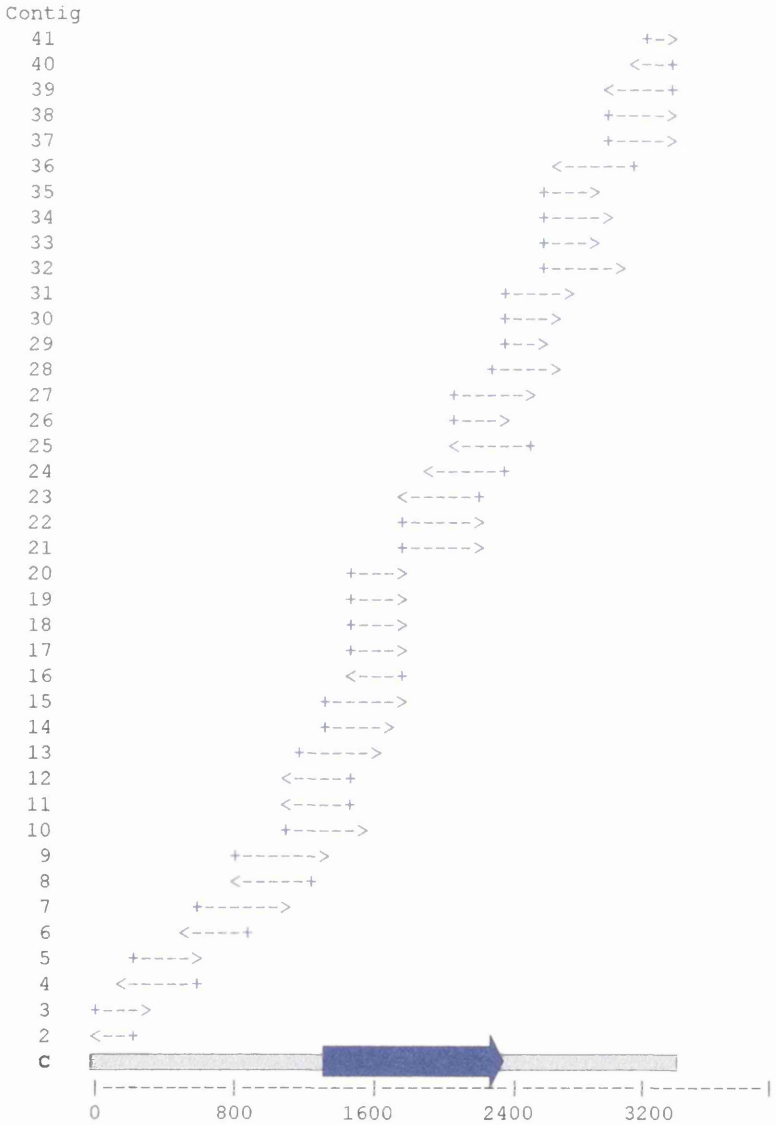
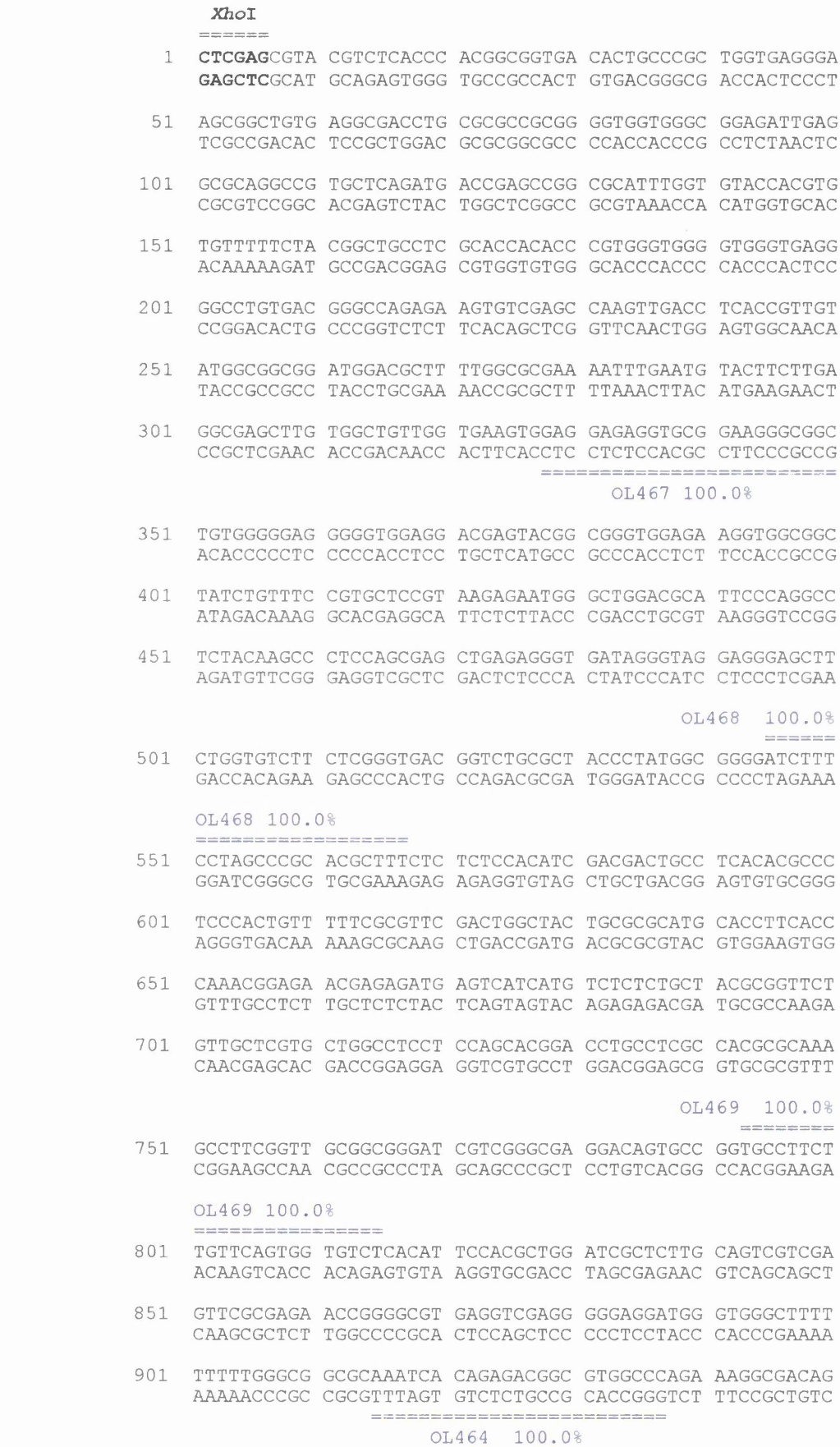
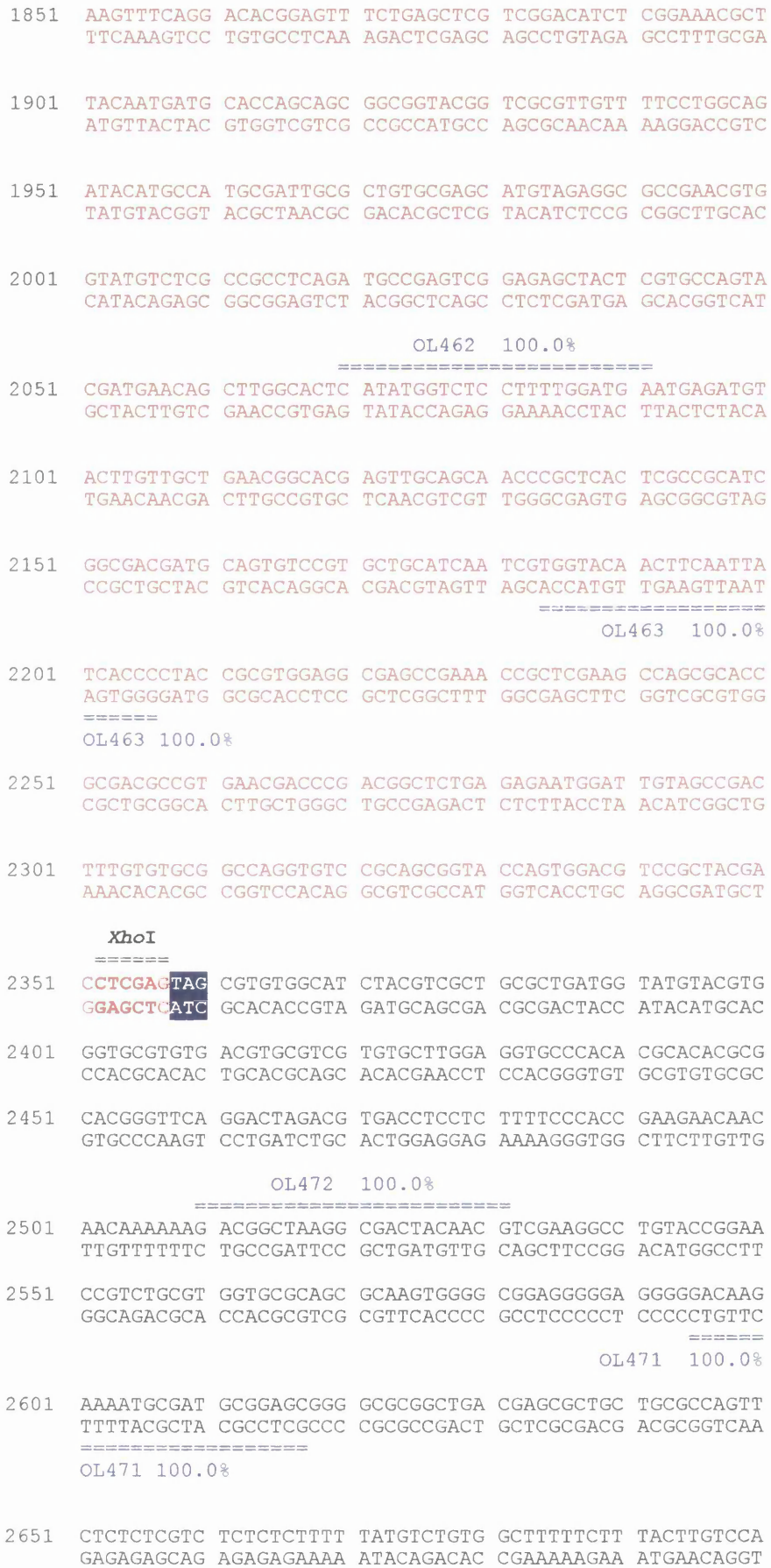

Contig
41
40
39
38
37
36
35
34
33
32
31
30
29
28
27
26
25
24
23
22
21
20
19
18
17
16
15
14
13
12
11
10
9
8
7
6
5
4
3
2
c
0
800
1600
2400
3200

Figure 3.8: The complete nucleotide sequence of *L. mexicana GPI8* locus and predicted amino acid sequence of the ORF. **(A)** The sequence is contained on a 3337bp *Xho*I-*Sal*I fragment (the restriction sites are marked). The positions of oligonucleotides used to obtain the sequence are shown in blue whilst the open reading frame of the gene is given in red. Note that the position of the stop codon is highlighted in dark blue. **(B)** The predicted ORF was translated using the Translate program (GCG, Wisconsin).

A


XhoI
1 CTCGAGCGTA CGTCTCACCC ACGGCGGTGA CACTGCCCGC TGGTGAGGGA
GAGCTCGCAT GCAGAGTGGG TGCCGCCACT GTGACGGGCG ACCACTCCCT
51 AGCGGCTGTG AGGCGACCTG CGCGCCGCGG GGTGGTGGGC GGAGATTGAG
TCGCCGACAC TCCGCTGGAC GCGCGGCGCC CCACCACCCG CCTCTAACTC
101 GCGCAGGCCG TGCTCAGATG ACCGAGCCGG CGCATTTGGT GTACCACGTG
CGCGTCCGGC ACGAGTCTAC TGGCTCGGCC GCGTAAACCA CATGGTGCAC
151 TGTTTTTCTA CGGCTGCCTC GCACCACACC CGTGGGTGGG GTGGGTGAGG
ACAAAAAGAT GCCGACGGAG CGTGGTGTGG GCACCCACCC CACCCACTCC
201 GGCCTGTGAC GGGCCAGAGA AGTGTCGAGC CAAGTTGACC TCACCGTTGT
CCGGACACTG CCCGGTCTCT TCACAGCTCG GTTCAACTGG AGTGGCAACA
251 ATGGCGGCGG ATGGACGCTT TTGGCGCGAA AATTTGAATG TACTTCTTGA
TACCGCCGCC TACCTGCGAA AACCGCGCTT TTAAACTTAC ATGAAGAACT
301 GGCGAGCTTG TGGCTGTTGG TGAAGTGGAG GAGAGGTGCG GAAGGGCGGC
CCGCTCGAAC ACCGACAACC ACTTCACCTC CTCTCCACGC CTTCCCGCCG
OL467 100.0%
351 TGTGGGGGAG GGGGTGGAGG ACGAGTACGG CGGGTGGAGA AGGTGGCGGC
ACACCCCCTC CCCCACCTCC TGCTCATGCC GCCCACCTCT TCCACCGCCG
401 TATCTGTTTC CGTGCTCCGT AAGAGAATGG GCTGGACGCA TTCCCAGGCC
ATAGACAAAG GCACGAGGCA TTCTCTTACC CGACCTGCGT AAGGGTCCGG
451 TCTACAAGCC CTCCAGCGAG CTGAGAGGGT GATAGGGTAG GAGGGAGCTT
AGATGTTCGG GAGGTCGCTC GACTCTCCCA CTATCCCATC CTCCCTCGAA
OL468 100.0%
501 CTGGTGTCTT CTCGGGTGAC GGTCTGCGCT ACCCTATGGC GGGGATCTTT
GACCACAGAA GAGCCCACTG CCAGACGCGA TGGGATACCG CCCCTAGAAA
OL468 100.0%
551 CCTAGCCCGC ACGCTTTCTC TCTCCACATC GACGACTGCC TCACACGCCC
GGATCGGGCG TGCGAAAGAG AGAGGTGTAG CTGCTGACGG AGTGTGCGGG
601 TCCCACTGTT TTTCGCGTTC GACTGGCTAC TGCGCGCATG CACCTTCACC
AGGGTGACAA AAAGCGCAAG CTGACCGATG ACGCGCGTAC GTGGAAGTGG
651 CAAACGGAGA ACGAGAGATG AGTCATCATG TCTCTCTGCT ACGCGGTTCT
GTTTGCCTCT TGCTCTCTAC TCAGTAGTAC AGAGAGACGA TGCGCCAAGA
701 GTTGCTCGTG CTGGCCTCCT CCAGCACGGA CCTGCCTCGC CACGCGCAAA
CAACGAGCAC GACCGGAGGA GGTCGTGCCT GGACGGAGCG GTGCGCGTTT
OL469 100.0%
751 GCCTTCGGTT GCGGCGGGAT CGTCGGGCGA GGACAGTGCC GGTGCCTTCT
CGGAAGCCAA CGCCGCCCTA GCAGCCCGCT CCTGTCACGG CCACGGAAGA
OL469 100.0%
801 TGTTCAGTGG TGTCTCACAT TCCACGCTGG ATCGCTCTTG CAGTCGTCGA
ACAAGTCACC ACAGAGTGTA AGGTGCGACC TAGCGAGAAC GTCAGCAGCT
851 GTTCGCGAGA ACCGGGGCGT GAGGTCGAGG GGGAGGATGG GTGGGCTTTT
CAAGCGCTCT TGGCCCCGCA CTCCAGCTCC CCCTCCTACC CACCCGAAAA
901 TTTTTGGGCG GCGCAAATCA CAGAGACGGC GTGGCCCAGA AAGGCGACAG
AAAAACCCGC CGCGTTTAGT GTCTCTGCCG CACCGGGTCT TTCCGCTGTC
OL464 100.0%

```
 951  CGCGCGAGTT GCACCGCTTG TGCACATGTG CCGTTACGGC GTTGATGTGA
      GCGCGCTCAA CGTGGCGAAC ACGTGTACAC GGCAATGCCG CAACTACACT

                                                 OL470  100.0%
                                                    ===
1001  GGAGTGCGTG TACGGTGCCA CTGCCGGTGC GCCGATGGGC TTGCCCAGTC
      CCTCACGCAC ATGCCACGGT GACGGCCACG CGGCTACCCG AACGGGTCAG

         OL470 100.0%
      ========================
1051  CCTCTGCTTC CCTTCCTCAT CTGATCCCTT TCTCCTCTAC CTCGCTCTCG
      GGAGACGAAG GGAAGGAGTA GACTAGGGAA AGAGGAGATG GAGCGAGAGC

1101  CTCACGCACG GAAGCTCCTT TTTTCCCTTT CGCTGGCTCT CCCTCGTTCT
      GAGTGCGTGC CTTCGAGGAA AAAAGGGAAA GCGACCGAGA GGGAGCAAGA

                                      OL460 100.0%
                                    ======================
1151  CCCGCTACTG GTGCGTGCCA CCTCTCTCTT GCGAATCGCG CAGACGGCGA
      GGGCGATGAC CACGCACGGT GGAGAGAGAA CGCTTAGCGC GTCTGCCGCT

1201  CACAGCTCAT CAGCCGCGGC AGCAGAAGGT AGTGGGAGAG GGGCGTGCGG
      GTGTCGAGTA GTCGGCGCCG TCGTCTTCCA TCACCCTCTC CCCGCACGCC

1251  GTGTGCCGGC TCTTCCTTTC ATATTCTAGA AAACGGCAGC TAGTGCAGTT
      CACACGGCCG AGAAGGAAAG TATAAGATCT TTTGCCGTCG ATCACGTCAA
                            ========================
                                  OL461  100.0%

1301  CCGGCTGTTT ATGCGCACCA CAGCTTACGT GATGACGTCG CCAACAAGGT
      GGCCGACAAA TACGCGTGGT GTCGAATGCA CTACTGCAGC GGTTGTTCCA

1351  GTATCGCGAC GGCATTGATC GTCTTTGCTT TTCTTGTTCT CACCGCCGCG
      CATAGCGCTG CCGTAACTAG CAGAAACGAA AAGAACAAGA GTGGCGGCGC

1401  GCCGCCGCGT CTGCGCCGCT GGGGGCGACA GGCAAGGGCC AGAGCAACAA
      CGGCGGCGCA GACGCGGCGA CCCCCGCTGT CCGTTCCCGG TCTCGTTGTT

1451  CTGGGCTGTC ATTGTCTCCT CCTCGCGATA CCTCTTCAAC TACCGCCACA
      GACCCGACAG TAACAGAGGA GGAGCGCTAT GGAGAAGTTG ATGGCGGTGT

1501  CCGCCAATGC GCTTACCATG TACCACCTCT TGCGTCAGCA CGGCATTGAC
      GGCGGTTACG CGAATGGTAC ATGGTGGAGA ACGCAGTCGT GCCGTAACTG

1551  GACGACCATA TTCTTCTCTT TTTGAGTGAC AGCTTCGCCT GCGACCCGCG
      CTGCTGGTAT AAGAAGAGAA AAACTCACTG TCGAAGCGGA CGCTGGGCGC
                                  ========================
                                        OL458  100.0%

1601  AAATGTGTAC CCTGCGGAGA TTTTTTCGCA GCCCCCCGGA GCACACGACG
      TTTACACATG GGACGCCTCT AAAAAAGCGT CGGGGGGCCT CGTGTGCTGC

1651  CGGATGGGCG CGCTAGCATG AATCTGTACG GCTGCAGCGC GCAGGTGGAC
      GCCTACCCGC GCGATCGTAC TTAGACATGC CGACGTCGCG CGTCCACCTG

                                            OL459 100.0%
                                          ======================
1701  TACGCGGGCA GCGACGTGGA CGTGCGCCGT TTTCTAAGTG TGTTGCAGGG
      ATGCGCCCGT CGCTGCACCT GCACGCGGCA AAAGATTCAC ACAACGTCCC

1751  CCGTTATGAC GAGAACACGC CGCCCACGCG GCGACTCCTT TCGGACAACA
      GGCAATACTG CTCTTGTGCG GCGGGTGCGC CGCTGAGGAA AGCCTGTTGT

1801  CATCGAACAT CATCATCTAC GTCGCGGGGC ACGGCGCCAA GTCGTACTTC
      GTAGCTTGTA GTAGTAGATG CAGCGCCCCG TGCCGCGGTT CAGCATGAAG
```


1851 AAGTTTCAGG ACACGGAGTT TCTGAGCTCG TCGGACATCT CGGAAACGCT
TTCAAAGTCC TGTGCCTCAA AGACTCGAGC AGCCTGTAGA GCCTTTGCGA
1901 TACAATGATG CACCAGCAGC GGCGGTACGG TCGCGTTGTT TTCCTGGCAG
ATGTTACTAC GTGGTCGTCG CCGCCATGCC AGCGCAACAA AAGGACCGTC
1951 ATACATGCCA TGCGATTGCG CTGTGCGAGC ATGTAGAGGC GCCGAACGTG
TATGTACGGT ACGCTAACGC GACACGCTCG TACATCTCCG CGGCTTGCAC
2001 GTATGTCTCG CCGCCTCAGA TGCCGAGTCG GAGAGCTACT CGTGCCAGTA
CATACAGAGC GGCGGAGTCT ACGGCTCAGC CTCTCGATGA GCACGGTCAT
OL462 100.0%
2051 CGATGAACAG CTTGGCACTC ATATGGTCTC CTTTTGGATG AATGAGATGT
GCTACTTGTC GAACCGTGAG TATACCAGAG GAAAACCTAC TTACTCTACA
2101 ACTTGTTGCT GAACGGCACG AGTTGCAGCA ACCCGCTCAC TCGCCGCATC
TGAACAACGA CTTGCCGTGC TCAACGTCGT TGGGCGAGTG AGCGGCGTAG
2151 GGCGACGATG CAGTGTCCGT GCTGCATCAA TCGTGGTACA ACTTCAATTA
CCGCTGCTAC GTCACAGGCA CGACGTAGTT AGCACCATGT TGAAGTTAAT
OL463 100.0%
2201 TCACCCCTAC CGCGTGGAGG CGAGCCGAAA CCGCTCGAAG CCAGCGCACC
AGTGGGGATG GCGCACCTCC GCTCGGCTTT GGCGAGCTTC GGTCGCGTGG
OL463 100.0%
2251 GCGACGCCGT GAACGACCCG ACGGCTCTGA GAGAATGGAT TGTAGCCGAC
CGCTGCGGCA CTTGCTGGGC TGCCGAGACT CTCTTACCTA ACATCGGCTG
2301 TTTGTGTGCG GCCAGGTGTC CGCAGCGGTA CCAGTGGACG TCCGCTACGA
AAACACACGC CGGTCCACAG GCGTCGCCAT GGTCACCTGC AGGCGATGCT
XhoI
2351 CCTCGAGTAG CGTGTGGCAT CTACGTCGCT GCGCTGATGG TATGTACGTG
GGAGCTCATC GCACACCGTA GATGCAGCGA CGCGACTACC ATACATGCAC
2401 GGTGCGTGTG ACGTGCGTCG TGTGCTTGGA GGTGCCCACA CGCACACGCG
CCACGCACAC TGCACGCAGC ACACGAACCT CCACGGGTGT GCGTGTGCGC
2451 CACGGGTTCA GGACTAGACG TGACCTCCTC TTTTCCCACC GAAGAACAAC
GTGCCCAAGT CCTGATCTGC ACTGGAGGAG AAAAGGGTGG CTTCTTGTTG
OL472 100.0%
2501 AACAAAAAAG ACGGCTAAGG CGACTACAAC GTCGAAGGCC TGTACCGGAA
TTGTTTTTTC TGCCGATTCC GCTGATGTTG CAGCTTCCGG ACATGGCCTT
2551 CCGTCTGCGT GGTGCGCAGC GCAAGTGGGG CGGAGGGGGA GGGGGACAAG
GGCAGACGCA CCACGCGTCG CGTTCACCCC GCCTCCCCCT CCCCCTGTTC
OL471 100.0%
2601 AAAATGCGAT GCGGAGCGGG GCGCGGCTGA CGAGCGCTGC TGCGCCAGTT
TTTTACGCTA CGCCTCGCCC CGCGCCGACT GCTCGCGACG ACGCGGTCAA
OL471 100.0%
2651 CTCTCTCGTC TCTCTCTTTT TATGTCTGTG GCTTTTTCTT TACTTGTCCA
GAGAGAGCAG AGAGAGAAAA ATACAGACAC CGAAAAAGAA ATGAACAGGT

```
2701  TGTGAAGTAC ACGTCTCCAC ATGCTTTCCC AGCAGCATGT ATGGGGGTGT
      ACACTTCATG TGCAGAGGTG TACGAAAGGG TCGTCGTACA TACCCCCACA

2751  AGCGTCGCAA ACCCGTGGTG GCGATGACTC CTGGTCGGCT TTCATGTTGC
      TCGCAGCGTT TGGGCACCAC CGCTACTGAG GACCAGCCGA AAGTACAACG

2801  CTTCAGGGGG GTGGGGTGGG GAGCATACGG GCACTCACTC GTACACGCGG
      GAAGTCCCCC CACCCCACCC CTCGTATGCC CGTGAGTGAG CATGTGCGCC

2851  AACGGCTCCC AATCTTGTGG CTCAGCGCGC CGTCTGCCTC TGTGTGGGGC
      TTGCCGAGGG TTAGAACACC GAGTCGCGCG GCAGACGGAG ACACACCCCG

2901  GATGCGTGGA ACCTGGTACC CATCTCTCTC TCTGTCTTCG ACCGCGCACA
      CTACGCACCT TGGACCATGG GTAGAGAGAG AGACAGAAGC TGGCGCGTGT

2951  ATCACCTCAC TATGTTTCGA TATCTTCCTC GCCCACTTTG CTGATTTTTC
      TAGTGGAGTG ATACAAAGCT ATAGAAGGAG CGGGTGAAAC GACTAAAAAG

3001  ACCCGTGGCC TTCCGAATCA CCCATGCGCA CCCCAAAATG GGATGCACCA
      TGGGCACCGG AAGGCTTAGT GGGTACGCGT GGGGTTTTAC CCTACGTGGT

3051  GCGAAGAGCG TTCGTTGGCT TCCTTTTCTA ATGCGCGCGT TTACGCTAGC
      CGCTTCTCGC AAGCAACCGA AGGAAAAGAT TACGCGCGCA AATGCGATCG

3101  TACTGGTGCG GCGCCGCGCC AACTCTTGGC CGTGGCGATG GCGACCGCAC
      ATGACCACGC CGCGGCGCGG TTGAGAACCG GCACCGCTAC CGCTGGCGTG

3151  GCATGTCGGC GTCATCGAGC ACGCCTGCTG AATCCTCCAC CCCAGTCGGT
      CGTACAGCCG CAGTAGCTCG TGCGGACGAC TTAGGAGGTG GGGTCAGCCA
                            ==========================
                                  OL473  100.0%

3201  GCGTTGCAAG CTGCAGCGGC GACGGGGCCG ACCGTGGCGC CGAAGAAGCC
      CGCAACGTTC GACGTCGCCG CTGCCCCGGC TGGCACCGCG GCTTCTTCGG

3251  ACGCCGTCGT ATCACGAGCA AGCTCGACGC YTCTGCCCTG CTTAAGAATG
      TGCGGCAGCA TAGTGCTCGT TCGAGCTGCG RAGACGGGAC GAATTCTTAC

                                              SalI
                                             ======
3301  TGCCGAGCGA ACCCGGCGCG CATGCGGAGC TGTCGAC
      ACGGCTCGCT TGGGCCGCGC GTACGCCTCG ACAGCTG
```

B

Amino acid sequence of LMGPI8

```
  1  MRTTAYVMTS PTRCIATALI VFAFLVLTAA AAASAPLGAT GKGQSNNWAV

 51  IVSSSRYLFN YRHTANALTM YHLLRQHGID DDHILLFLSD SFACDPRNVY

101  PAEIFSQPPG AHDADGRASM NLYGCSAQVD YAGSDVDVRR FLSVLQGRYD

151  ENTPPTRRLL SDNTSNIIIY VAGHGAKSYF KFQDTEFLSS SDISETLTMM

201  HQQRRYGRVV FLADTCHAIA LCEHVEAPNV VCLAASDAES ESYSCQYDEQ

251  LGTHMVSFWM NEMYLLLNGT SCSNPLTRRI GDDAVSVLHQ SWYNFNYHPY

301  RVEASRNRSK PAHRDAVNDP TALREWIVAD FVCGQVSAAV PVDVRYDLE
```

Figure 3.9: Amino acid sequence alignment of the complete predicted amino acid sequence of LMGPI8 with *S. cerevisiae* and *H. sapiens* GPI8 and legumain. Coloured arrows indicate the conserved regions to which the original PCR primers were designed (OL76, OL77, JATR3 and OL78). Conserved cysteine (*), histidine (#) and serine (^) residues are indicated. *L. mexicana* GPI8 has three potential N-glycosylation sites at positions N^{163}, $N^{268}$and N^{307} (+). Amino acid numbering is according to the LMGPI8 sequence. All of the proteins compared are larger than LMGPI8 and are only shown up to amino acid 349.

Abbreviations:

Lm = *L. mexicana* GPI8

Sc = *S. cerevisiae* GPI8

Hs = *H. sapiens* GPI8

Lg = Legumain (*Canavalia ensiformis* asparaginyl endopeptidase)

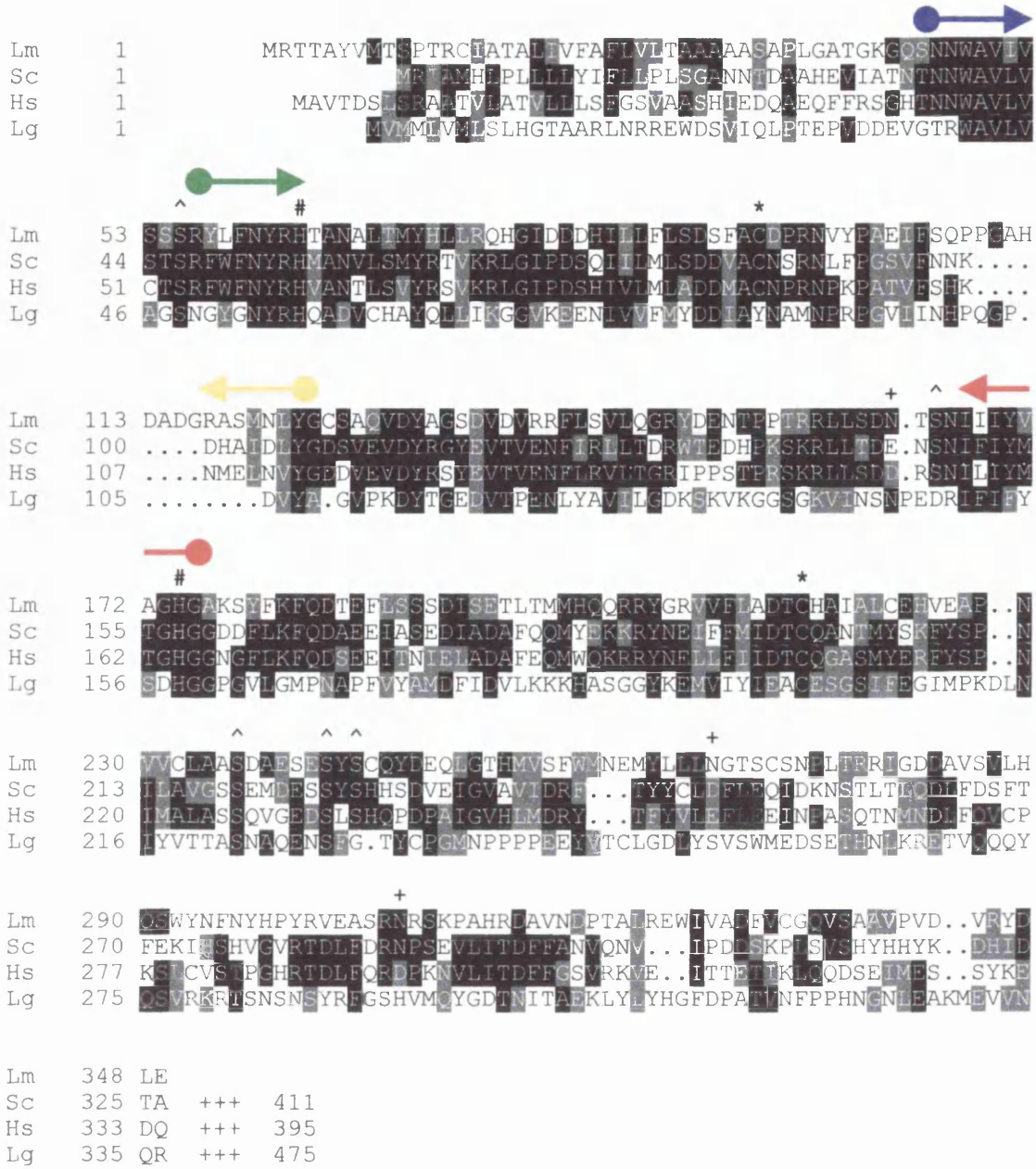

Lm 1 MRTTAYVMTSPTRCIATALIVFAFLVLTAAAAASAPLGATGKGQSNNWAVLV
Sc 1 MRTAMHLPLLLLYIFLLPLSGANNIDAAHEVIATNTNNWAVLV
Hs 1 MAVTDSLSRAAIVLATVLLLSFGSVAASHIEDQAEQFFRSGHTNNWAVLV
Lg 1 MVMMLVMLSLHGTAARLNRREWDSVIQLPTEPVDDEVGTRWAVLV
Lm 53 SSSRYLFNYRHTANALIMYHLLRQHGIDDDHILIFLSDSFACDPRNVYPAEIFSQPPGAH
Sc 44 STSRFWFNYRHMANVLSMYRTVKRLGIPDSQIILMLSDDVACNSRNLFPGSVFNNK....
Hs 51 CTSRFWFNYRHVANTLSVYRSVKRLGIPDSHIVLMLADDMACNPRNPKPATVFSHK....
Lg 46 AGSNGYGNYRHQADVCHAYQLLIKGGVKEENIVVFMYDDIAYNAMNPRPGVIINHPQGP.
Lm 113 DADGRASMNLYGCSAQVDYAGSDVDVRRFLSVLQGRYDENTPPTRRLLSDN.TSNIIIYV
Sc 100 ....DHAIDLYGDSVEVDYRGYEVTVENFIRLLTDRWTEDHPKSKRLLTDE.NSNIFIYM
Hs 107 ....NMELNVYGEDVEVDYRSYEVTVENFLRVLTGRIPPSTPRSKRLLSDD.RSNILIYM
Lg 105 ....DVYA.GVPKDYTGEDVTPENLYAVILGDKSKVKGGSGKVINSNPEDRIFIFY
Lm 172 AGHGAKSYFKFQDTEFLSSSDISETLTMMHQQRRYGRVVFLADTCHAIALCEHVEAF..N
Sc 155 TGHGGDDFLKFQDAEEIASEDIADAFQQMYEKKRYNEIFFMIDTCQANTMYSKFYSP..N
Hs 162 TGHGGNGFLKFQDSEEITNIELADAFEQMWQKRRYNELLFIIDTCQGASMYERFYSP..N
Lg 156 SDHGGPGVLGMPNAPFVYAMDFILVLKKKFASGGYKEMVIYIEACESGSIFEGIMPKDLN
Lm 230 VVCLAASDAESESYSCQYDEQLGTHMVSFWWNEMYLLLNGTSCSNPLTRRIGDQAVSILH
Sc 213 ILAVGSSEMDESSYSHHSDVEIGVAVIDRF...TYYCLDFLEQIDKNSTLTEQDLFDSFT
Hs 220 IMALASSQVGEDSLSHQPDPAIGVHLMDRY...TFYVLEFLEEINPASQTNMNDLFDVCP
Lg 216 IYVTTASNAQENSFG.TYCPGMNPPPPEEYYTCLGDLYSVSWMEDSETHNLKRETVLQQY
Lm 290 QEWYNFNYHPYRVEASRNRSKPAHRDAVNLPTALREWIVALFVCGQVSAAVPVD..VRYD
Sc 270 FEKIESHVGVRTDLFDRNPSEVLITDFFANVQNV...IPDLSKPISVSHYHHYK..DHID
Hs 277 KSICVSTPGHRTDLFQRDPKNVLITDFFGSVRKVE..ITTETIKLQQDSEIMES..SYKE
Lg 275 QSVRKRTSNSNSYRFGSHVMQYGDTNITAEKLYLYHGFDPATINFPPHNGNLEAKMEVVN
Lm 348 LE
Sc 325 TA +++ 411
Hs 333 DQ +++ 395
Lg 335 QR +++ 475

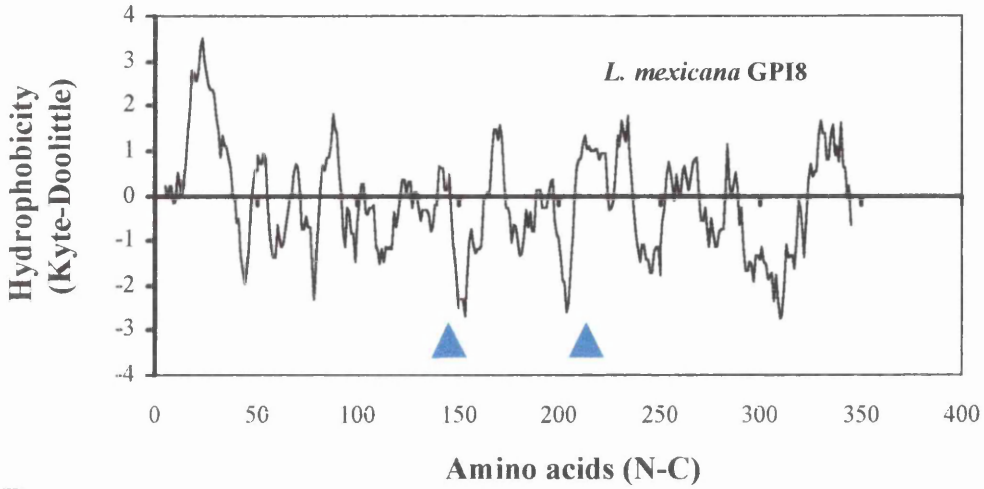
Figure 3.10: Hydrophobicity plot of GPI8 proteins. The plots were calculated by the "Prot-Scale" programme from the EXPASY Internet site (see text for details) according to the Kyte-Doolittle algorithm. The data was entered into Microsoft Excel which was used to plot the profile graphically with Amino Acid Scores (which are either greater than zero for hydrophobic residues and less than zero for hydrophilic) are given on the y-axis while the amino acid positions (from the amino to the carboxy terminus) are represented on the x-axis. Areas of notable similarity are indicated between blue arrows.

A: *L.mexicana* GPI8

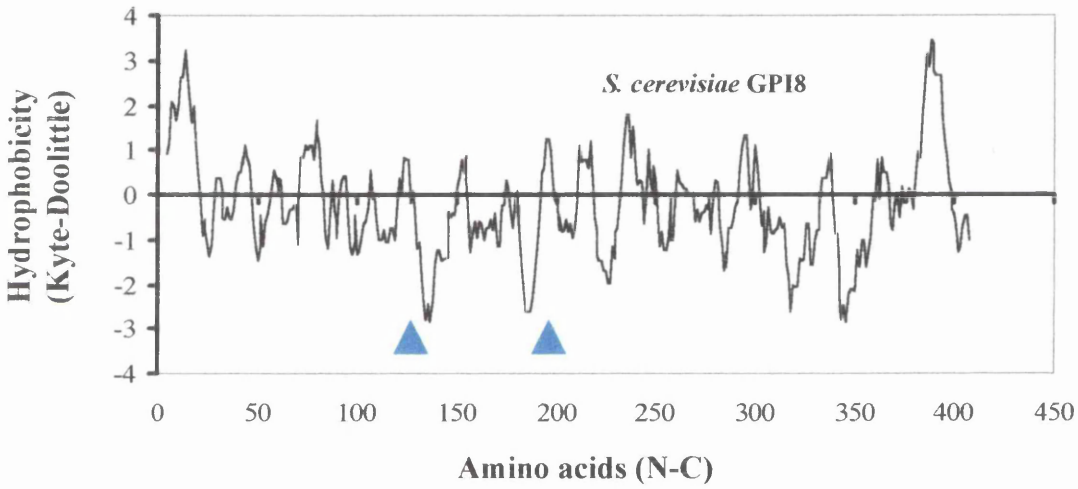
B: *S. cerevisiae* Gpi8p

C: *H. sapiens* GPI8

A


Hydrophobicity (Kyte-Doolittle)
L. mexicana GPI8
4
3
2
1
0
-1
-2
-3
-4
0
50
100
150
200
250
300
350
400
Amino acids (N-C)

B


Hydrophobicity (Kyte-Doolittle)
S. cerevisiae GPI8
4
3
2
1
0
-1
-2
-3
-4
0
50
100
150
200
250
300
350
400
450
Amino acids (N-C)

C

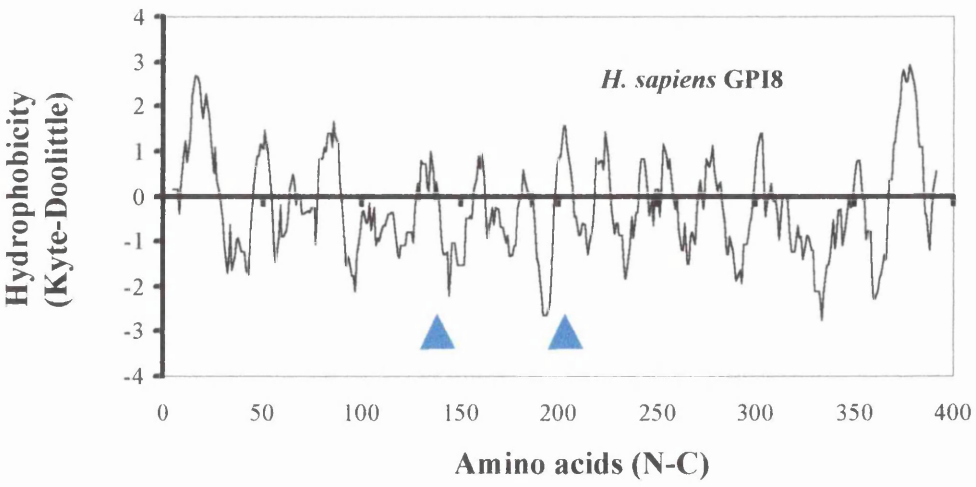

Hydrophobicity (Kyte-Doolittle)
H. sapiens GPI8
4
3
2
1
0
-1
-2
-3
-4
0
50
100
150
200
250
300
350
400
Amino acids (N-C)

Figure 3.11: Southern analysis of *L. mexicana* DNA using the *LMGPI8* ORF as probe. 5μg of DNA was digested with *Hin*cII, *Hin*dIII, *Pst*I, *Xba*I and *Xho*I. A schematic of the gene indicating the relevant restriction sites is shown **(A)**. The digests were electrophoresed on a 0.7% agarose gel, blotted onto Hybond N and probed with the *LMGPI8* ORF contained on a 1.1kb *Xba*I/*Xho*I fragment **(B)**.

Lane 1: *Hin*cII

Lane 2: *Hin*dIII

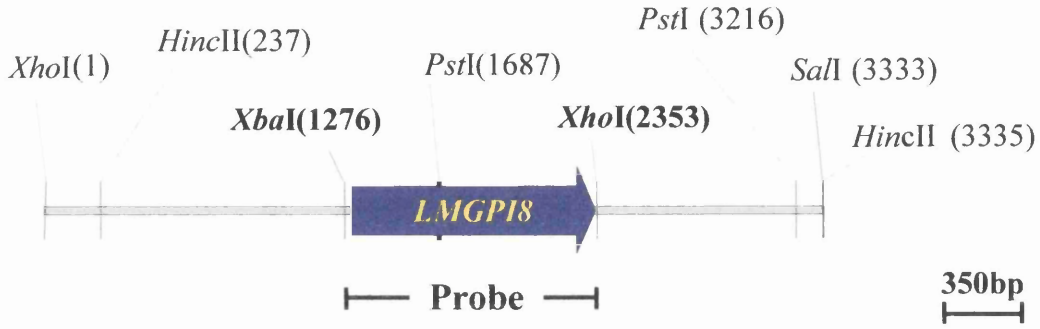
Lane 3: *Pst*I

Lane 4: *Xba*I

Lane 5: *Xho*I

The positions of the DNA marker fragments (1kb ladder, GibcoBRL) are shown.

A


XhoI(1)
HincII(237)
XbaI(1276)
PstI(1687)
XhoI(2353)
PstI (3216)
SalI (3333)
HincII (3335)
LMGPI8
Probe
350bp

B

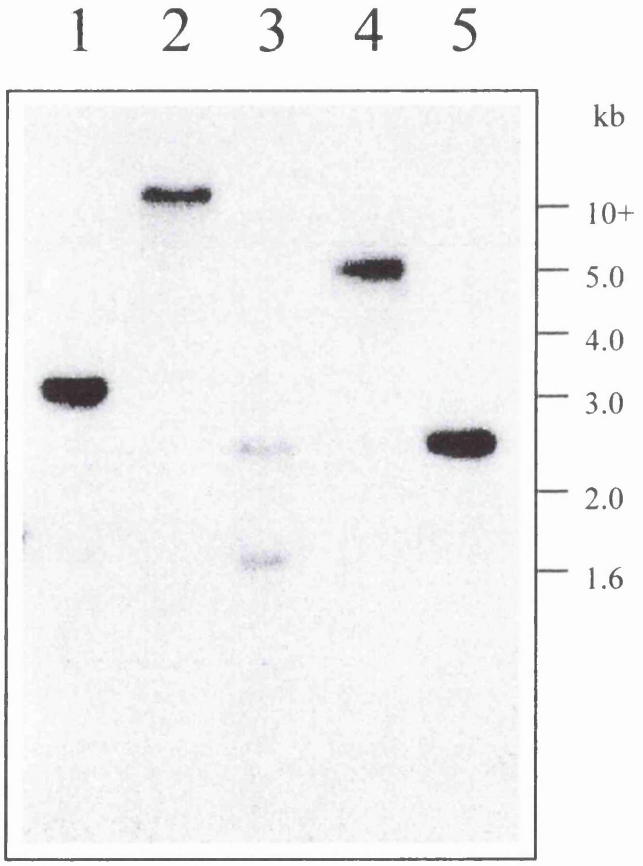

1 2 3 4 5
kb
10+
5.0
4.0
3.0
2.0
1.6

Figure 3.12: Southern blot analysis of restriction digested λ DNA containing *TBGPI8* using *LMGPI8* as a heterologous probe. 0.5μg of each of the two purified λ DNA clones was digested with *Hin*dIII, *Sal*I, *Xba*I and *Xho*I and electrophoresed on a 0.8% agarose gel, which was stained with ethidium bromide **(A)** prior to blotting onto Hybond N. This was done to determine if the two λ clones were identical.

Lane 1: λ**1**, *Hin*dIII

Lane 2: λ**1**, *Sal*I

Lane 3: λ**1**, *Xba*I

Lane 4: λ**1**, *Xho*I

M: DNA markers, 1 kb ladder (GibcoBRL)

Lane 5: λ**2**, *Hin*dIII

Lane 6: λ**2**, *Sal*I

Lane 7: λ**2**, *Xba*I

Lane 8: λ**2**, *Xho*I

(B) The blot of λ2 was probed with the 1.1kb *Xba*I/*Xho*I fragment containing *L. mexicana GPI8* ORF at lower stringency (55°C, 2xSSC, 0.1% SDS).

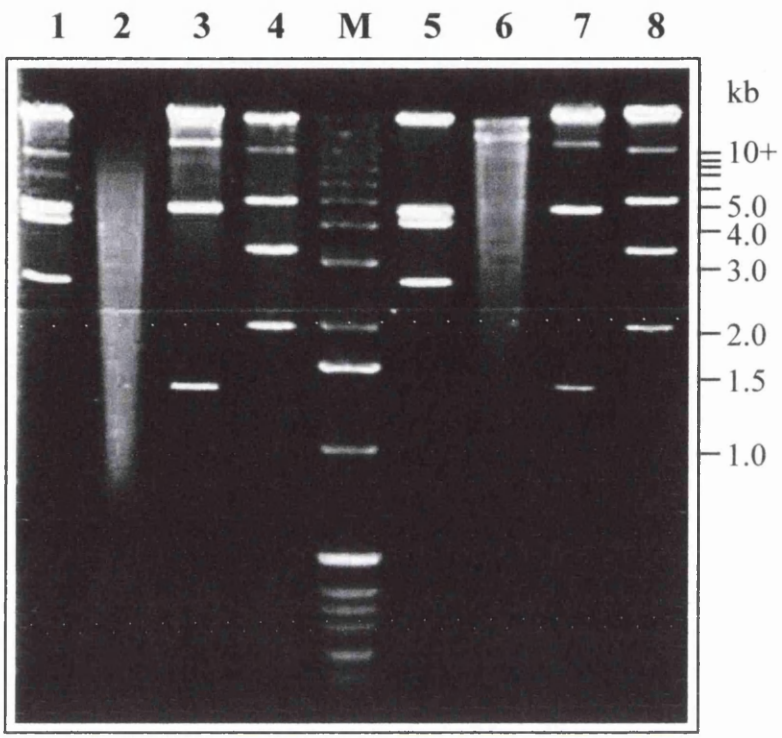
Lane 1: λ**2**, *Hin*dIII

Lane 2: λ**2**, *Sal*I

Lane 3: λ**2**, *Xba*I

Lane 4: λ**2**, *Xho*I

A



B

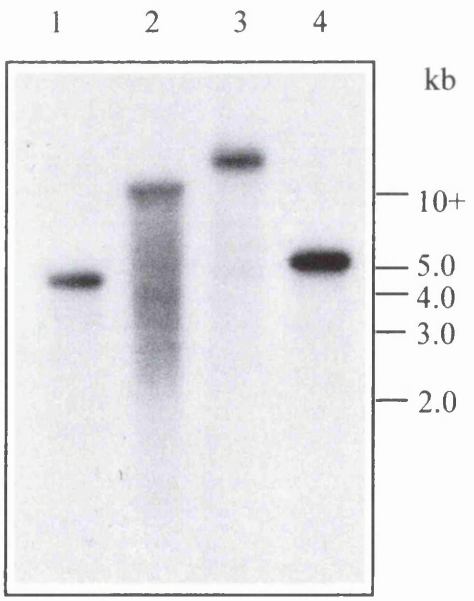
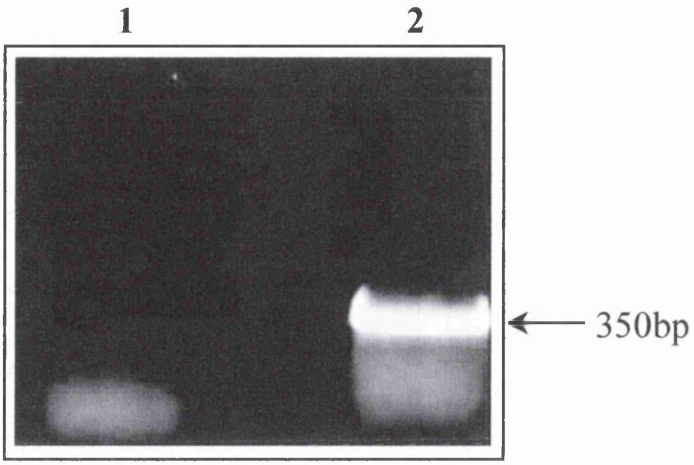


Figure 3.13: Characterisation of the PCR product from *T. brucei* λ clone 1. The PCR was carried out with two sets of primer pairs on DNA purified from *T. brucei* λ clone 1 that hybridised to the *LMGPI8* ORF. **(A)** Primers OL47 and OL48 (lane 1) and OL76 and OL48 (lane 2) were used in this experiment (see Tables 3.1 and 3.4 for details of primers). **(B)** A cloned PCR product (in pGL446) was sequenced in double stranded form using the universal sequencing primers T7 and SP6. Translation of this nucleotide sequence produced an ORF and the amino acid sequence is given. Primer positions are indicated by arrows (Blue = OL76 and Red = OL48). **(C)** A BLAST search for similar proteins using the BLASTX programme at the NCBI web site (http://www.ncbi.nlm.nih.gov/BLAST/) shows the most closely related amino acid sequences to those encoded by the *T. brucei* fragment in all six reading frames. **(D)** The similarity of the amino acid sequence encoded by this fragment (tb) to the same region of *L. mexicana* GPI8 (lm) was determined using the GAP programme (GCG Molecular Analysis Package, Wisconsin).

A



B

Nucleic Acid Sequence

```
  1  CCAACAACTG GGCAGTGCTC GTCTCCTCCC GTTACTTCTT TAATCTCCGT
 51  CATACAACGA ATGCGCTGGC CATGTACCAT TTGTGTCGAA AACACGGGAT
101  GGATGACGAC CACATCCTTG TCTCTCTTAG CGATAGCTAC GCCTGCGATC
151  CCCGCAAACC CAACCCAGCG ACAATTTACG GAGCACCTGC CCAAGCGGAG
201  CAACCGAATC TATACGGCTG CAATATTAGA GTGGACTACG CGAGCTACGA
251  TGTCGGTGTT CGCCGGTTCC TTGGCGTCCT GCAGGGCCGA TATGATGAGA
301  ATACTCCTCC TTCACGGCGT CTCGACACCG ACGAGAACTC GAACATAATC
351  ATTTATGTAG CCGGCCATGG CGC
```

Amino Acid Sequence

```
  1  NNWAVLVSSR YFFNLRHTTN ALAMYHLCRK HGMDDDHILV SLSDSYACDP
 51  RKPNPATIYG APAQAEQPNL YGCNIRVDYA SYDVGVRRFL GVLQGRYDEN
101  TPPSRRLDTD ENSNIIIYVA GHGA
```

C

Sequences producing significant alignments:	(bits)	Value
gi|2558891 (AF022913) **GPI transamidase [Homo sapiens]**	125	1e-28
emb|CAA68871| (Y07596) **gpi8 [Homo sapiens]**	125	1e-28
sp|P49048|YJ96_CAEEL HYPOTHETICAL 36.9 KD PROTEIN T05E11.6 IN C	120	5e-27
sp|P49018|GPI8_**YEAST GPI-ANCHOR TRANSMIDASE** >gi|2132503|pir||S5.	114	2e-25
emb|CAA15687.1| (AL009192) /prediction=(method:; /prediction=(m.	108	2e-23
gi|3063464 (AC003981) F22O13.26 [Arabidopsis thaliana]	94	4e-19
sp|P49043|VPE_CITSI VACUOLAR PROCESSING ENZYME PRECURSOR (VPE) .	61	3e-09
gi|4154281 (AF082347) C13 endopeptidase NP1 precursor [Zea mays]	60	6e-09
emb|CAB01126| (Z77653) Similarity to Arabadopsis vacuolar proce.	60	8e-09
sp|P09841|HGLB_SCHMA HEMOGLOBINASE PRECURSOR (ANTIGEN SM32) >gi.	57	5e-08
gi|3413718 (AC004747) alpha-vacuolar processing enzyme [Arabido.	56	9e-08
dbj|BAA76744.1| (D89971) asparaginyl endopeptidase (VmPE-1) [Vi.	56	9e-08
sp|P42665|HGLB_SCHJA HEMOGLOBINASE PRECURSOR (ANTIGEN SJ32) >gi.	56	1e-07
sp|P49044|VPE_VICSA VACUOLAR PROCESSING ENZYME PRECURSOR (VPE) .	55	2e-07
emb|CAA21203| (AL031804) gamma-VPE (vacuolar processing enzyme).	54	3e-07
dbj|BAA18924| (D61395) gamma-VPE [Arabidopsis thaliana]	54	3e-07
emb|CAA04439| (AJ000990) legumain [Mus musculus]	54	3e-07
dbj|BAA09530| (D55696) cysteine protease [Homo sapiens]	54	4e-07
sp|P49047|VPEA_ARATH VACUOLAR PROCESSING ENZYME, ALPHA-ISOZYME .	53	6e-07
pir||S60050 vacuolar processing enzyme (EC 3.4.22.-) isozyme be.	53	8e-07
gi|161019 (M21308) hemoglobinase [Schistosoma mansoni]	52	1e-06
sp|Q99538|LEGU_HUMAN LEGUMAIN PRECURSOR (ASPARAGINYL ENDOPEPTID.	52	1e-06
emb|CAB17078| (Z99956) asparagine-specific endopeptidase precur.	52	2e-06
sp|P49042|VPE_RICCO VACUOLAR PROCESSING ENZYME PRECURSOR (VPE) .	51	3e-06
emb|CAB16318| (Z99174) cysteine proteinase precursor [Vicia nar.	48	2e-05
emb|CAB17079| (Z99957) legumain-like proteinase precursor [Phas.	48	2e-05
sp|P49046|LEGU_CANEN LEGUMAIN PRECURSOR (ASPARAGINYL ENDOPEPTID.	48	3e-05
emb|CAB42650.1| (AJ238880) putative preprolegumain [Nicotiana t.	46	1e-04
dbj|BAA76745.1| (D89972) asparaginyl endopeptidase (VmPE-1A) [V.	46	1e-04
emb|CAA07639| (AJ007743) cysteine proteinase precursor [Vicia s.	45	2e-04
emb|CAB42655.1| (AJ238882) putative preprolegumain [Vicia narbo.	45	2e-04
sp|P49045|VPE_SOYBN VACUOLAR PROCESSING ENZYME PRECURSOR (VPE) .	43	7e-04

D

```
tb  1 RYFFNLRHTTNALAMYHLCRKHGMDDDHILVSLSDSYACDPRKPNPATIY 50
      || || ||| ||| |||| |.||.||||||. ||||:|||||   || |:
lm  1 RYLFNYRHTANALTMYHLLRQHGIDDDHILLFLSDSFACDPRNVYPAEIF 50
                  .         .         .         .         .
tb 51 GAPAQAEQP......NLYGCNIRVDYASYDVGVRRFLGVLQGRYDENTPP 94
        |  |         |||||. .|||| || ||||| ||||||||||||
lm 51 SQPPGAHDADGRASMNLYGCSAQVDYAGSDVDVRRFLSVLQGRYDENTPP 100
                 .         .
tb  95 SRRLDTDENSNIIIYVAGHGA 115
       .||| .|  ||||||||||
lm 101 TRRLLSDNTSNIIIYVAGHG. 120
```

Figure 3.14: Nested PCR with *Plasmodium falciparum* DNA. Nested PCR was carried out as follows: first round of PCR, three reactions were set up:

Reaction A: Primer OL80 only (control)

Reaction B: Primer OL82 only (control)

Reaction C: Primers OL80 and 82

The reactions were run as described in Figure 3.2 legend. Reactions were diluted 10 fold with water and these were used as template in a second round of PCR with primer combinations as follows:

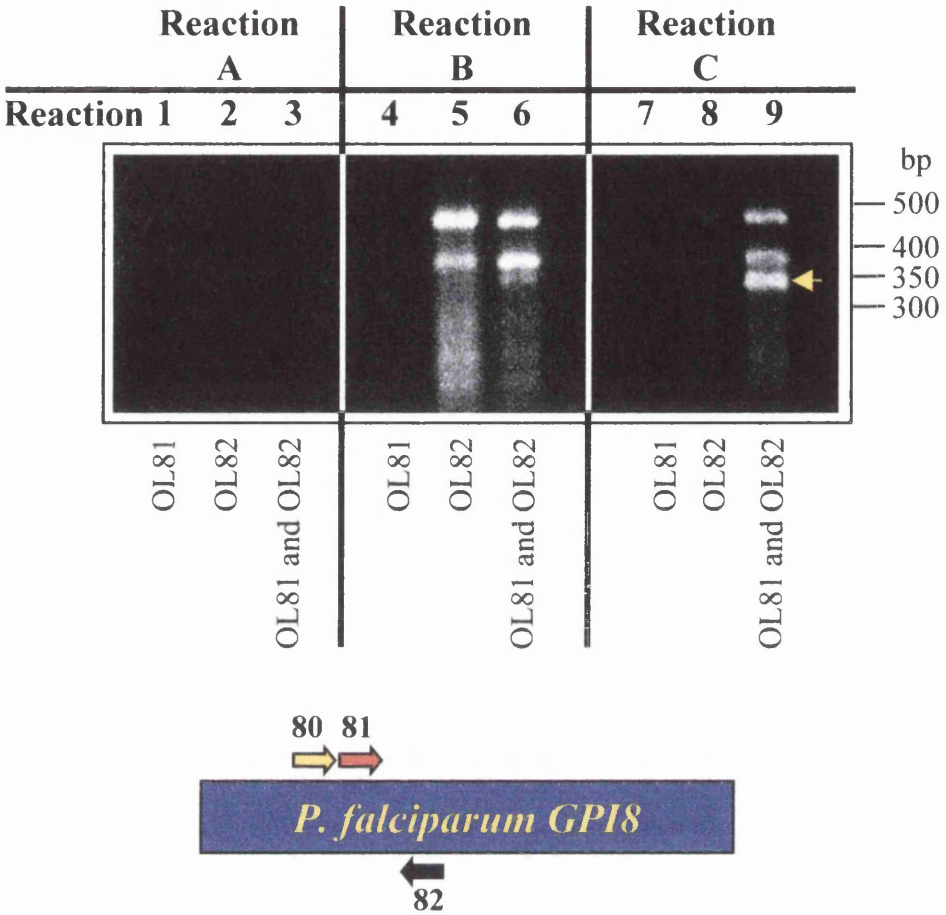
Reaction A template: (1) OL81 only, **(2)** OL82 only **(3)** OL81 and 82 (all controls)

Reaction B template: **(4)** OL81 only, **(5)** OL82 only, **(6)** OL81 and 82 (all controls)

Reaction C template: **(7)** OL81 only, **(8)** OL82 only, **(9)** OL81 and 82 (7&8 are controls)

The reactions were electrophoresed on a 1.2% acrylamide gel and stained with ethidium bromide **(A)**. The primer orientation with respect to *PFGPI8* is also shown. **(B)** A 350bp fragment (marked by yellow arrow in (A)) was gel-purified, cloned into pGEM-T, sequenced and translated to give the amino acid sequence. A BLASTx database search was carried out with the nucleotide sequence **(C)** and a sequence comparison of the *P. falciparum* with *H. sapiens* and *S. cerevisiae*, produced by the GAP programme (GCG Wisconsin) is shown **(D)**.

A



B

Nucleotide Sequence

```
  1  TTTTGGTTTA ATTATAGACA TATGGCCAAC ACTCTTAGTA TGTATCGTAC
 51  TGTTAAGAGA TTAGGTATAC CAGATTCACA GATTATATTA ATGCTAGCGG
101  ATGATGTTGC TTGCAATCCA AGAAACGCAT TTCCTGGTTT TGTATTCAAC
151  AATCAAGATA GACAGTTAGA ACTATACGGT GATAATATTG AGGTTGATTA
201  TCGTGGATAT GAGGTTACCG TTGAAAATTT TATCAGATTA TTAACTGACC
251  GCTGGAGTGA AGACCAACCT AGATCAAAAA GATTATTAAC CGATGAAAAT
301  TCGAATATTT TTATTTATCT TACTGGACAC GGAGG
```

Amino Acid Sequence

```
  1  FWFNYRHMAN TLSMYRTVKR LGIPDSQIIL MLADDVACNP RNAFPGFVFN
 51  NQDRQLELYG DNIEVDYRGY EVTVENFIRL LTDRWSEDQP RSKRLLTDEN
101  SNIFIYLTGH G
```

C

```
sp|P49018|GPI8_YEAST  GPI-ANCHOR TRANSMIDASE >gi|2132503|pir...   210  4e-54
emb|CAA68871|  (Y07596) gpi8 [Homo sapiens]                      171  2e-42
gi|2558891  (AF022913) GPI transamidase [Homo sapiens]           171  2e-42
emb|CAA15687.1|  (AL009192) /prediction=(method:; /predictio...   170  3e-42
gi|3063464  (AC003981) F22O13.26 [Arabidopsis thaliana]          148  1e-35
sp|P49048|YJ96_CAEEL  HYPOTHETICAL 36.9 KD PROTEIN T05E11.6 ...  146  4e-35
emb|CAB55340.1|  (AJ242865) GPI:protein transamidase [Leishm...   107  4e-23
emb|CAB01126|  (Z77653) Similarity to Arabadopsis vacuolar p...    88  2e-17
emb|CAA04439|  (AJ000990) legumain [Mus musculus]                 76  9e-14
ref|NP_005597.1|PPRSC1|  protease, cysteine, 1 (legumain) >g...   75  3e-13
dbj|BAA09530.1|  (D55696) cysteine protease [Homo sapiens]        75  3e-13
emb|CAB51545.1|  (AJ243876) vacuolar processing enzyme [Lyco...   72  1e-12
sp|P09841|HGLB_SCHMA  HEMOGLOBINASE PRECURSOR (ANTIGEN SM32)...   71  2e-12
sp|P42665|HGLB_SCHJA  HEMOGLOBINASE PRECURSOR (ANTIGEN SJ32)...   69  1e-11
gi|161019  (M21308) hemoglobinase [Schistosoma mansoni]           68  3e-11
sp|P49047|VPEA_ARATH  VACUOLAR PROCESSING ENZYME, ALPHA-ISOZ...   65  2e-10
gi|3413718  (AC004747) alpha-vacuolar processing enzyme [Ara...   64  3e-10
gi|4154281  (AF082347) C13 endopeptidase NP1 precursor [Zea ...   64  3e-10
```

D

P. falciparum* vs *H. sapiens

```
  1 .....FWFNYRHMANTLSMYRTVKRLGIPDSQIILMLADDVACNPRNAFP 45
         |||||||.|||||.||.|||||||||| |:|||||.||||||  |
 56 .....FWFNYRHVANTLSVYRSVKRLGIPDSHIVLMLADDMACNPRNPKP 100
               .         .         .         .         .
 46 GFVFNNQDRQLELYGDNIEVDYRGYEVTVENFIRLLTDRWSEDQPRSKRL 95
      ||.... :| .|||.:|||||| ||||||||:|.|| |     ||||||
101 ATVFSHKNMELNVYGDDVEVDYRSYEVTVENFLRVLTGRIPPSTPRSKRL 150
               .         .         .         .         .
 96 LTDENSNIFIYLTGHG 111
    |.|: ||| ||:||||
151 LSDDRSNILIYMTGHG 166
```

P. falciparum* vs *S. cerevisiae

```
  1 ..............................................FWF 3
                                                   |||
 48 ..............................................FWF 50
               .         .         .         .         .
  4 NYRHMANTLSMYRTVKRLGIPDSQIILMLADDVACNPRNAFPGFVFNNQD 53
    ||||||| ||||||||||||||||||||.|||||| || ||| ||||.|
 51 NYRHMANVLSMYRTVKRLGIPDSQIILMLSDDVACNSRNLFPGSVFNNKD 100
               .         .         .         .         .
 54 RQLELYGDNIEVDYRGYEVTVENFIRLLTDRWSEDQPRSKRLLTDENSNI 103
      ::||||.:|||||||||||||||||||||.|| |:||||||||||||
101 HAIDLYGDSVEVDYRGYEVTVENFIRLLTDRWTEDHPKSKRLLTDENSNI 150
               .         .         .         .         .
104 FIYLTGHG 111
    |||:||||
151 FIYMTGHG 158
```

CHAPTER 4

TARGETED GENE REPLACEMENT OF *LMGPI8* AND MUTANT PHENOTYPE ANALYSIS

4.1 INTRODUCTION

One method of investigating the function of a gene product is to remove the gene by targeted gene replacement and determine what effect or effects this may have on the organism. Targeted gene replacement was first described in *L. major* by Beverley and co-workers where the *NEO* gene was inserted into the *DHFR-TS* locus (Cruz and Beverley, 1990) and was then used to remove both allelic copies of the gene giving rise to *DHFR-TS* null mutants that require the addition of thymidine for growth (Cruz *et al.*, 1991). It was also described in *L. enrietti* by Wirth and co-workers (Tobin and Wirth, 1992) after demonstrating homologous recombination between two plasmids containing independent mutations within the chloramphenicol acetyltransferase (*CAT*) gene, each causing a lack of activity (Tobin *et al.*, 1991). The detection of high levels of CAT activity demonstrated that recombination had occurred between plasmids yielding an active *CAT* gene and Southern analysis and PCR confirmed this. Subsequent experiments resulted in stable integration of the NEO^r gene (previously demonstrated to be expressed reliably in *L. enrietti* (Laban *et al.*, 1990)) flanked by α-tubulin flanking regions at the α-tubulin locus of *L. enrietti*, with the result that the cells were resistant to the antibiotic Geneticin (Tobin and Wirth, 1992). Further experiments led to the deletion of an entire array of α-tubulin genes (Curotto de Lafaille and Wirth, 1992).

The gene-targeting process involves the exploitation of naturally occurring, homologous recombination of chromosomal DNA by introducing into the parasite an engineered knockout cassette consisting of a selectable marker, such as an antibiotic-resistance gene, flanked by gene-specific regions of DNA. These regions also flank or include part of the open reading frame of the gene of interest and therefore target the knockout cassette to the correct locus within the chromosome (recombination of DNA is more likely to occur between homologous stretches) and replacing or disrupting the target gene. Knockout cassettes are introduced into the parasite by electroporation, a technique used to allow plasmids to be taken up by *Leishmania* (LeBowitz, 1994). Integration of such knockout constructs causes cells to express the gene and thus become resistant to the antibiotic. The procedure has been used extensively to disrupt or remove both individual genes and multi-copy gene arrays. Individual genes, such as those encoding two cysteine proteinases of *L. mexicana* (CPA and CPC), were disrupted in this way (Souza *et al.*, 1994; Bart *et al.*, 1997). Further, the entire gene array encoding cysteine proteinase b (CPB) in *L. mexicana* (comprising 19 copies of the gene tandemly arranged over 50kb of a chromosome) was disrupted using this method (Mottram *et al.*, 1996a). Studies in *Leishmania* species determined that the optimum length of flanking regions is approximately 1kb for high efficiency homologous recombination (Papadopoulou and Dumas, 1997). *Leishmania* is a diploid organism, therefore two successive rounds of transfection with two independent selectable markers are required to delete each allelic copy of a gene.

Genes that are essential for the viability of *Leishmania* species can be identified using this procedure, since the only cells that survive transfections are those that retain a

copy of an essential gene. Attempts to knock out an essential gene such as the cyclin-dependent kinase *CRK*1 of *L. mexicana* (Mottram *et al.*, 1996b) or the dihydrofolate reductase-thymidylate synthase (*DHFR-TS*) gene without providing the required thymidine (Cruz *et al.*, 1993) can result in clones that have altered their ploidy in order to retain at least a single copy of the gene. Alternatively, rearrangement of the genome can take place such that the gene "jumps" to a different location within the same or a different chromosome, as was seen with *L. donovani* (Dumas *et al.*, 1997).

There are several selectable marker genes that confer resistance to a number of antibiotics that can be used with *Leishmania*. These are shown in Table 4.1.

Table 4.1:
Antibiotic-resistance genes used as selectable markers in *Leishmania* integration experiments.

Gene	Encoded protein	Antibiotic	Reference
NEO	Neomycin phosphotransferase	Geneticin or G418	(Laban *et al.*, 1990)
SAT	Streptothricin acetyltransferase	Nourseothricin	(Joshi *et al.*, 1995)
PAC	Puromycin N-acetyltransferase	Puromycin	(Freedman and Beverley, 1993)
HYG	HygromycinB phosphotransferase	Hygromycin B	(Lee and Van der Ploeg, 1991)
BLE	Bleomycin hydrolase	Phleomycin (Zeocin)	(Souza *et al.*, 1994)

There are alternative selectable markers, such as the Green Fluorescent Protein (GFP), from the jellyfish *Aequorea victoria*, which can be detected by fluorescence. Cells that express GFP can be detected and isolated from those that do not by using fluorescence activated cell sorting (FACS) (Ha *et al.*, 1996). This provides an alternative to the production of antibiotic-resistant cell lines. Such mutants could be used potentially as attenuated live vaccines since they avoid the problem of possible

transmission of antibiotic-resistance genes to other pathogenic organisms (Gueiros-Filho and Beverley, 1994).

It was the aim of this section of work to delete *GPI8* from *L. mexicana* and analyse the phenotype of resulting mutants. This would determine if the gene is essential and also confirm that LMGPI8 is a homologue of yeast Gpi8p and a component of the GPI:protein transamidase.

Our approach was to use the technique that has already been used successfully in this laboratory to disrupt several cysteine proteinase genes (Souza *et al.*, 1994; Mottram *et al.*, 1997; Bart *et al.*, 1997). The preparation and transfection of targeted gene replacement constructs specific for the *GPI8* locus and the subsequent deletion of *GPI8* is described. A combination of PCR, Southern blotting and flow cytometry was used to confirm the removal of *GPI8* and thus the generation of *GPI8* null mutants. Phenotype analysis was performed using immunofluorescence and Western blotting to determine if GPI-anchoring was functional in the mutants. The most abundant surface protein of *L. mexicana* promastigotes is GP63 and is attached via a GPI anchor (Medina-Acosta *et al.*, 1989); GP63 antibodies (Button *et al.*, 1991) were available to examine the presence of GP63 in the mutants. The ability of null mutants to infect macrophages *in vitro* was assessed and compared with wild type infections and *in vivo* mice infections were used to determine the mutant's virulence compared with wild type parasites (Bates *et al.*, 1992).

4.2 METHODS

4.2.1 Preparation of all constructs described in this chapter

4.2.1.1 Preparation of *GPI8* knockout constructs

Constructs prepared originally by Dr. Darren Brooks (WCMP, Glasgow) for the removal of the multi-copy gene array encoding the major lysosomal cysteine proteinases (CPB) were used as the backbone for the *GPI8*-targeting knockout constructs. These constructs carry the *PAC* (pGL207) and *SAT* (pGL51) genes flanked by *CPB* 5` and 3` flanks as well as dihydrofolate reductase-thymidylate synthase (*DHFR-TS*) flanking regions. These *DHFR-TS* flanks were included in the original vector to provide the signals required for trans-splicing and polyadenylation of the antibiotic-resistance gene transcript (LeBowitz *et al.*, 1994).

The aim, however, was to remove both the 5` and 3` *CPB* and *DHFR-TS* flanking regions and replace them with the 5` and 3` flanking regions of *L. mexicana GPI8*. This would allow the utilisation of the naturally occurring trans-splicing and polyadenylation sites of *GPI8* in the expression of the antibiotic-resistance genes in *L. mexicana* promastigotes. As GPI-anchored proteins are found on the surface of the promastigote form, then one would predict that GPI8 must be expressed in this life-cycle stage. Thus the trans-splicing and polyadenylation sites of *GPI8* would be used to process the antibiotic-resistance mRNA following integration in the *GPI8* locus.

Primers were designed to allow the amplification of the 5` and 3` flanking regions of *GPI8* (Figure 4.1) by PCR and these are given in Table 4.2. The 5` flank had *Hin*dIII and *Spe*I sites engineered onto the 5` and 3` ends respectively whilst the 3` flank had *Bam*HI and *Bgl*II sites.

Table 4.2: Primers used to amplify the 5` and 3` flanking regions of *GPI8*.

Primer	Sequence	Description
OL83	GCCAAGCTTTGCTCAGATGACCGAGCCGGCGC	5` flank 5` primer - *Hin*dIII site
OL84	GACTAGTAAACAGCCGGAACTGCACTAGCT	5` flank 3` primer – *Spe*I site
OL85	CGCGGATCCCGTGTGGCATCTACCTCCCTGCG	3` flank 5` primer - *Bam*HI site
OL132	AGATCTGCAACGCACCGACTGGGGTG	3` flank 3` primer - *Bgl*II site

Primers outwith the amplified regions (see Figure 4.4A) were available making it possible to screen for integration at the correct locus by PCR (see Section 4.3.1.1).

PCR amplification of both flanks was carried out as follows:

94°C, 4 minutes	x1
94°C, 1 minute	
55°C, 1 minute	x20
72°C, 1 minute	
72°C, 7 minutes	x1

The proof-reading DNA polymerase VENT (NEB) was used, which, together with a relatively low number of cycles (15), maximised the fidelity of the amplified DNA fragments. Adenosine-overhangs were added to products by incubation for a further 10 minutes at 72°C with Taq polymerase to allow cloning into pGEM-T vector (Promega). Positive plasmid clones were digested with the appropriate restriction enzymes to drop out the 5` and 3` flanks, which were gel-purified. The original pX53 vectors (pGL207 and pGL51) supplied by Dr. Darren Brooks (WCMP, Glasgow) were initially digested with *Hin*dIII and *Spe*I to drop out the 5` flank and the plasmid

backbone was gel-purified. *GPI8* 5` flank was ligated with this vector and positive clones were identified by colony PCR using a 5` flank sense primer (OL474) and a *PAC* (OL15) or *SAT* (OL17) antisense primer. These clones were grown up and plasmid was purified by the Qiagen (Tip20) Mini-Prep kit. The 3` *CPB* and *DHFR-TS* flank was removed by digestion with *Bam*HI and *Bgl*II and replaced with the similarly digested *GPI8* 3` flank as before, resulting in the constructs pGL236 (*PAC*) and pGL237 (*SAT*), which are shown schematically in Figure 4.2A.

Constructs which were identified as positive by colony PCR screening (not shown) were purified and subjected to a number of test restriction digests to confirm that the constructs are complete and correct (Figure 4.2B). The predicted restriction fragment pattern was obtained for both clones (described in Figure 4.2 legend) indicating that the knockout constructs were correct.

4.2.1.2 Preparation of episomal copy of *LMGPI8*

An episomal copy of *GPI8* was prepared in pBluescript (pBS) using the constructs containing the lambda fragments described in the previous chapter as shown in Figure 4.3. *GPI8* and its 5` and 3` flanking regions were removed from pBS by digestion with *Hin*cII and subcloned into the pXG vector (designated pGL269) which was used to transform XL1-blue supercompetent *E. coli* (Stratagene).

4.2.2 Targeted gene replacement of *LMGPI8*

4.2.2.1 Preparation of knockout cassettes for replacement of first allele

Medium-scale production of pGL236 and pGL237 was performed using the Qiagen Tip100 Midi-Prep kit with 25ml of *E. coli* overnight cultures. 60μg of each plasmid

was digested overnight with 4µl each of *Hin*dIII (80U) and *Bgl*II (40U) in a 200µl reaction volume. Digests were electrophoresed on 0.7% agarose TBE gels. The plasmid backbone and the excised knockout cassette were indistinguishable on the gel. The 300bp difference in size is sufficient to separate fragments but the amount of DNA was so large in this instance that the two fragments co-migrated on a medium-sized gel. Thus both fragments were co-purified using Spin X columns. The eluted DNA was ethanol precipitated and dried before being resuspended in 10µl of sterile water. Yields of DNA were around 10 - 15µg of which about half would have been plasmid backbone.

4.2.2.2 Transfection of *L. mexicana* with knockout cassette and selection for transfectants

Mid-log phase *L. mexicana* promastigote cells were washed in electroporation buffer (EPB) before being resuspended in the same buffer at a density of $1x10^7$ cells ml^{-1}. $4x10^7$ cells were mixed with DNA and placed in 0.2cm electrocuvettes as follows:

1) WILD TYPE cells + No DNA (control)

2) WILD TYPE cells + *Hin*dIII/*Bgl*II fragment from pGL236 (*PAC*)

3) WILD TYPE cells + *Hin*dIII/*Bgl*II fragment from pGL237 (*SAT*)

Following electroporation, cells were allowed to recover on ice for 10 minutes before being transferred to 10ml of HOMEM medium containing 10% (v/v) HI-FCS. Cultures 2) and 3) were immediately divided into two to provide two independent cultures from each transfection. The cells would not have had sufficient time to undergo a cell division hence transfectants produced from individual cultures would

have arisen from independent integration events. All cultures were grown overnight at 25°C and were then exposed to antibiotics as follows:

(1) Control Transfection.

Solid Medium: 5ml of culture was pelleted, resuspended in 400µl of supernatant and spread equally on 1 agar plate containing 25µg/ml nourseothricin, and 1 agar plate containing 10µg/ml puromycin.

Liquid medium: 2.5ml of culture was added to 7.5ml of medium (HOMEM) containing of 10µg/ml puromycin (final concentration) and 2.5ml into medium containing of 25µg/ml nourseothricin (final concentration).

(2) Transfection with knockout cassette containing *PAC* gene

Solid medium: 2.5ml of each overnight culture was pelleted, resuspended in 200µl of supernatant and spread equally on two agar plates, each containing 10µg/ml puromycin.

Liquid medium: 2.5ml of each culture were added to 7.5ml of medium containing 10µg/ml puromycin (final concentration).

(3) Transfection with knockout cassette containing *SAT* gene.

Solid medium: 2.5ml of each overnight culture was pelleted, resuspended in 200µl of supernatant and spread equally on two agar plates, each containing 25µg/ml nourseothricin.

Liquid medium: 2.5ml of each culture were added to 7.5ml of HOMEM with 25µg/ml nourseothricin (final concentration).

Plates were allowed to absorb the liquid medium for 5 minutes. They were then sealed with Parafilm, inverted and incubated at 25°C for 2-3 weeks.

4.2.3 *In vitro* macrophage infections

Freshly culled BALB/c mice were injected intraperitoneally with 5 ml of ice-cold RPMI medium (GibcoBRL) complete with 10U/ml penicillin and 10μg/ml streptomycin sulphate (GibcoBRL). After 5 min gentle massaging of the mouse abdomen, the RPMI containing macrophages was harvested and kept on ice. Macrophages were counted and diluted to a cell density of 5×10^5 cells/ml with ice-cold RPMI. 2×10^5 macrophages (400μl) were placed in each chamber of an eight-chamber Permanox slide (Labtek) and allowed to adhere overnight at 32°C under 5% CO_2/95% air. Chambers were washed with RPMI to remove non-adhered cells before 2×10^5 stationary phase promastigotes were added in RPMI with 10% foetal calf serum (FCS) and incubated at 32°C under 5% CO_2/95% air. After four hours, slides were washed to remove free promastigotes and fresh RPMI with 10% FCS was added. Infection was allowed to proceed for seven days at 32°C under 5% CO_2/95% air. Individual chambers were fixed with 100% methanol for 5 min at 8 h, 1 day, 3 days and 7 days after addition of promastigotes. Slides were stained with Modified Giemsa Stain (Sigma Diagnostics), diluted 20-fold with distilled water, for 30 min and viewed by light microscopy using a Zeiss Axioplan microscope

4.2.4 Mutation of episomal copy of *LMGPI8*

His^{63}, Cys^{94}, His^{174} and Cys^{216} of LMGPI8 were selected for mutation by site directed mutagenesis using the QuikChange Site Directed Mutagenesis Kit (Stratagene). The

primers shown in Table 4.3 were designed to bring about the mutations of the episomal copy of *GPI8* in pGL269 using PCR according to the manufacturer's protocol such that cysteine was mutated to glycine and histidine was mutated to alanine. Note that for each primary primer used, a complementary primer (including the mutation) was used as the second primer in the PCR and that a 20 minute elongation time for each cycle of PCR was used due to the size of pGL269 (10kb).

Table 4.3:
Primers used to mutate potential active site residues (mutated nucleotides are shown in bold)

Primer	Sequence (5'-3')	Mutation
OL330	CTCTTCAACTACCGC**GC**CACCGCCAATGCGC	**His^{63}-Ala**
OL332	GACAGCTTCGCC**GG**CGACCCGCGAAATG	**Cys^{94}-Gly**
OL334	CTACGTCGCGGGG**GC**CGGCGCCAAGTC	**His^{174}-Ala**
OL336	CCTGGCAGATACA**GG**CCATGCGATTGCG	**Cys^{216}-Gly**

Following PCR, plasmid was treated according the QuikChange Kit protocol and used to transform *E. coli*. Plasmids were purified from several colonies using Tip20 mini-preps (Qiagen) and sequenced to check if the mutation had been incorporated.

4.3 RESULTS

4.3.1 Targeted gene replacement of *GPI8*

No growth was observed on the control plates or in the control liquid cultures indicating that antibiotic selection was effective. Transfection (2) yielded three puromycin-resistant clones, two from one plate and one from the second (designated P1, P2 and P3). Transfection (3) yielded just one nourseothricin-resistant clone on solid medium (S1). Growth was observed in all liquid cultures of these transfections

but since clones were obtained on solid medium, these cultures, which were not clonal, were not examined further. Clones were transferred into liquid media with appropriate antibiotics and were grown up for further analysis.

4.3.1.1 Check for integration using PCR on whole cells

Primers designed to anneal upstream and downstream of the 5` and 3` flanking regions of the knockout cassettes were synthesised and used to check for integration by PCR (Figure 4.4).

OL133	GGGTGGTGGGCGGAGATTGAG	**5` (Sense) Primer**
OL86	GAAGATCTTTGCTCGTGATACGACGGCGTGG	**3`(Antisense) Primer**

The *SAT* and *PAC* genes are about 550bp and 450bp, respectively, smaller than *GPI8* and these differences can be detected by electrophoresis. PCR of wild type *L. mexicana* produced a single fragment of 3.1kb, corresponding to the wild type *GPI8* alleles (Figure 4.4, lane 1). Clones P1, P2, P3 and S1 all yielded a 3.1kb fragment corresponding to wild type *GPI8* as well as one fragment of 2.6-2.7kb (lanes 2-5, respectively) which correspond to integration of *PAC* or *SAT* cassettes.

4.3.1.2 Second round transfections of P1, P2, P3 and S1 clones

The three puromycin-resistant clones, P1, P2 and P3, and the nourseothricin-resistant clone S1 were grown up for transfection as before. The puromycin-resistant clones were transfected with the *SAT* knockout cassette, which was prepared by restriction digest and gel-purification as described in the previous section. One alteration to the procedure was that *Pvu*I was included in the digestion. This enzyme bisects the

plasmid backbone allowing most of the plasmid to be separated from the knockout cassette by electrophoresis, thus minimising the possibility of plasmid re-ligation which would inhibit chromosomal integration of the cassette. S1 was transfected with the *PAC* knockout cassette prepared in the same way. P2 and P3 produced several clones on HOMEM plates containing both nourseothricin and puromycin (25μg/ml and 10μg/ml, respectively). On further analysis by PCR these clones, which were nourseothricin- and puromycin-resistant, were found to retain a copy of *GPI8* (Figure 4.5A). S1 produced no double resistant clones at all. P1, however, yielded three clones, which were analysed by PCR and appeared to have lost the *GPI8* gene (Figure 4.5B). The 3.1kb fragment that corresponds to the *GPI8* allele is absent in all of these clones; only the 2.6-2.7kb fragment corresponding to the *PAC* and *SAT* cassettes are detectable (Figure 4.5B – lanes 2, 3 and 4). The clones were named P1S1, P1S2 and P1S3.

4.3.1.3 Southern analysis of potential *GPI8* null mutants

To confirm the PCR results, Southern blot analysis was performed. Genomic DNA was extracted from two of the three potential *GPI8* null mutants (P1S1 and P1S3; P1S2 was not analysed further in this study) together with wild type and P1, the heterozygote parent cell line of P1S1 and P1S3. Use was made of an *Xho*I site immediately 5` to the stop codon in the *GPI8* ORF (see Figure 4.6A). This site is not present in either the *SAT* or *PAC* genes and would be predicted to be lost if correct integration of the knockout cassettes had occurred. Thus 5μg of DNA from each cell line was digested with *Xho*I and electrophoresed on 0.7% agarose TBE gel. The gel was stained with ethidium bromide prior to blotting, confirming the complete digestion of DNA (Figure 4.6B).

The blot was probed with the *XbaI/XhoI* fragment containing the complete ORF of the gene (Figure 4.6C). A strongly hybridising band of 2.4kb was detected with wild type DNA (lane 1). A weaker signal was detected with the *GPI8* heterozygote, P1 (lane 2), as expected given that the amount of *GPI8* DNA would be expected to be half that of wild type. No hybridisation could be detected in either of the putative *GPI8* null mutants (lanes 3 and 4) suggesting that *GPI8* had been lost. To confirm this finding, several other probes were used. The blot was stripped by the boiling SDS method and re-probed with a $[\alpha]^{32}$P-CTP labelled *SAT* gene (Figure 4.6D). 3.6kb fragments could be detected in both the putative null mutant lanes (Figure 4.6D, lanes 3 and 4) but not in wild type or P1 lanes (lanes 1 and 2 respectively). This is what we would expect, given that the heterozygote is puromycin-resistant (not nourseothricin-resistant). The restriction digests were repeated, electrophoresed and blotted. Probing with the *PAC* gene produced bands of 3.7kb with P1, P1S1 and P1S3 (Figure 4.6E, lanes 2, 3 and 4) as expected. One final hybridisation was carried out to make certain that integration had occurred at the correct locus. The previous blot was stripped as before and probed with the 5' flanking region. The probe, which was produced by PCR with primers OL85 and OL132 (Figure 4.1), detected a single 2.4kb fragment in wild type as expected from the map (Figure 4.6F). We would expect fragments to hybridise in all four lanes but to differ from the wild type DNA in size, due to the loss of the *XhoI* site in the *GPI8* ORF. A 2.4kb fragment was labelled by this probe in both wild type and P1 heterozygotes (4.6F, lanes 1 and 2) corresponding to the 5' flank and *GPI8* ORF contained on a single *XhoI* fragment (as shown in Figure 3.11A, Chapter 3). A second band of approximately 3.7kb was also

detected with P1 DNA corresponding to the replaced allele (containing the *PAC* gene). Only the 3.7kb fragment was detected in the putative null mutants, P1S1 and P1S3 (Figure 4.6F – lanes 3 and 4).

These data indicate that integration has taken place at the correct locus and that *GPI8* has been deleted. P1S1 and P1S3 are *GPI8* null mutants (Δ*GPI8*).

4.3.1.4 Transfection of episomal copy of *GPI8* into null mutants

P1S1 and P1S3 cells were transfected with pGL269. After the overnight recovery stage, transfected cultures were spread on agar plates containing 10μg/ml puromycin and 25μg/ml of both nourseothricin and Geneticin (G418 – GibcoBRL). One triple-resistant clone was analysed from each parent cell line by Southern Blotting. DNA was isolated from each cell line (P1S1[pGL269] and P1S3[pGL269]), digested with *Xho*I, electrophoresed and blotted. Blots were probed with [α]^{32}P-CTP-labelled *GPI8* ORF. Fragments of approximately 5kb were labelled in each lane (Figure 4.7) and these correspond to the expected size of the episomal fragments, given that pGL269 contains two *Xho*I sites (Figure 4.6). The episome is also multi-copy in *L. mexicana*, reflected in the high intensity of the hybridising signal.

Thus P1S1[pGL269] and P1S3[pGL269] are *GPI8* null mutants re-expressing the gene from an episome.

4.3.1.5 Confirmation of ploidy by flow cytometry

Genetic manipulation of *Leishmania* by transfection can result in changes of ploidy, particularly if the gene concerned is essential (Cruz *et al.*, 1993; Mottram *et al.*, 1996b). To assess the DNA content of the mutants, flow cytometry was used. This was done for wild type, P1 (heterozygote), P1S1, P1S3, P1S1[pGL269] and P1S3[pGL269] cell lines (Figure 4.8). FACS produces a trace with two peaks, one corresponding to cells in the first gap phase of the cell cycle (G1) after mitosis, and the other to the second gap phase following DNA synthesis but before mitosis (G2). By comparison with the trace for diploid wild type cells, all of the mutant cell lines appeared diploid (see Figure 4.8 legend for details). This result confirms that no changes in ploidy had taken place during deletion of *GPI8* and that the gene is not essential to the viability of *L. mexicana* under these *in vitro* conditions.

4.3.2 Phenotype analysis of *GPI8* null mutants and re-expresser cell lines

At this point of the work, a *L. mexicana* gene with homology to both yeast and human *GPI8* had been identified, characterised and knocked out. However, it was still not certain whether this gene encoded a component of the *L. mexicana* GPI:protein transamidase. Analysis of the phenotype of the null mutants and re-expresser cell lines compared with wild type cells would provide further evidence on whether *LMGPI8* is indeed the homologue of yeast and human *GPI8* and a key component of the transamidase complex.

4.3.2.1 Growth of *GPI8* null mutants and re-expressers compared with wild type cells

To determine the growth rate of null mutants and re-expresser promastigote cell lines, three 12ml cultures were seeded with $1x10^5$ cells ml^{-1} for each cell line (wild type,

P1S1 and P1S1[pGL269]) and their growth was monitored by recording their spectrophotometric absorbance at 600nm on a daily basis (Figure 4.9). An unseeded control culture was also prepared. Prior to measurements, cells were fixed by the addition of 50µl of 40% formaldehyde to 950µl of culture. The *in vitro* growth of *GPI8* null mutant promastigotes was found to be comparable with both wild type and the re-expresser cell lines. A slight lag in P1S1 growth was found after day 1, however the cells recovered and reached stationary phase at around the same time as the other cell lines. The reason for this is unclear, although it is possible that the cultures were not at the same stage of the growth cycle when used to inoculate the fresh medium for this experiment. Further experiments are needed to clarify this.

4.3.2.2 *In vitro* invasion of mouse peritoneal exudate cells (PEC) by *GPI8* null mutants

Mouse macrophages, harvested from peritoneal flushing of BALB/c mice, were exposed to wild type, P1S1 and P1S1[pGL269] promastigotes at a 1:1 ratio in 8 well Permanox chamber slides (Labtek). Time points were taken at 8 hours, 1 day, 3 days and 7 days after which the cells were fixed by immersion in methanol. Slides were stained with Giemsa's stain and viewed by light microscopy at 1000x magnification using an oil immersion lens. Approximately 100 macrophages were examined for infection at each time point and the following measurements were recorded:

- Number of macrophages counted
- Number of macrophages infected
- Number of amastigotes in infected macrophages

The ratio of infected to uninfected macrophages was plotted against the time over which data were recorded (Figure 4.10A). The level of infection by the *GPI8* null mutants was similar to that of wild type and re-expresser cell lines. P1S3 and P1S3[pGL269] infections, unlike P1S1 and P1S1[pGL269], were slightly reduced with respect to wild type, however the cell lines were still capable of infecting and proliferating within macrophages. The average number of P1S1 and P1S3 amastigotes per infected macrophages increased over the 7 days to between 3 and 4 amastigotes per infected macrophage (Figure 4.10B). A similar increase was found with wild type and both *GPI8* re-expresser cells.

These P1S1 cells have a similar growth phenotype to wild type cells. In addition, episomal re-expression of *GPI8* has no obvious effect on cell growth. These data show that GPI-anchored proteins are not required for entry of promastigotes into macrophages, their subsequent differentiation into amastigotes or their survival under these *ex vivo* conditions.

4.3.2.3 *In vivo* infection of BALB/c mice by *GPI8* null mutants

To assess the ability of *GPI8* null mutants to survive under *in vivo* conditions, wild type, P1 (heterozygote), P1S1, P1S1[pGL269], P1S3 and P1S3[pGL269] were each inoculated into six BALB/c mice and lesion volume in each mouse was monitored over a period of up to 11 months. Stationary phase promastigotes were prepared for inoculation by Mr. David Laughland; cells were resuspended in PBS at $2.5x10^7$ cells/ml and 0.2ml ($5x10^6$ cells) was injected subcutaneously into the shaven rump of

each mouse. Inoculations and monitoring of lesion data was carried out by staff at the University of Glasgow Animal Facility.

The data are shown as plots of lesion volume over the 47 week (11 month) period (Figure 4.11).

Mice inoculated with wild type cells had begun to produce lesions within 7-9 weeks of inoculation and the volume of the lesions increased throughout the 9 month period except for mouse 4 (Figure 4.11A). Mice were culled at the 34th week due to the size of the lesions.

The *GPI8* heterozygote P1 began to form lesions approximately 11 weeks after inoculation and interestingly, while lesion volume was seen to increase over time, it was on average, about 10 fold smaller than wild type by the end of the 6th month (Figure 4.11B). By 47 weeks, however, lesions ranged from $30mm^3$ to $680mm^3$. The largest (mouse 6) was three-fold smaller than the largest wild type. However, it should be noted that these measurements were made at week 47 when mice infected with wild type parasites had been culled.

Lesions for P1S1 and P1S3 cell lines were first noticed at around the 20th week post-infection, however, these remained very small for the first six months (Figure 4.11C and D; note the difference in the scale of y-axes compared with 4.11A and B). Thereafter, lesions were observed to increase in volume. P1S1 lesions ranged from none to $240mm^3$ and P1S3 were between 50 and $230mm^3$ (all P1S3 mice had

produced lesions, whereas two P1S1 mice had not). Thus the largest was 8-fold smaller than the largest wild type lesion (at the 34 week stage).

GPI8 re-expresser cell lines were found to have a similar or even poorer growth profile than the null mutant, indicating that complementation *in vivo* was not achieved by episomal expression of *GPI8* (Figure 4.11E and F). By the 47th week following inoculation, P1S1[pGL269] mutants failed to produce lesions in three of the six mice and only one had a lesion of appreciable volume (67mm^3). P1S3[pGL269] mutants produced similar results. One mouse developed a lesion of 80mm^3 (culled after 36 weeks) whilst another developed a lesion of 195mm^3, that is, around 10- to 20-fold less than wild type.

It should be noted that a number of mice infected with the different cell lines were observed to exhibit a sudden increase in lesion volume between weeks 43 and 44. It seems likely that this increase was an artefact, perhaps due to human error.

4.3.2.4 Effects on surface expression of the GPI-linked protein GP63

The major surface metalloproteinase, GP63, is the most abundant GPI-linked protein in the insect (promastigote) stage of the life cycle of the *Leishmania* species (Frommel *et al.*, 1990; Medina-Acosta *et al.*, 1989). If *GPI8* encodes a key component of the transamidase complex, then the surface expression of GP63 should be affected when *GPI8* is deleted. Therefore, the presence of GP63 was assessed by immunofluorescence and Western blotting using monoclonal GP63 antibodies.

4.3.2.4.1 Immunofluorescence with anti-GP63 antibodies

Wild type, P1S1 and P1S1[pGL269] cells were fixed on slides with 2.5% paraformaldehyde. The primary anti-GP63 antibody used was a monoclonal antibody named L3.6 that had been raised against *L. mexicana* GP63 and was a gift of Dr. David Russell, Washington University, St. Louis, USA (see Medina-Acosta *et al.*, (1989)). Cells were visualised by fluorescence microscopy (Figure 4.12). Wild type cells fluoresced strongly whilst the control cells exhibited very little fluorescence, indicating that the non-specific interactions of the secondary antibody were minimal and that the signals viewed were due to the binding of primary antibody. By contrast, P1S1 cells exhibited little or no fluorescence, indicating that GP63 was not expressed in detectable levels on the surface of these cells. Re-expression of the gene from the pXG episome restored the surface expression of GP63.

4.3.2.4.2 Western blotting with anti-GP63 antibodies

If GP63 is not being expressed on the cell surface, what is the fate of the protein in these null mutants? To attempt to answer this question, extracts of wild type, P1S1 and P1S1[pGL269] were prepared by the lysis of 10^8 cells for SDS-PAGE and Western blotting. Following electrophoresis on a 10% acrylamide gel, the gel was electroblotted onto Hybond C (Amersham) and the membrane blocked overnight in Blocking Buffer 2. The primary antibody used was a mouse monoclonal raised against *L. major* GP63, which cross-reacts with *L. mexicana* GP63, a gift of Dr. Robert McMaster, University of British Columbia, Vancouver, Canada (see Button *et al.*, (1991)). This antibody worked well with Western Blotting but was not suitable for immunofluorescence since it is unable to detect native GP63 on fixed cells.

Blots were washed and exposed to secondary antibody (anti-mouse IgG-horseradish peroxidase conjugated - Promega) at a 1 in 5000 dilution for 50 minutes. After washing, antibody binding was detected using Enhanced Chemiluminescence (ECL) reagents (Pierce). Blots were exposed to X-ray film from 15 seconds to 2 minutes and these are shown in Figure 4.13. A protein of approximately 60kDa (GP63) was detected in wild type and P1 cells (Figure 4.13A – lanes 1 and 2, respectively) but was absent from *GPI8* null mutant cells (lane 3). Episomal re-expression of *GPI8* restored GP63 to these cells (lane 4). When the blot was overexposed, a small amount of GP63 was detected in the null mutant in comparison with wild type (Figure 4.13B). Note that a well was left empty between samples on the gel, to minimise the risk of sample over-spill into adjacent wells. As a control, the same lysates were electrophoresed, blotted and probed with an anti-CRK1 antibody. CRK1 is found at relatively constant levels throughout the life cycle of *L. mexicana* (Mottram *et al.*, 1996b) and therefore makes a good control for protein loadings. A similar signal was detected in all samples (Figure 4.13C) indicating that the loss of GP63 is not due to underloading of the mutant cell lysates.

4.3.3 The fate of GP63 in *GPI8* null mutants

Several fates of GP63 in *GPI8* null mutants could be envisaged. It could be secreted into the growth medium, which has been observed during an artificially induced GPI-negative phenotype (Mensa-Wilmot *et al.*, 1994), or it could be retained and degraded, as is seen with incorrectly transamidase-processed proteins in mammalian systems (Wainwright and Field, 1997), and is mediated by the cytoplasmic proteasome (Wilbourn *et al.*, 1998). It is known that GPI anchor attachment plays a

determining role in the forward transport of an artificially expressed protein in *T. brucei* (Bangs *et al.*, 1997). To begin to address this, attempts were made to detect secreted GP63 in *GPI8* null mutant growth medium. Inhibitors of the proteasome and the major lysosomal proteinases were added to growing cultures and the effects on GP63 in *GPI8* null mutant cells were determined.

4.3.3.1 Do *GPI8* null mutants secrete GP63?

To begin to address this question, 10^8 late-log phase promastigotes (wild type, P1, P1S1 and P1S1[pGL269]) were added to 5ml of serum-free HOMEM and incubated for 7 hours. Cells were then pelleted and media was filter-sterilised and concentrated 50-80 fold using Vivaspin 10kDa molecular weight cut-off columns (Vivascience Ltd. Lincoln, UK). Media samples were analysed by Western blotting with anti-GP63 antibodies but no GP63 could be detected in any of the samples (not shown). Thus GP63 does not appear to be secreted from *GPI8* null mutants, although this method may not be sensitive enough to detect secreted proteins.

4.3.3.2 Inhibition of degradative enzymes and the effect on GP63 loss in *GPI8* null mutants

The proteasome has been implicated as being involved in the quality control of GPI-anchoring in mammalian systems (Wilbourn *et al.*, 1998). To determine if the *L. mexicana* proteasome is involved in GP63 degradation, wild type, P1, P1S1 and P1S1[pGL269] promastigotes were grown for 24 hours in medium containing the potent proteasome inhibitor MG132 (10μM). Lysates were prepared and analysed by Western blotting with anti-GP63 antibodies (Figure 4.14). GP63 was readily detected in both wild type and P1S1[pGL269] (Figure 4.14A - lanes 1 and 3), however the

loss of GP63 from *GPI8* null mutant cells had not been affected (Figure 4.14A - lane 2). This indicates that GP63 was not being degraded by the proteasome.

Perhaps degradation occurs via the lysosome pathway. The peptidyl diazomethane cysteine proteinase (CP) inhibitor Z-Phe-Arg-CHN_2 is known to affect CPs of both mammals (Kirschke and Barrett, 1987) and parasitic protozoa (North *et al*, 1990a; Scory *et al.*, 1999), including the known *L. mexicana* CPs. The inhibitor was incubated with wild type, P1, P1S1 and P1S1[pGL269] promastigotes for 48 hours and lysates were analysed for GP63 content by Western blotting (Figure 4.14B). GP63 levels in *GPI8* null mutants, although apparently slightly recovered, were not fully restored when cells were grown in the presence of this inhibitor (Figure 4.14B – lane 2) compared with the control (lane 5) indicating that GP63 is unlikely to be degraded solely by the lysosomal cysteine proteinases. The differences could also be due to slight variations in the amount of material loaded in each lane.

4.3.4 Investigation of potential residues of importance to GPI8

GPI8 belongs to the cysteine proteinase family C13 (Riezman and Conzelmann, 1998), which includes the cysteine proteinase legumain, which was first identified in plants such as Jack Bean (Abe *et al.*, 1993) and later in mammals (Chen *et al.*, 1997). It has been shown that the *T. brucei* GPI:protein transamidase is sensitive to sulphydryl alkylating agents suggesting the involvement of an essential cysteine residue (Sharma *et al.*, 1999a). If GPI8 is a cysteine proteinase, then it most likely possesses the classical dyad residues of cysteine and histidine that would mediate activity. Four potential active site residues of LMGPI8 had been identified previously (see Chapter 3), Cys^{94}, Cys^{216}, His^{63} and His^{174}. These residues are conserved

between all of the GPI8 molecules analysed whereas Cys^{216} and both histidine residues are also conserved in the cysteine proteinase legumain (Figure 3.9). As the *Leishmania* GPI8 is also likely to have an essential cysteine residue, Cys^{216} would seem to be the most likely candidate.

Attempts to mutate these potential active site residues were made (Section 4.2.1.3) and of the four mutations only C216G was successful. This plasmid was designated pGL403. Insufficient time was available for further attempts to produce and confirm the other mutations but the C216G mutant plasmid was transformed into the P1S1 (*GPI8* null mutant) cell line and a single clone was obtained. This was analysed by Western blotting with anti-GP63 antibody and little GP63 could be detected (Figure 4.15 – lane 4). P1S1[pGL403] appears to have a similar phenotype as P1S1 *GPI8* null mutant cells (lane 2). However, to confirm that Cys216 is essential for GPI:protein transamidase activity, it is necessary to show that GPI8 protein is expressed in the P1S1[pGL403] mutant. Unfortunately, monospecific antibodies were not available to confirm the results at the time. This work is ongoing.

4.4 DISCUSSION

In the work described in the previous chapter, several circumstantial lines of evidence pointed to the *L. mexicana* GPI8 being a functional homologue of yeast Gpi8p. These were based on the relatively high degree of homology between these proteins as well as the conservation of particular residues that may be involved in transamidase activity. The aim of this section of work was to provide further evidence that *GPI8* of *L. mexicana* encodes the GPI:protein transamidase by investigating a possible role for the protein in the attachment of GPI anchors to proteins.

PCR and Southern data provide evidence that P1S1 and P1S3 are indeed *GPI8* null mutants. The absence of hybridisation with the *GPI8* ORF probe together with the observed hybridisation with the 5` flank probe, which demonstrated the loss of the *Xho*I site within the ORF, indicated that the *GPI8* ORF had been lost. Further experiments that would support the absence of GPI8 in P1S1 and P1S3 would be the detection of GPI8 directly in wild type cell lysates by Western blotting and the failure to do so in P1S1 and P1S3 extracts. Northern blotting could also be employed to demonstrate the absence of transcript in these mutant lines.

GPI8 null mutant promastigotes, which grow as well as wild type cells in liquid culture, have the ability to infect macrophages, differentiate to amastigotes and grow and replicate within macrophages under *in vitro* conditions. Recent studies suggest that the GPI biosynthesis is essential for growth of *L. mexicana* promastigotes, however it is proposed that it is the free GPIs (glycoinositolphospholipids or GIPLs) that are the essential membrane components (Ilgoutz *et al.*, 1999). It was found that dolichol phosphate mannose synthase (DPMS) is an essential gene catalysing the formation of dolichol phosphate mannose, the sugar donor of mannose addition in all GPI biosynthesis. This includes the production of other GPI structures such as the carbohydrate-linked GPIs (the lipophosphoglycans or LPGs) and the free GPIs called glycoinositolphospholipids (GIPLs). This study supports the idea that it is the free GPIs that are essential for growth of *Leishmania* promastigotes (Ilgoutz *et al.*, 1999) and amastigotes (Mensa-Wilmot *et al.*, 1999) since, in collaborative studies (data not shown) it has been determined that the levels of LPG and GIPL are unaffected in *GPI8* null mutants (Hilley, J.D., Zawadzki,J., McConville, M.J. Coombs, G.H. and

Mottram, J.C. manuscript submitted). This is consistent with the idea that the biosynthesis of these non-protein associated GPIs is independent from the biosynthesis of protein GPIs in *L. mexicana* (Ralton and McConville, 1998).

Nevertheless the mice-infectivity studies show that *GPI8* null mutants have difficulty in establishing infection in BALB/c mice, resulting in lesion volumes of at least 8-fold smaller than the largest wild type lesion. Thus, although GPI-anchored proteins do not seem to be required for parasite survival and macrophage invasion under *in vitro* conditions, they appear to play some role in *in vivo* survival. It is possible that the transfection and selection processes used to generate these mutants could be, in some way, compromising the ability of cells to survive *in vivo* by reduction in virulence. To investigate this *L. mexicana* wild type, P1 (heterozygote), P1S3 (*GPI8* null mutant) and P1S3[pGL269] amastigotes have been purified from mouse lesions and re-injected into BALB/c mice. Increased infection would probably mean that the virulence of the mutant promastigotes was indeed compromised by the transfection/selection process.

Interestingly, episomal expression of *GPI8* could not rescue this *in vivo* mutant phenotype. GPI-anchored proteins may be necessary when amastigotes emerge from dead macrophages into the lesion and are available to re-infect more macrophages. GP63 surface expression is dramatically down-regulated in the amastigote stage of some *Leishmania* species (Bahr *et al.*, 1993) and, to date, only the GPI-anchored promastigote surface antigen 2 (PSA2) family of proteins, also known as GP46 (Lohman *et al.*, 1990; Murray *et al.*, 1989) are believed to be expressed on amastigote surface albeit at a considerably lower level than promastigotes (Handman

et al., 1995). Since antibiotics are not present to provide additional selection pressure, the episome could be lost while the parasites are existing intracellularly. As an alternative approach to determine whether re-expression of *GPI8* can restore virulence, it should be possible to integrate a copy of the gene back into the chromosome, perhaps at the original locus or a different locus such as that of the *L. mexicana* cysteine proteinase c (Bart *et al.*, 1995). This non-essential gene does not contribute to virulence (Bart *et al.*, 1997) and integration at an alternative locus, such as this, would allow a third selectable marker to be used to screen for integration. Integration at the *GPI8* locus would result in the deletion of one of the present antibiotic-resistance markers making screening for transfectants more difficult. Constructs for this are already available (prepared by Dr. Hubert Denise, University of Glasgow); *GPI8* has been ligated into these and the work has begun but re-integration has not yet been achieved.

If the transfection/selection procedure has compromised the virulence of the mutant promastigotes, then differences in *in vivo* infectivity between the *GPI8* null mutants and re-expressers may not be apparent. The observation that the *GPI8* re-expresser mutant lines are apparently even less capable of lesion formation in mice than the *GPI8* null mutants could be explained by the fact that the re-expressers have undergone three rounds of transfection compared with two for the null mutants. The amastigote re-infection experiment described previously may provide further evidence for this.

The *GPI8* heterozygote P1 also produced smaller lesions in BALB/c mice than did wild type parasites, despite the retention of one copy of the gene. A possible

explanation for this stems from the observation with *Trypanosoma cruzi* that, during an artificially induced GPI-deficiency proteins that normally are GPI-anchored have divergent fates (Garg *et al.*, 1997). The fates of four GPI-anchored proteins were examined. Two proteins, gp50/55 and p60 were observed to be absent or decrease over a 20 hour period in cells expressing exogenous GPI-PLC and these proteins could not be detected in medium. This indicates that these proteins were retained and degraded intracellularly. Two other proteins, ssp-4 and p75 were detected in the growth medium indicating secretion. It was proposed that during GPI deficiency, i.e. when one of the transamidase substrates was limiting, GPI-signal sequences could have different "strengths of interaction" with the transamidase. The two secreted proteins interacted more readily with the transamidase, out-competing the other two proteins, which were retained and degraded. Ssp-4 and p75 may have been processed by the transamidase without receiving an anchor, producing soluble forms of each protein. Water could have replaced the anchor as the nucleophile required for completion of the reaction, as was seen with the modified transamidase substrate preprominiPLAP in mammalian cell-free systems (Maxwell *et al.*, 1995a). These surface protein-deficient *T. cruzi* were able to infect macrophages but were unable to replicate. In the case of the *L. mexicana GPI8* heterozygote (P1), perhaps the levels of the transamidase are less than normal since one copy of the gene is absent. If the proteins are competing for the transamidase, then it could be that a protein important for survival of amastigotes in the bloodstream when they emerge from one macrophage to re-infect another is not getting to the cell surface and this results in a less well-established infection of the mice. To investigate whether limiting transamidase levels or perhaps the transfection/selection process causes reduced virulence, two of the three remaining *GPI8* heterozygotes described in Section 4.3.1,

S1 and P2 were inoculated into BALB/c mice to compare with P1. Time constraints for this project mean that the data from this experiment cannot be included here.

GPI-anchored proteins may be essential during the sandfly stages of infection, however passage of these mutants through the sandfly stage of the life cycle remains to be tested.

The ability of the *GPI8* null mutants P1S1 and P1S3 and the re-expresser cell lines P1S1[pGL269] and P1S3[pGL269] to express GPI-anchored proteins on their surface was investigated. GP63 is an excellent protein to study in terms of its anchor attachment because it is the major GPI-linked protein of *Leishmania* promastigotes. The effects of deletion of *GPI8* on this protein were dramatic, with the apparent total loss of GP63 expression on the cell surface as determined by immunofluorescence on whole cells (Figure 4.12). Western blotting confirmed this; only very low levels of GP63 could be detected in *GPI8* null mutants (Figure 4.13). This remaining GP63 could be either unprocessed GP63 or the non-GPI form of GP63, expressed in both promastigotes and amastigotes of *L. mexicana* by the C1 gene class (Medina-Acosta *et al.*, 1989; Medina-Acosta *et al.*, 1993; Ilg *et al.*, 1993). Collaborative work within the laboratory of Dr. Malcolm McConville has demonstrated that incorporation of [^{3}H]ethanolamine into GP63 (i.e. GPI anchor addition) was not detectable in P1S1 mutants, indicating a loss of GPI anchor attachment to GP63 (Figure 4.16). This is fully concurrent with the GP63 immunofluorescence and Western data (Hilley *et al.*, 1999).

The fate of the non-anchored GP63 is not clear. Two possibilities exist; rapid secretion of the unprocessed GP63 or its retention and rapid degradation within the cell. These have not been investigated extensively and further study is required. Attempts to detect the secreted protein in serum-free medium were unsuccessful indicating a lack of secretion. The hydrophobic C-terminal GPI-signal of GP63 is probably not removed in the absence of the transamidase thus causing GP63 to remain associated with the ER membrane and be retained within the cell. A similar case has been reported in mammalian systems where human growth hormone (hGH) fused to the GPI anchor signal sequence of decay accelerating factor (DAF) protein with a mutated cleavage/attachment (ω) site remained unprocessed and was degraded (Moran and Caras, 1992). In contrast, ectopic expression of *T. brucei* PI-PLC caused the cleavage of GPI anchors before they entered the ER, resulting in an artificial GPI-negative phenotype (Mensa-Wilmot *et al.*, 1994). As a result, no GPIs were available for attachment to protein, which led to the secretion of GP63. This situation differs from that of the *GPI8* null mutants in that in the other study the processing of the C-terminal signal sequence could still occur, since the transamidase was still present/active. In yeast and trypanosomal *in vitro* systems, a degree of processing without anchor addition occurs; small nucleophiles such as water, hydrazine and hydroxylamine act as alternatives to the ethanolamine component of the GPI anchor and are able to attack the activated carbonyl formed during transamidation causing the release of processed, but GPI-free protein (Ramalingham *et al.*, 1996; Sharma *et al.*, 1999a; Maxwell *et al.*, 1995b). Interestingly, mutation of the ω site of GP63 (Asn^{577}) also caused GP63 to be secreted although it was not clear if the protein was indeed being processed (McGwire and Chang, 1996). Attempts were made to influence GP63 degradation by inhibition of the *L. mexicana*

proteasome and major lysosomal CPs (Figure 4.14). Inhibition of the proteasome (using MG132) had no apparent effect on the loss of GP63 whilst the inhibition of the cysteine proteinases (by Z-FA) had a slight effect. In both cases further experiments are required to confirm that full inhibition was accomplished, although these conditions have been used successfully in the past (Mutomba *et al.*, 1997; Frame, 1997). Perhaps another class of proteinase degrades GP63. Clearly the fate of GP63 in *L. mexicana GPI8* null mutants merits further study.

Further work within Dr. McConville's laboratory demonstrated an accumulation of GPI anchor precursors in *GPI8* null mutants (Hilley *et al.*, 1999), which is the phenotype observed in yeast *gpi8-1* mutants (Benghezal *et al.*, 1996). Collectively, the data strongly suggest that *L. mexicana* GPI8 is the GPI:protein transamidase but it remains to be seen if all GPI-anchoring of proteins has been prevented in *GPI8* null mutants.

Because of its similarity to the C13 family CPs (Chapter 3) and because *T. brucei* transamidase activity requires an essential cysteine (Sharma *et al.*, 1999a) it is hypothesised that GPI8 is a cysteine proteinase and as such it will probably rely on the classical active site dyad (cysteine and histidine) for activity (Barrett and Rawlings, 1996). The cysteine is required as a nucleophile mediating proteolytic activity. GPI8 belongs to the C13 family of cysteine proteinases, which includes the plant cysteine proteinases known as the asparaginyl endopeptidases (Riezman and Conzelmann, 1998). The most likely residue for the catalytic cysteine of LMGPI8 is Cys^{216} as it is conserved between the three GPI8 proteins (yeast, human and *L. mexicana*) as well as legumain. When the C216G mutant was used to transfect the *GPI8* null mutant cell

line P1S1, it was found that surface GPI-linked GP63 could not be detected in this cell line. GPI8 antibodies capable of detecting the protein in cell lysates were not available (Chapter 5), thus the only evidence that the mutant gene was expressed lies in the fact that the cells were resistant to Geneticin (the antibiotic-resistance marker encoded from the episome). If the protein was actually expressed in the cell line then it is not active providing evidence that Cys^{216} is the active site cysteine. However, Cys^{216} could be involved in disulphide bridge formation to maintain protein tertiary structure rather than catalysis itself and there are a total of 9 cysteine residues within the protein. Clearly it is important to mutate these other cysteines to see if they have any effect. If not, then we have identified the active site residue. Since nothing is known about the three dimensional structures of GPI8, it is not possible to predict which residues, if any, are involved with disulphide bridge formation.

The wild type GPI8 and C216G GPI8 could be expressed as 6xHis fusion proteins in *E. coli* and both enzymes could be assessed for activity using an adapted version of GPI-anchoring cell free assay (Sharma *et al.*, 1999a).

The validity of GPI8 as a potential anti-parasite drug target is open to question, however, the decreased ability of the parasites to infect *in vivo*, associated with the loss of protein GPI-anchoring suggests that the enzyme may still be exploitable. A non-essential gene can still encode a potential target (Barrett *et al.*, 1999). Cysteine proteinase inhibitors have been used effectively against *Leishmania* (Selzer *et al.*, 1997) despite the fact that, of all the CPs identified to date, none have been shown to be essential (Souza *et al.*, 1994; Mottram *et al.*, 1996a; Bart *et al.*, 1997). Inhibition of a *T. cruzi* cysteine proteinase was shown to prevent its autocatalytic processing,

leading to accumulation of the precursor molecule in the Golgi complex (Engel *et al.*, 1998). This had a lethal effect as the build up prevented normal functioning, which in turn led to cell death.

Although *GPI8* may not be essential in *Leishmania*, the gene is likely to be essential in the related organism *T. brucei*, which depends on GPI anchors for the attachment of the variant surface glycoproteins (VSGs), used in the avoidance of the host immune response (see (Barry, 1997) and Pays and Nolan (1998) for reviews). *L. mexicana GPI8* null mutants could be exploited as a model system for the study of the *T. brucei* homologue (*TBGPI8*). Once *TBGPI8* has been isolated and characterised, it can be determined if it is able to functionally replace the *L. mexicana* gene. If so, the gene can then be subject to modifications and the GPI-anchoring ability can be assessed using GP63 surface expression as a marker in a transamidase null background (Alternatively, VSG could be expressed in *Leishmania* promastigotes providing a more "natural" substrate). Fluorescence due to anti-GP63 antibody conjugated to FITC can be quantified by flow cytometry (Joshi *et al.*, 1998) and therefore GP63 surface expression can be quantified. The efficacy of potential inhibitors of the transamidase could also be determined in this system.

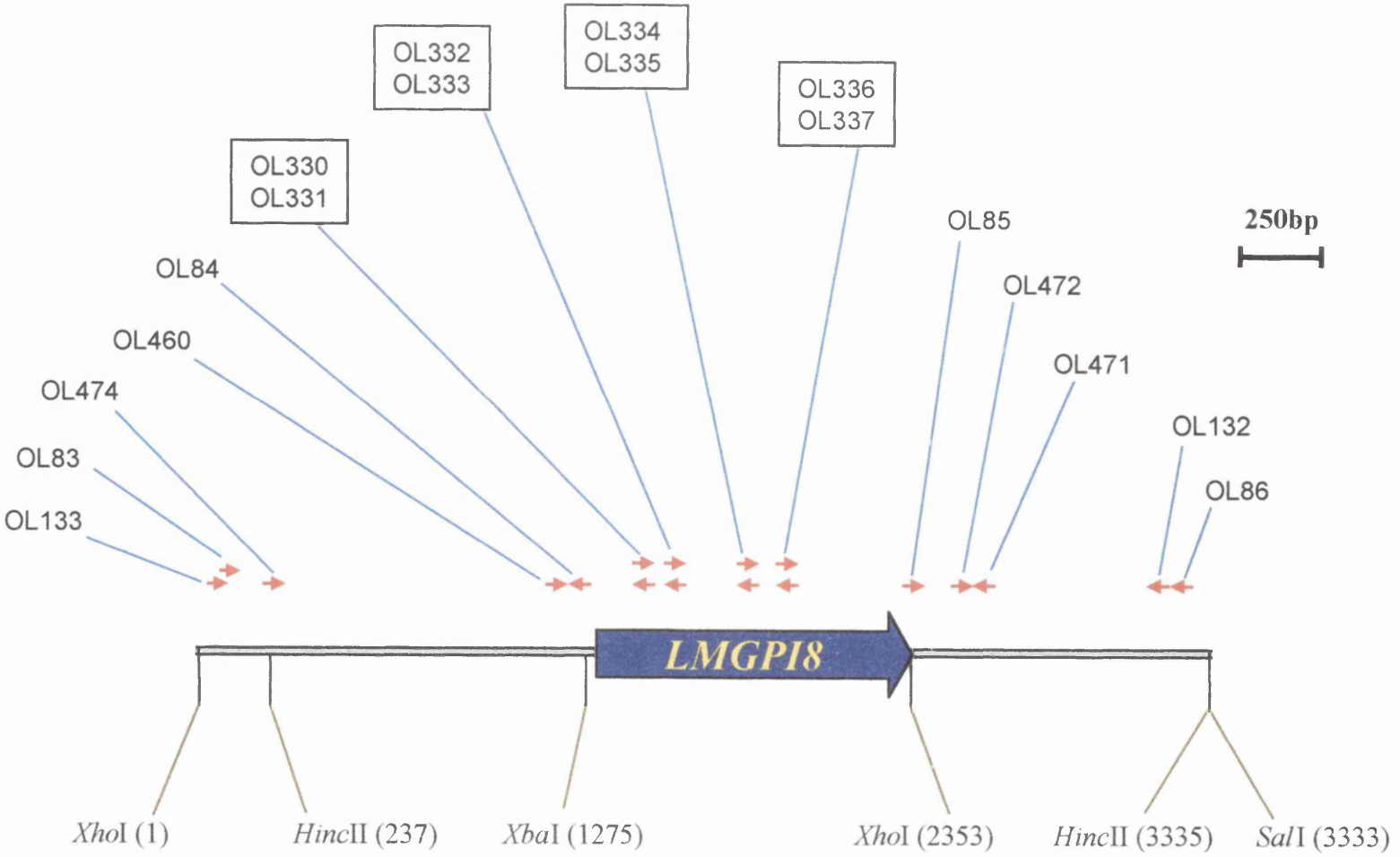


Figure 4.1: Schematic of the *GPI8* locus showing positions of primers described in Chapter 4. Red arrows mark the position and orientation of primers used. The primer pairs used for mutagenesis of potential active site residues are indicated by boxed text.

Figure 4.2: Schematic of constructs containing *PAC* and *SAT* genes flanked by *GPI8* 5` and 3` flanking regions. The constructs pGL236 and pGL237 were synthesised from the existing pX53*PAC* (pGL207) and pX53*SAT* (pGL51) constructs, by removing the existing 5` and 3` flanks of the antibiotic resistance genes (*DHFT-TS* and *LMCPB* flanking sequences) by restriction digests and replacing them with those of *GPI8*. A map of the constructs and predicted fragment sizes are shown (**A**). Constructs were tested by restriction digest to confirm that they were complete. Digests were electrophoresed on a 0.8% agarose gel and stained with ethidium bromide (**B**).

Lanes 1-5 were pGL236 digests:

1 *Hin*dIII/*Bgl*II,

2 *Hin*dIII/*Spe*I,

3 *Bam*HI/*Bgl*II,

4 *Spe*I/*Bgl*II,

5 *Sma*I only

Lanes 6-10 were pGL237 digests:

6 *Hin*dIII/*Bgl*II,

7 *Hin*dIII/*Spe*I,

8 *Bam*HI/*Bgl*II,

9 *Spe*I/*Bgl*II,

10 *Sma*I only

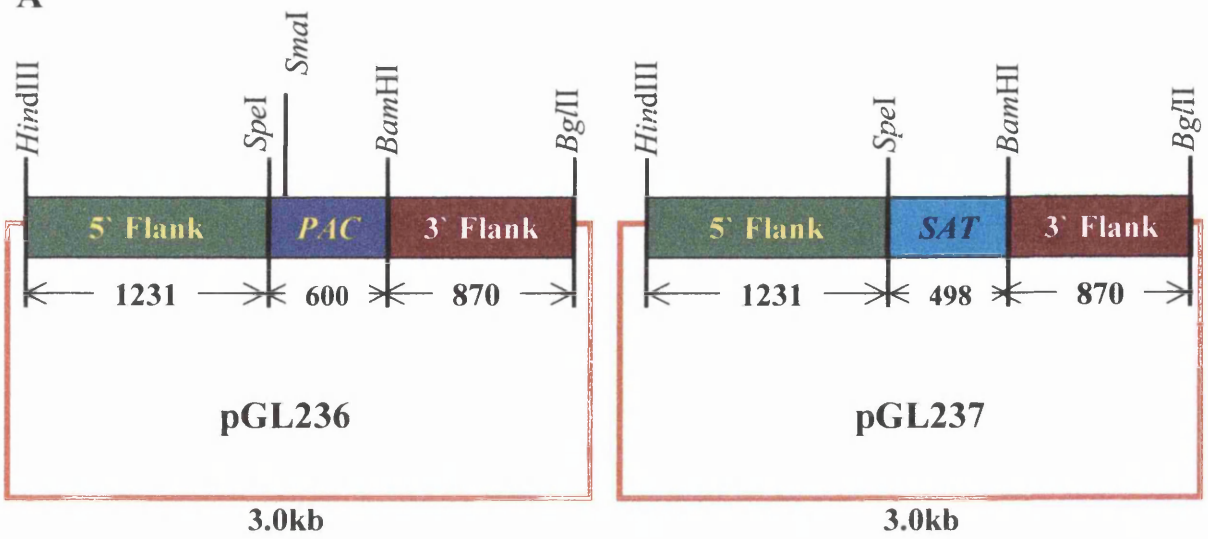
Lanes 11 and 12 were uncut plasmid:

11 pGL236 (uncut),

12 pGL237 (uncut)

DNA markers (**M**) used: 1kb ladder (GibcoBRL).

A



Predicted Restriction Fragments

Restriction Digests	pGL236 fragments	pGL237 fragments
*Hin*dIII/*Bgl*II	2.7kb & 3.0kb	2.6kb & 3.0kb
*Hin*dIII/*Spe*I	1.2kb & 4.5kb	1.2kb & 4.4kb
*Bam*HI/*Bgl*II	0.9kb & 4.8kb	0.9kb & 4.7kb
*Spe*I/*Bgl*II	1.5kb & 4.2kb	1.4kb & 4.2kb
*Sma*I	5.7kb (linearised plasmid)	uncut plasmid

B

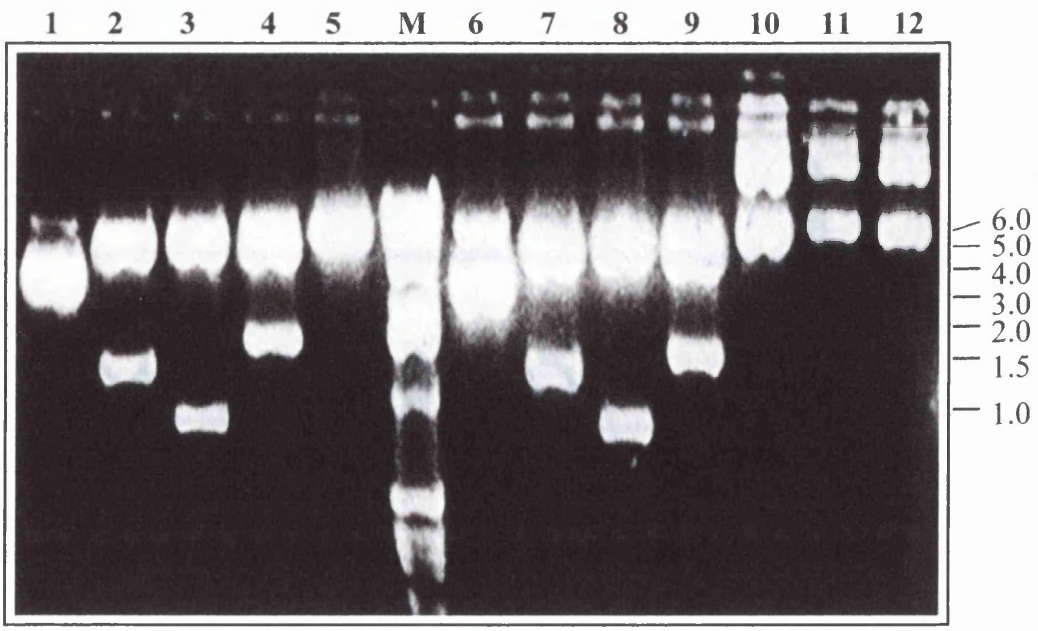


Figure 4.3: Construction of episomal copy of *GPI8* in pXG. The gene together with 5` and 3` flanks was not available on a single construct and had to be produced by cloning fragments of the gene from the two plasmid subclones generated from the original lambda clones. pGL185 contained the 2011bp *Sal*I fragment and pGL186 contained the 2351bp *Xho*I lambda fragment. 1 **pGL185** was digested with *Xho*I to remove the majority of the *GPI8* ORF and phosphatase-treated. The *Xho*I digested plasmid was gel purified. 2 **pGL186** was digested with *Xho*I to drop out the 5` flank and *GPI8* ORF (shown as green block), which was gel-purified. 3 The *Xho*I digested pGL185 was ligated with the *Xho*I fragment from pGL186 and used to transform *E. coli* producing **pBS-*GPI8***. Ampicillin-resistant colonies were screened for both cloning and orientation using specific primers (OL460 and OL471) marked on pBS-*GPI8* by yellow arrows. 4 The *Hin*cII fragment from pBS-*GPI8* was gel-purified and cloned into *Sma*I-digested, gel-purified **pXG** 5 to give **pGL269** 6. Primers OL472 and T3 were used to check for cloning and orientation.

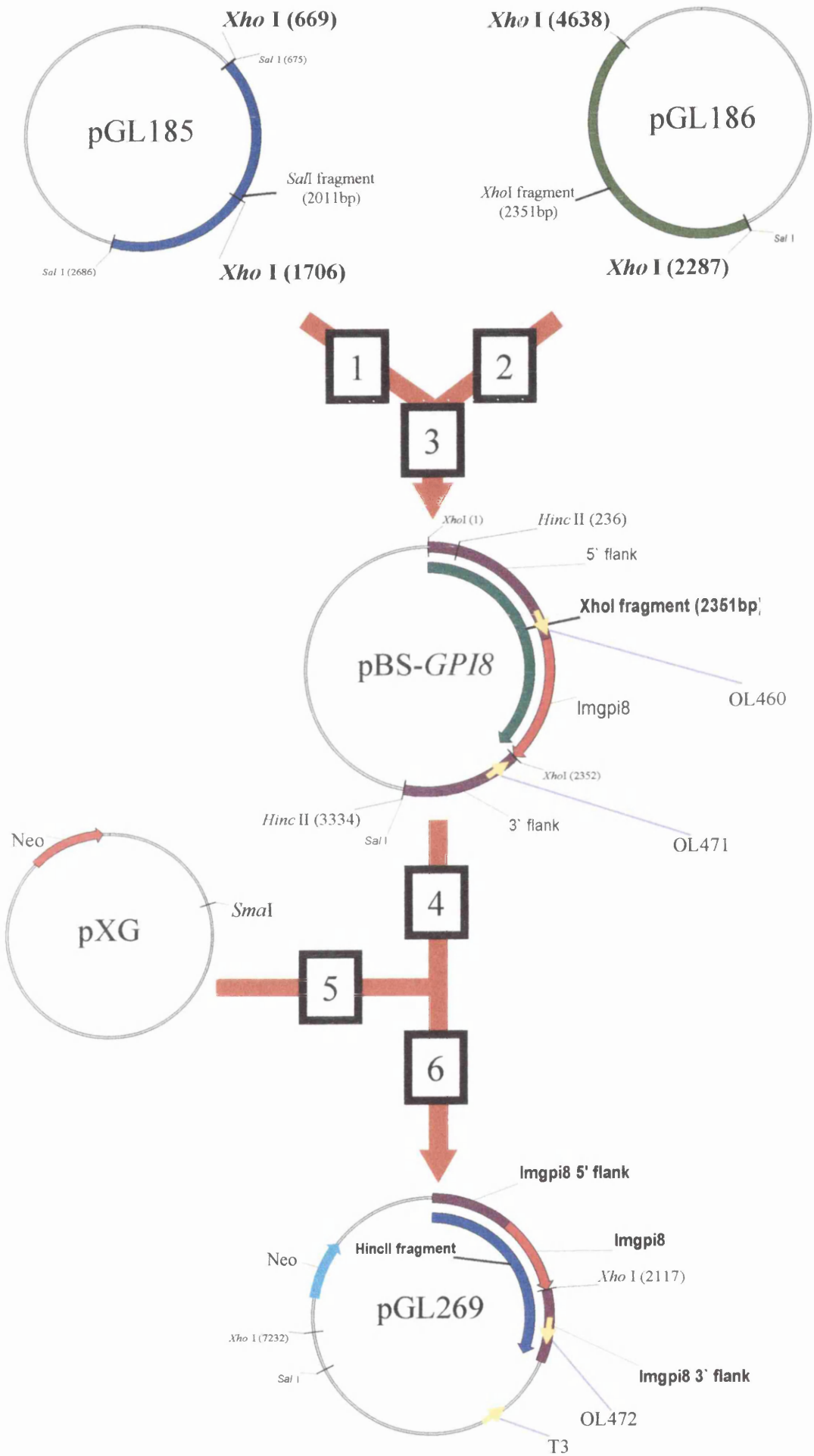

Xho I (669)
Sal I (675)
pGL185
SalI fragment
(2011bp)
Sal I (2686)
Xho I (1706)
Xho I (4638)
pGL186
XhoI fragment
(2351bp)
Sal I
Xho I (2287)
1
2
3
XhoI (1)
HincII (236)
5` flank
XhoI fragment (2351bp)
pBS-GPI8
OL460
Imgpi8
XhoI (2352)
HincII (3334)
SalI
3` flank
OL471
Neo
pXG
SmaI
4
5
6
Imgpi8 5' flank
Imgpi8
Neo
HincII fragment
Xho I (2117)
pGL269
Xho I (7232)
Sal I
Imgpi8 3` flank
OL472
T3

Figure 4.4: PCR analysis of clones derived from first round transfection. Nourseothricin- or puromycin-resistant clones were isolated and grown to near-stationary phase. 10µl of each cell line (wild type, P1, P2, P3 and S1) was added directly to 100µl of sterile water for hypotonic lysis. 2µl of each lysate was used as DNA template in 20µl PCR reactions with Taq polymerase (Applied Biosystems) and primers OL133 and OL86 under the following conditions:

94°C, 4 minutes	x1
94°C, 45 seconds 60°C, 45 seconds 72°C, 3 minutes	x28
72°C, 7 minutes	x1

A map of the *GPI8* locus showing primer positions is shown **(A)**.

The reactions were electrophoresed on a 0.8% agarose gel, which was stained with ethidium bromide (**B**):

Lane 1: wild type

Lane 2: P1

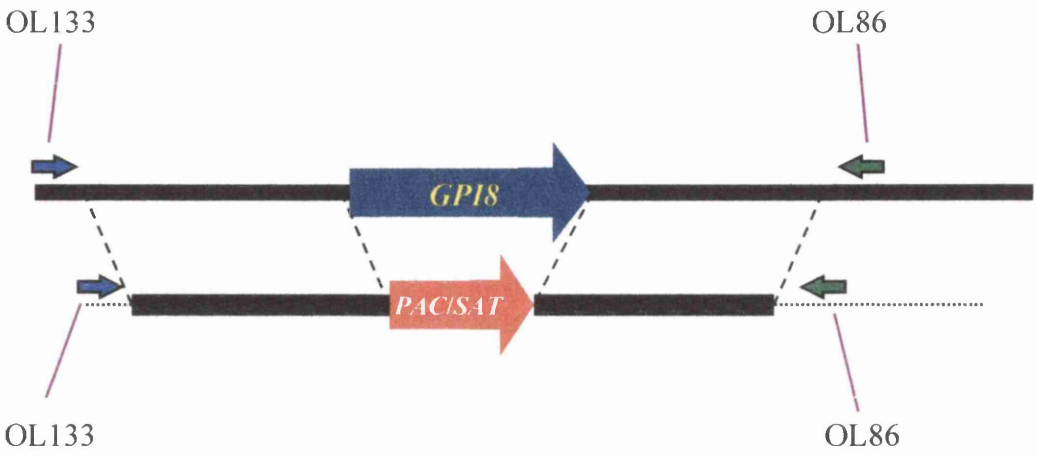
Lane 3: P2

Lane 4: P3

Lane 5: S1

DNA markers used were 1kb Ladder (GibcoBRL)

A



B

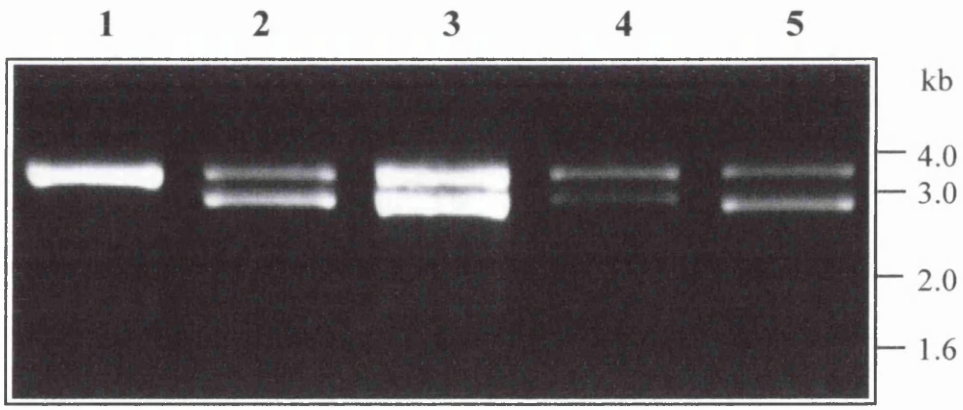


Figure 4.5: PCR analysis of clones derived from second round transfections. Puromycin- and nourseothricin- resistant promastigote clones were isolated on solid medium and subjected to whole cell PCR with primers OL133 and OL86 as described in Figure 4.4. PCR conditions were exactly as described in Figure 4.4 legend.

(A) PCR of double-resistant clones from transfections with P2 and P3 *GPI8* heterozygotes.

Lane 1: Wild type

Lane 2: P2S1

Lane 3: P2S2

Lane 4: P2S3

Lane 5: P3S1

Lane 6: P3S2

(B) PCR analysis of three double-resistant clones from transfection with the P1 *GPI8* heterozygote.

Lane 1: Wild type

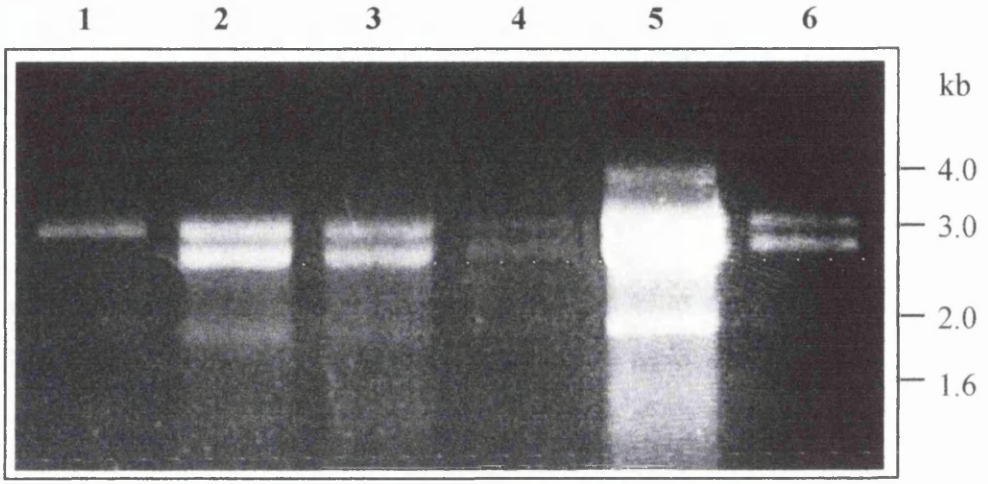
Lane 2: P1S1

Lane 3: P1S2

Lane 4: P1S3

DNA Markers: 1 kb Ladder (GibcoBRL)

A



B

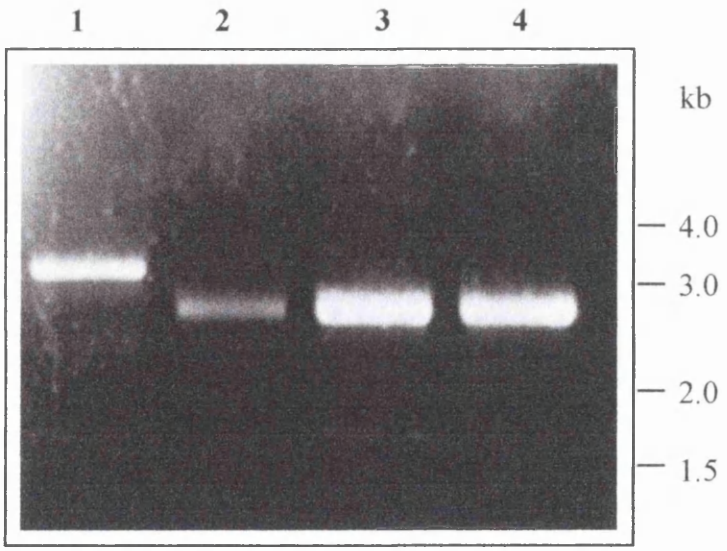


Figure 4.6: Southern analysis of *GPI8* mutants. The schematic structure of the *GPI8* locus is shown with the predicted size of *Xho*I fragments (**A**). 5μg of gDNA isolated from wild type, P1, P1S1 and P1S3 cells was digested with *Xho*I, electrophoresed on a 0.7% agarose gel, which was stained with ethidium bromide (**B**), and blotted onto Hybond N (Amersham). Markers used were 1kb Ladder (GibcoBRL).

(**C**) The blot was probed with [α]^{32}P-CTP labelled *GPI8* ORF and hybridisation was detected by autoradiography. The *GPI8* ORF was prepared by digesting pGL136 (see Figure 4.5) with *Xho*I and *Xba*I and gel-purifying the 1.1kb fragment prior radio-labelling.

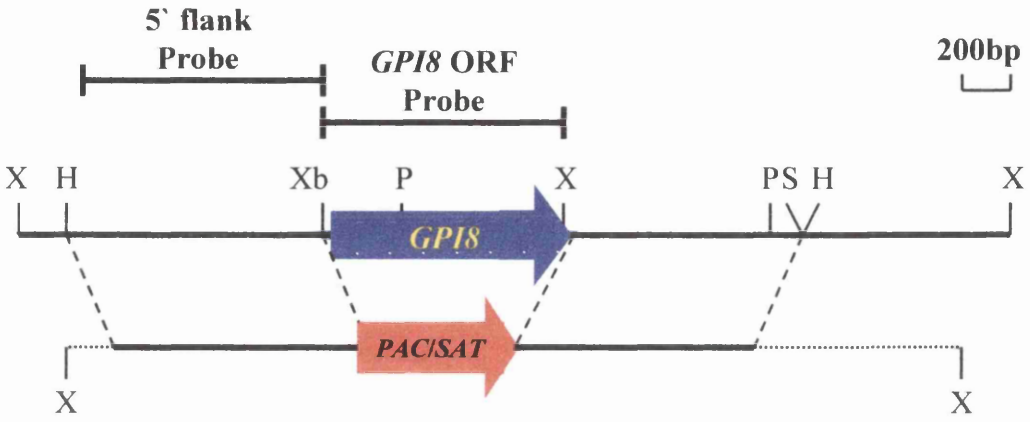
(**D**) This blot was stripped by the boiling SDS method and reprobed with similarly labelled *PUR* gene.

(**E**) More gDNA from these cell lines was *Xho*I digested, electrophoresed and blotted and this time was probed with the *SAT* gene.

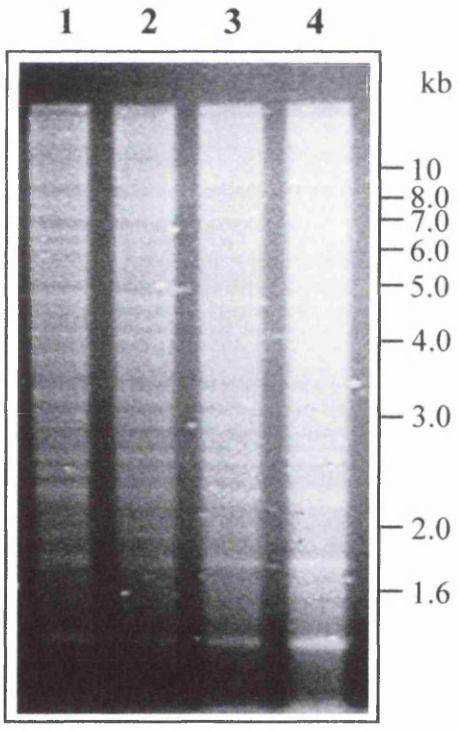
The *PUR* and *SAT* genes were prepared for hybridisation by digesting the KO constructs pGL236 and pGL237, respectively, with *Spe*I and *Bam*HI (see Figure 4.1) and gel-purifying the fragments.

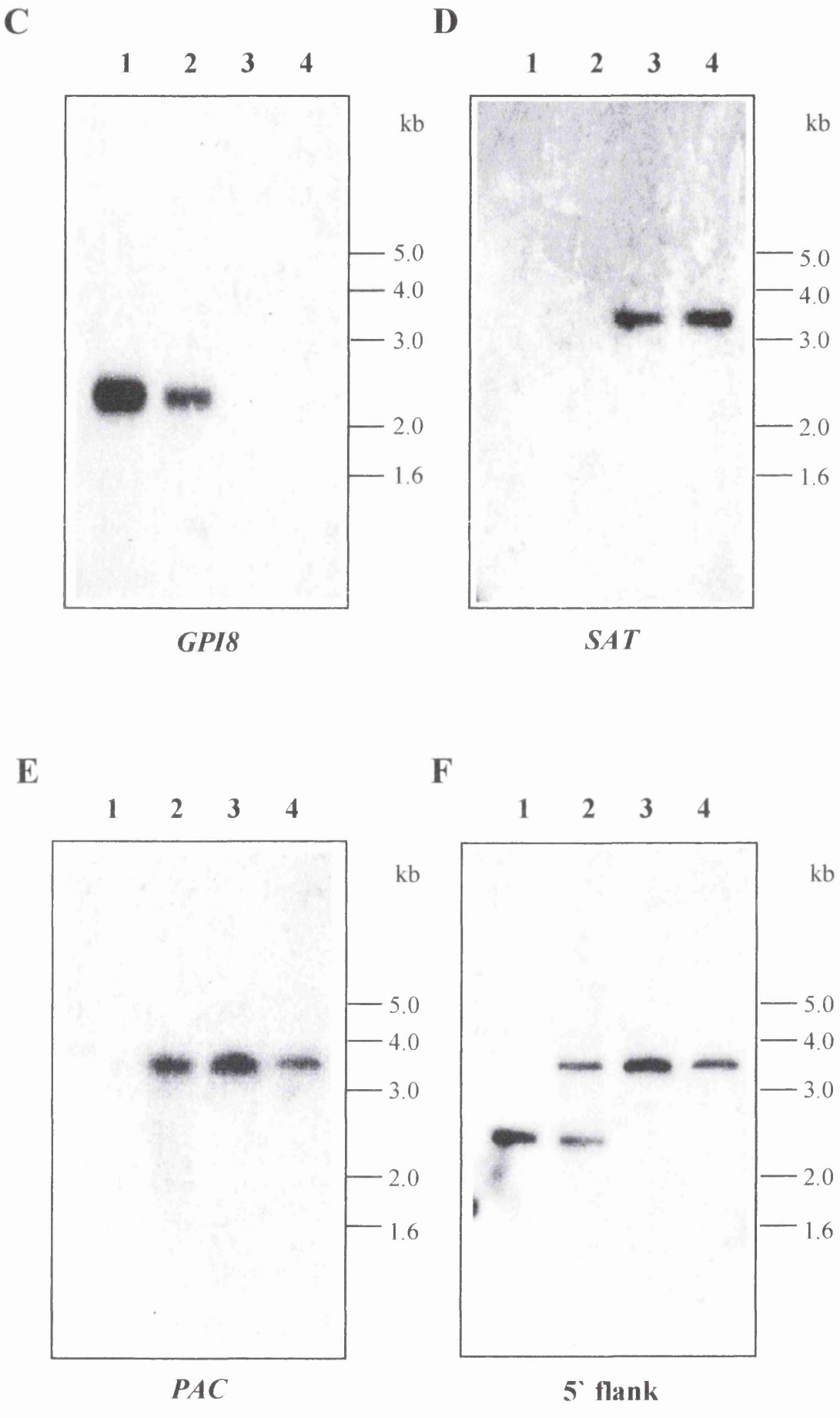
(**F**) Finally, the blot was stripped as before and reprobed with the 5` flanking region of *GPI8*. The 5` flank probe was produced by PCR (OL83 and OL84), using pGL186 as a template. The PCR product was gel-purified prior to labelling.

A


5` flank
Probe
GPI8 ORF
Probe
200bp
X H Xb P X PS H X
GPI8
PAC/SAT
X
X

B


1 2 3 4
kb
10
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.6


C
1 2 3 4
kb
5.0
4.0
3.0
2.0
1.6
GPI8
D
1 2 3 4
kb
5.0
4.0
3.0
2.0
1.6
SAT
E
1 2 3 4
kb
5.0
4.0
3.0
2.0
1.6
PAC
F
1 2 3 4
kb
5.0
4.0
3.0
2.0
1.6
5` flank

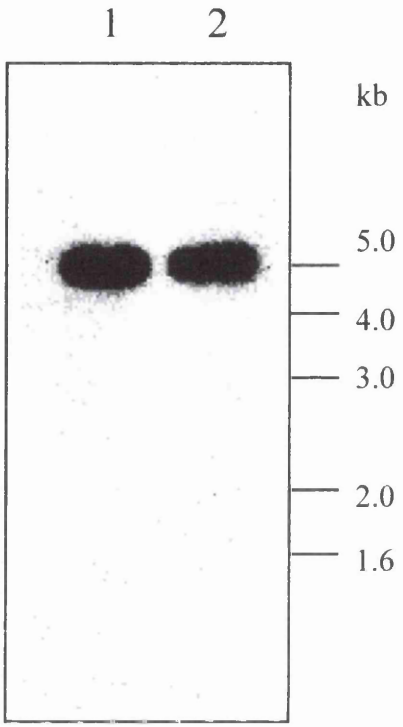


Figure 4.7: Southern analysis of the *GPI8* null mutants, transfected with an episomal copy of the gene. 5μg of total DNA, isolated from clones P1S1[pGL269] and P1S3[pGL269], were digested with *Xho*I, electrophoresed on a 0.7% agarose gel and blotted. The blot was probed with *GPI8* ORF (Figure 4.4) and exposed to X-ray film overnight.

Lane 1: P1S1[pGL269]

Lane 2: P1S3[pGL269].

DNA markers: 1kb Ladder (GibcoBRL).

Figure 4.8: Flow cytometry was used to assess DNA content of *GPI8* mutant cell lines. Mid-log phase promastigotes were harvested (2000xg, 5 minutes at 4°C) and washed in 10ml PBS. Cells were fixed in 70% methanol/30% PBS for an hour. Cells were pelleted by slow-speed centrifugation (1000xg, 10 minutes at 4°C), washed in 10ml PBS before resuspending in 1ml of PBS + 10μg/ml RNAse A and 10μg/ml propidium iodide and incubation for 45 minutes at 37°C. Stained cells were analysed using a Becton Dickinson FACScalibur flow cytometer. Fluorescence was recorded for 10000 cells from each cell line and the data were plotted using the CellQuest software (Becton Dickinson). The **2N** peak represents the fluorescence due to diploid cells and the **4N** peak represents fluorescence due to dividing cells that have doubled their DNA content but have yet to undergo mitosis.

The following cell lines were analysed:

WT

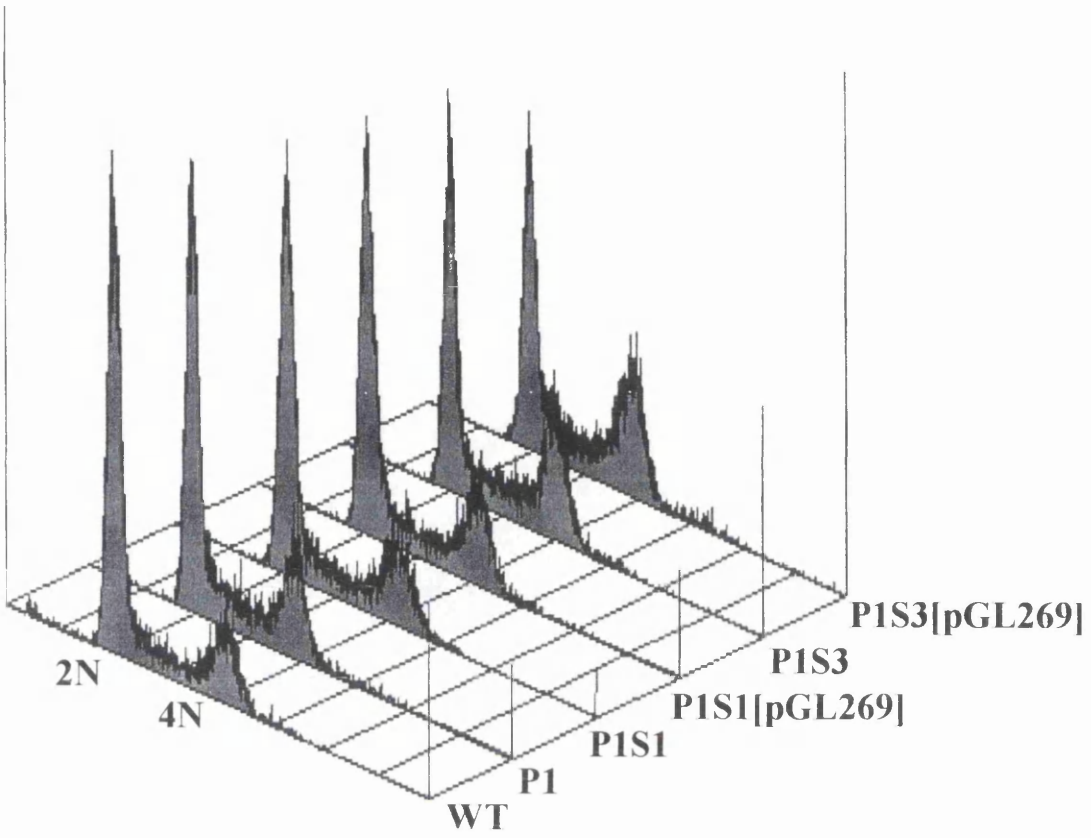
P1

P1S1

P1S1[pGL269]

P1S3

P1S3[pGL269]


2N
4N
WT
P1
P1S1
P1S1[pGL269]
P1S3
P1S3[pGL269]

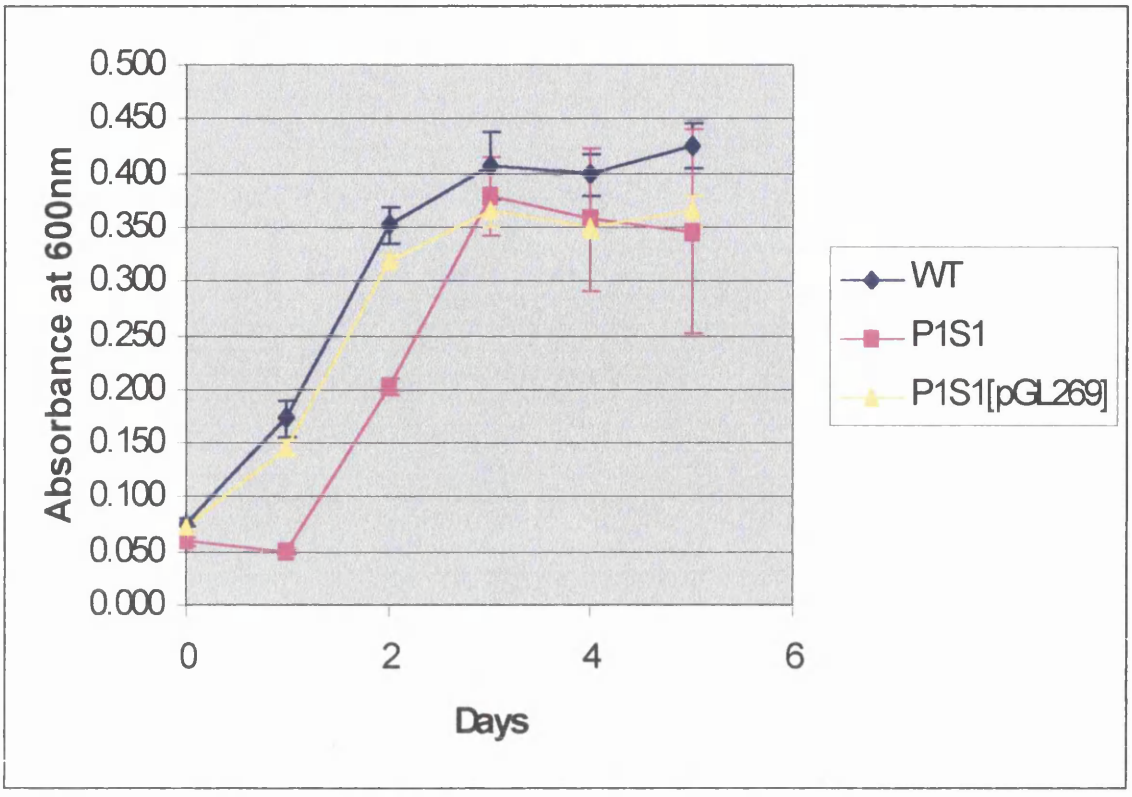


Figure 4.9: The rate of growth of *GPI8* null mutant and re-expresser cell lines was compared with wild type cells. Wild type, P1S1 and P1S1[pGL269] were inoculated into 3x12 ml cultures each, at a density of $1x10^5$ cells ml^{-1}. Absorbance readings at 600nm were taken daily with a spectrophotometer after the cells were fixed by incubating 950µl culture with 50µl 40% formaldehyde. Culture medium was used as a blank.

Data are means ± standard deviation from 3 replicates.

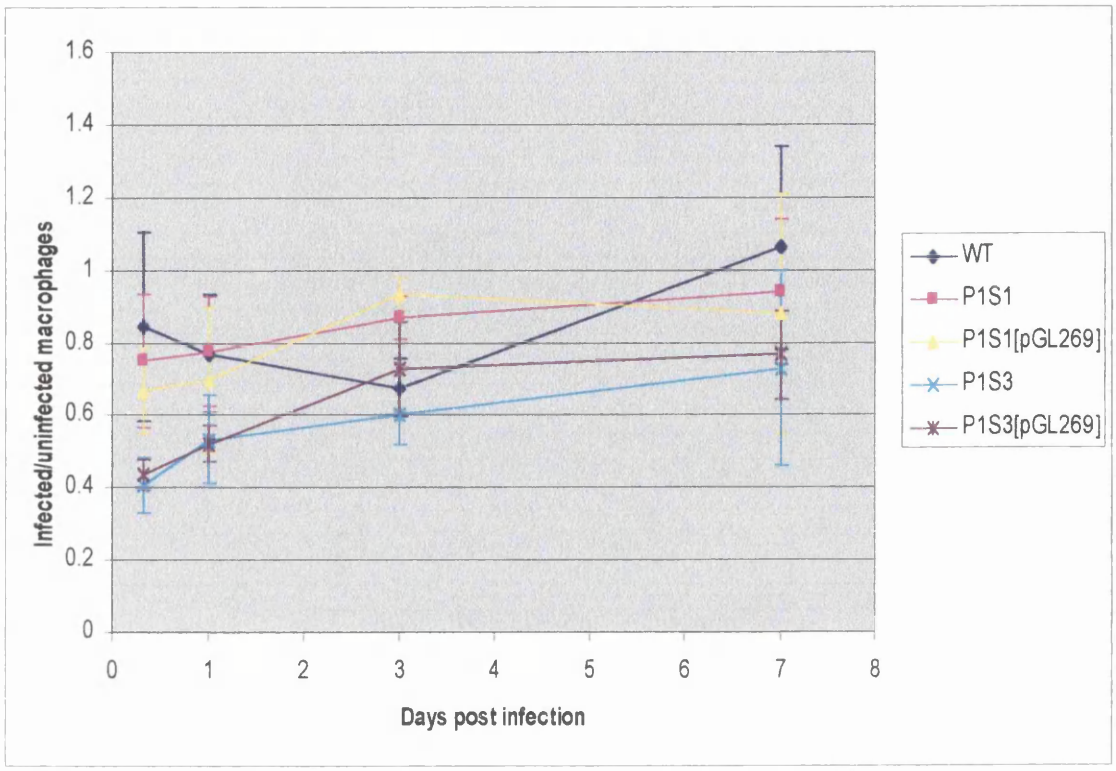
Figure 4.10: *In vitro* macrophage infection by wild type, *GPI8* null mutants and *GPI8* re-expresser cell lines. $2x10^5$ macrophages (Peritoneal exudate cells harvested from freshly killed mice) were added to each chamber of an eight chamber slide (Labtek) and adhered to the slide overnight before exposure to *L. mexicana* promastigotes from each cell line at a 1:1 ratio. After four hours incubation, cells were washed to remove free promastigotes and incubated at 32°C for seven days, with infections being monitored at 8 hours, 1 day, 3 days and 7 days. Cells were methanol-fixed at each time point. Slides were stained with Giemsa's stain and viewed by light microscopy. Greater than 100 macrophages were counted for each cell line, and the number of infected macrophages recorded together with the number of amastigotes per infected cell. The experiment was repeated three times.

(**A**) The ratio of infected to uninfected macrophages was plotted to allow comparison of the ability of each cell line to infect macrophages.

(**B**) The average number of amastigotes per infected macrophage was plotted against time.

Data are means ± standard deviation from 3 experiments.

A


Infected/uninfected macrophages
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5
6
7
8
Days post infection
WT
P1S1
P1S1[pGL269]
P1S3
P1S3[pGL269]

B

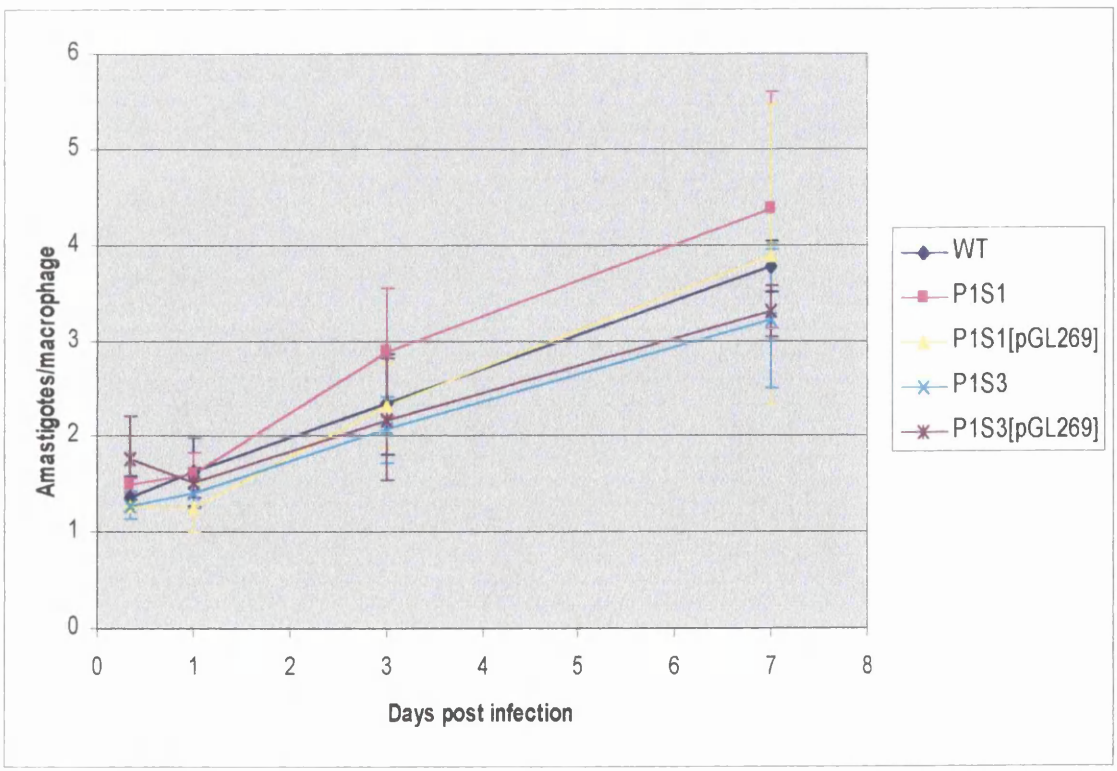

Amastigotes/macrophage
6
5
4
3
2
1
0
0
1
2
3
4
5
6
7
8
Days post infection
WT
P1S1
P1S1[pGL269]
P1S3
P1S3[pGL269]

Figure 4.11: BALB/c mice were inoculated with the *GPI8* mutant cell lines described below, to compare virulence. Each cell line was inoculated into 6 mice and infection was monitored over a period of 47 weeks by the University of Glasgow Animal Facility. Lesion volumes were recorded on a weekly basis as follows:

Lesions are assumed to be half-spheres and the volume of a half-sphere is calculated by $((4/3\pi)/8)$ x length x width x height. $(4/3\pi)/8 = 0.5236$.

Therefore **Lesion volume = 0.5236 x length x width x height**

(A) Wild type

(B) *GPI8* heterozygote, P1

(C) *GPI8* null mutant, P1S1

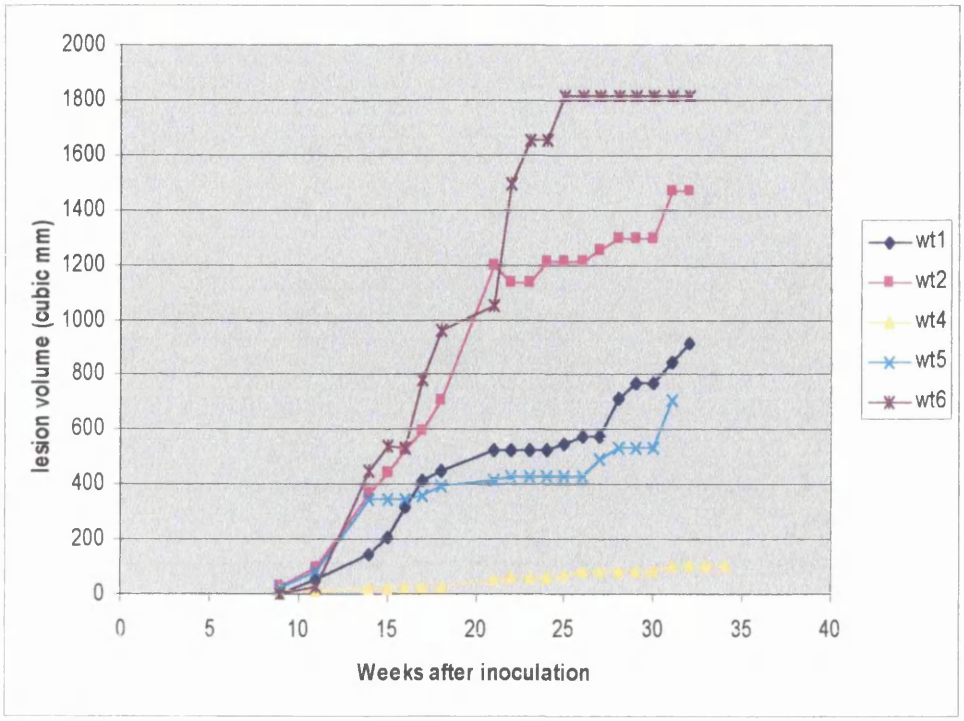
(D) *GPI8* null mutant, P1S3

(E) *GPI8* re-expresser, P1S1[pGL269]

(F) *GPI8* re-expresser, P1S3[pGL269]

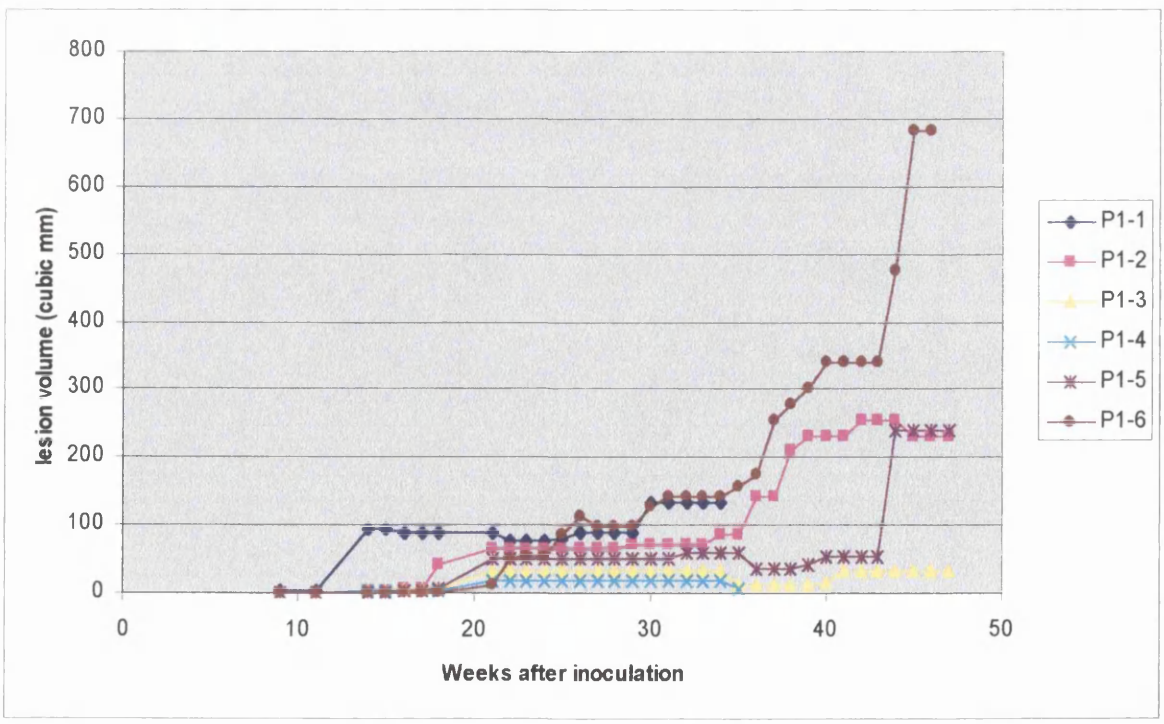
A

Wild type mice infections



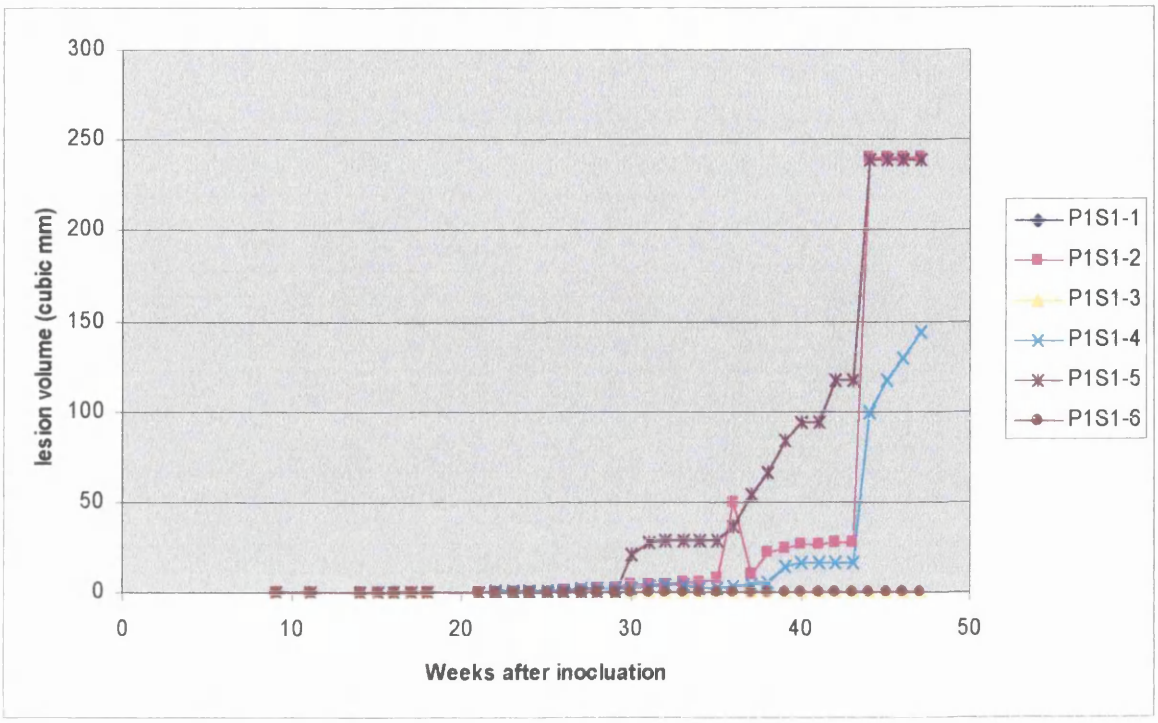
B

Heterozygote (P1) mice infections



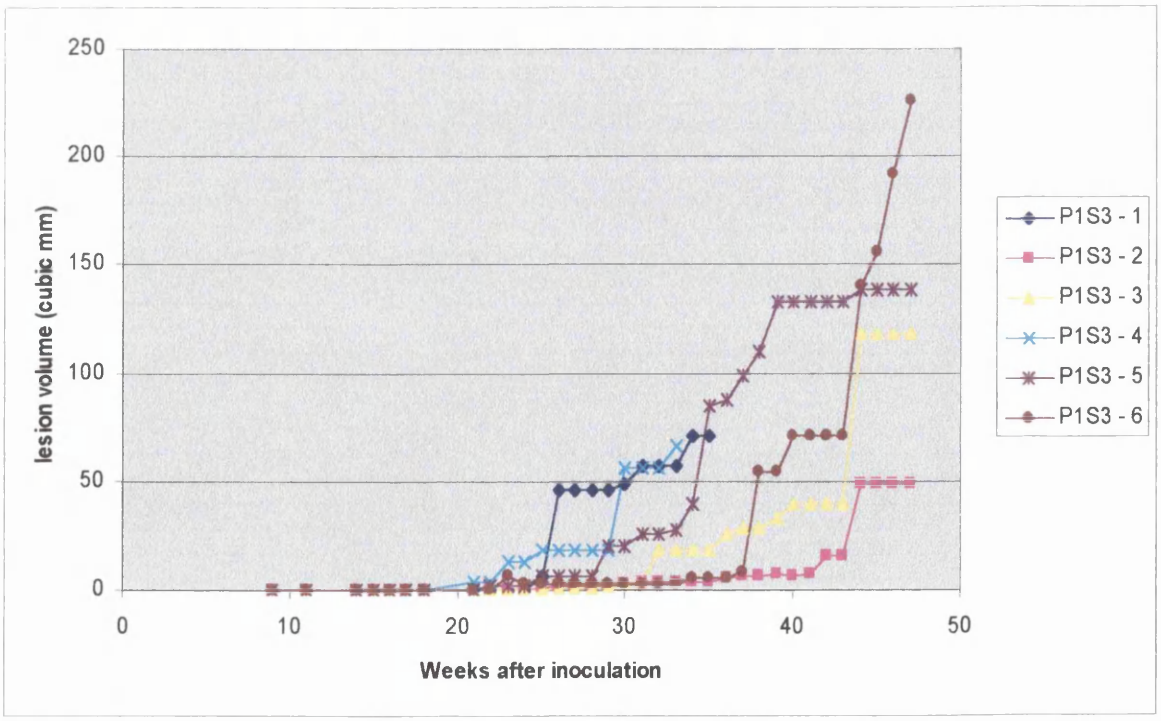
C

GPI8 null mutant (P1S1) mice infections



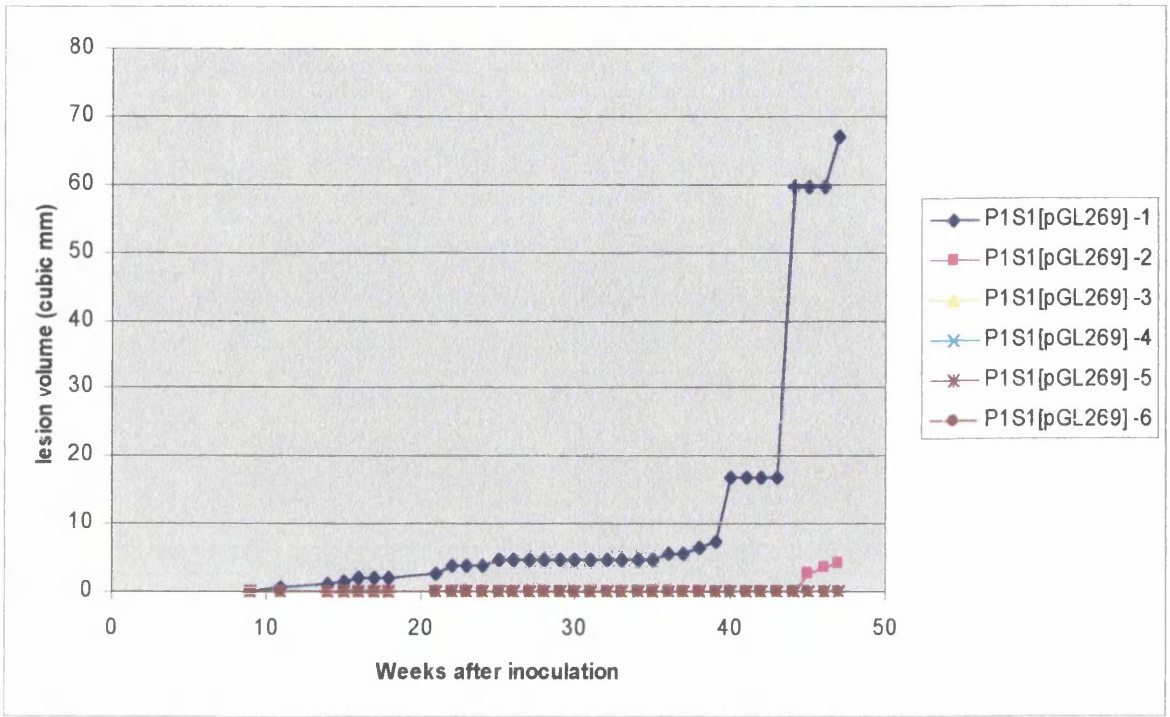
D

GPI8 null mutant (P1S3) mice infections



E

***GPI8* re-expresser (P1S1[pGL269]) mice infections**



F

***GPI8* re-expresser (P1S3[pGL269]) mice infections**

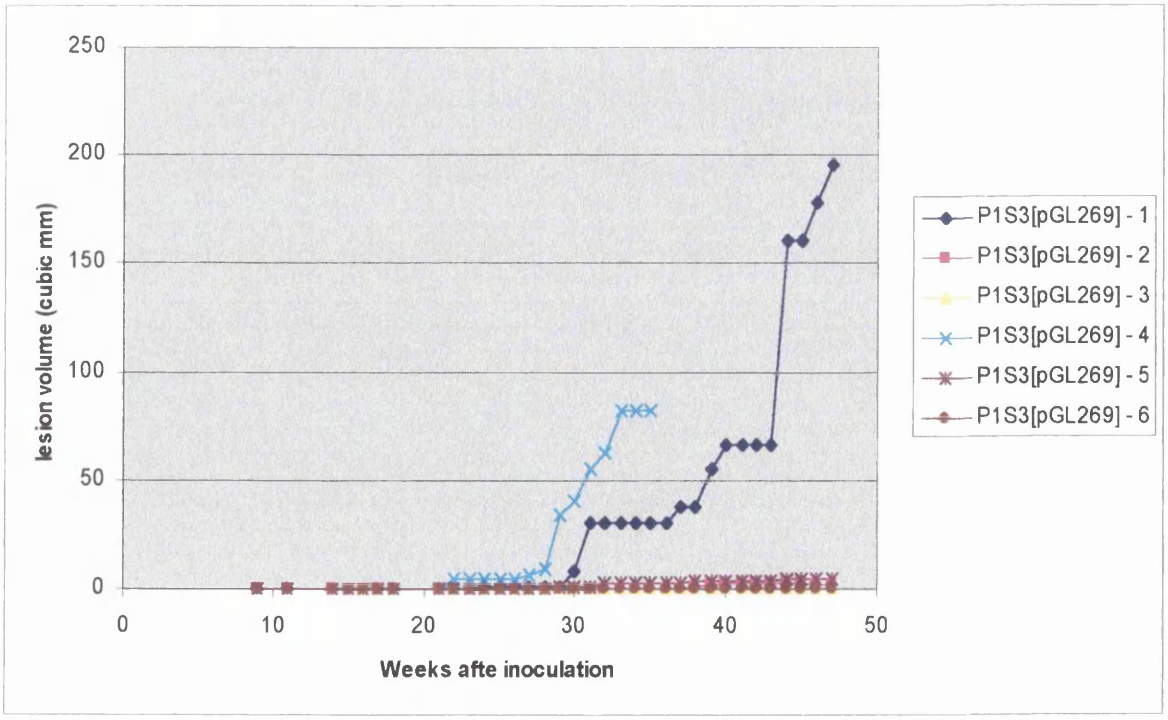


Figure 4.12: Immunofluorescence of GP63 in the *GPI8* null mutant. GP63 was detected by immunofluorescence with anti-GP63 antibodies. Wild type, P1S1 and P1S1[pGL269] cells were smeared on slides and fixed with paraformaldehyde before exposure to a 1 in 100 dilution of L3.6 monoclonal antibody (raised against GP63 from *L. mexicana* (Medina-Acosta *et al.*, 1989)) in PBS for 30 minutes or, as a negative control, just PBS. Cells were washed and blocked (10% HI-FCS in PBS for 30 minutes) before exposure to a 1 in 500 dilution of secondary antibody (anti-mouse IgG-FITC conjugate (Sigma)) and 20μg/ml DAPI in PBS for 30 minutes in the dark. After washing in PBS, 20μl of the anti-quenching agent MOWIOL/DABCO was dropped onto each slide under a glass coverslip and cells were viewed using a Zeiss Axioplan fluorescence microscope with a Hamamatsu Digital Camera using the OpenLab Software Package (Improvision, University of Warwick).

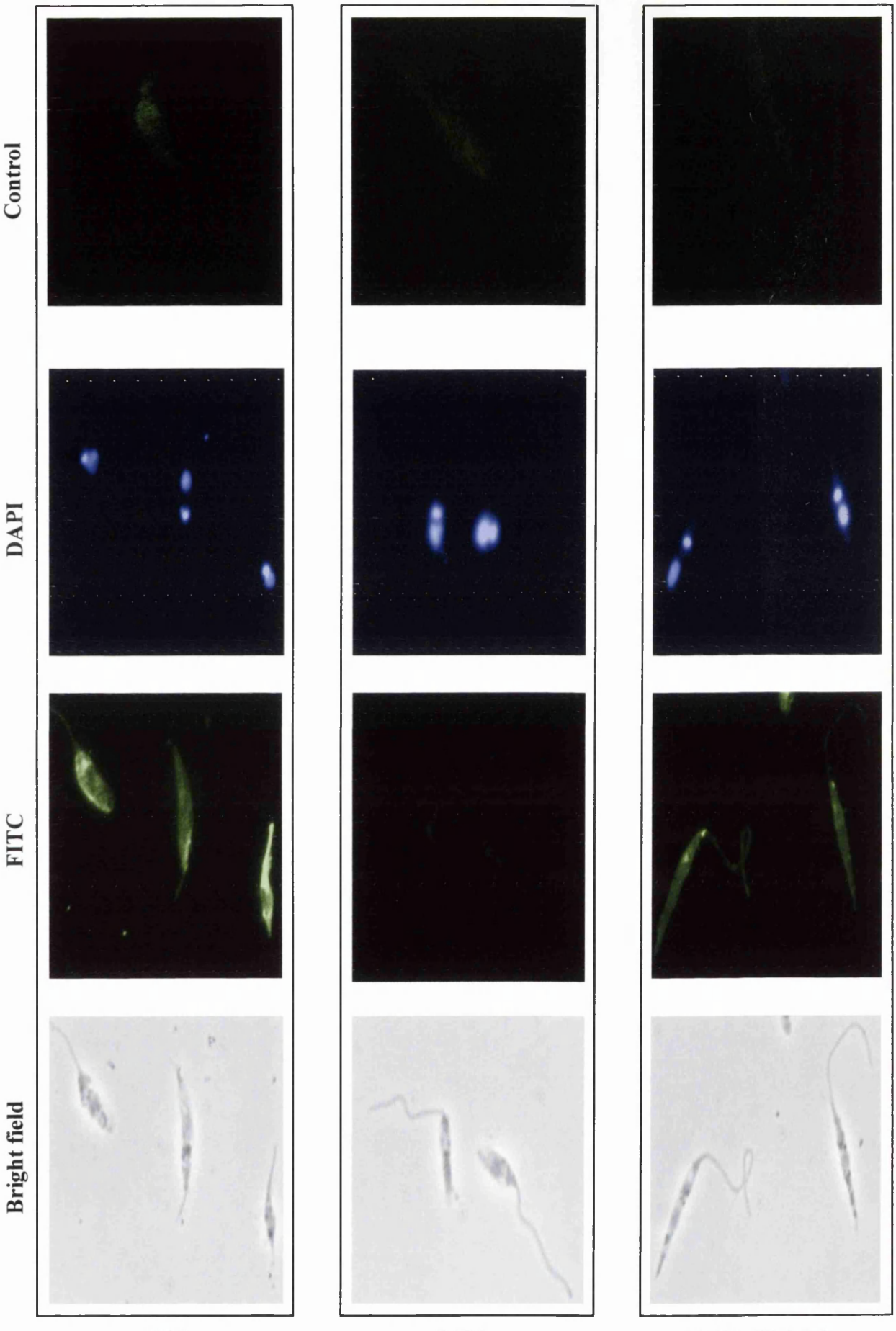

Control
DAPI
FITC
Bright field
WT
P1S1
P1S1[pGL269]

Figure 4.13: Western analysis of GP63 in *GPI8* null mutants. GP63 was analysed by Western blotting. Cells were lysed in 0.25% Triton X-100, Laemmli sample buffer and lysates were electrophoresed (SDS-PAGE) and electroblotted onto Hybond C (Amersham). Blots were exposed to a mouse monoclonal antibody raised against *L. major* GP63 (after overnight blocking) at a 1 in 25 dilution in blocking buffer for 60 minutes. Secondary antibody (anti-mouse IgG-HRP conjugate (Promega)) was used at a 1 in 5000 dilution and antibody binding was detected by ECL (Pierce) (**A**). The blot was overexposed to detect any faint bands (**B**). The same lysates were electrophoresed, electroblotted and exposed to affinity-purified anti-CRK1 antibody (1 in 50 dilution) as a loading control (**C**).

Lane 1: wild type

Lane 2: P1

Lane 3: P1S1

Lane 4: P1S1[pGL269].

Molecular weight markers: MultiMark Multi-Colored Standard (Novex)

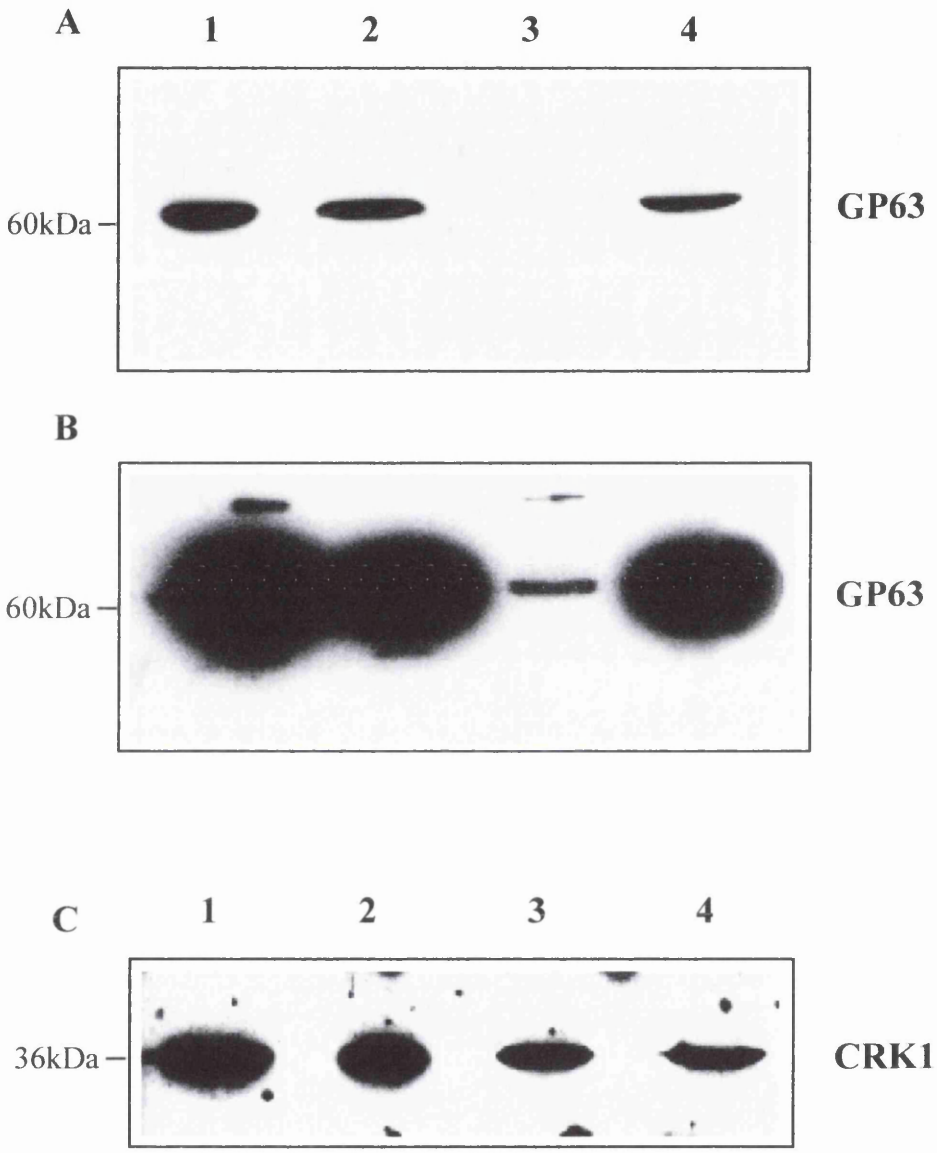

A
1
2
3
4
60kDa
GP63
B
60kDa
GP63
C
1
2
3
4
36kDa
CRK1

Figure 4.14: The effects of inhibition of the proteasome and lysosomal cysteine proteinases on the loss of GP63 in *GPI8* null mutants. **(A)** Wild type, P1S1 and P1S1[pGL269] promastigotes were resuspended at $1x10^7$ cells ml^{-1} in 5ml of fresh medium and were incubated in both the absence and the presence of the proteasome inhibitor MG132 at 10μM. Cells were incubated at 25°C for 24 hours and cell counts were performed. Lysates were prepared and fractionated by SDS-PAGE before electroblotting and detection with anti-*L. major* GP63 antibodies at 1 in 25 dilution. **(B)** Promastigotes were resuspended at $1x10^7$ cells ml^{-1} in 5ml of fresh medium in both the presence of 10mM Z-phe-ala-CHN_2 (Z-FA) and its absence. Cells were incubated for 48 hours at 25°C before harvesting and lysis. Lysates were analysed by Western blotting with anti-*L. major* GP63 as before.

In both cases, samples were electrophoresed as follows:

Lane 1: wild type plus inhibitor

Lane 2: P1S1 plus inhibitor

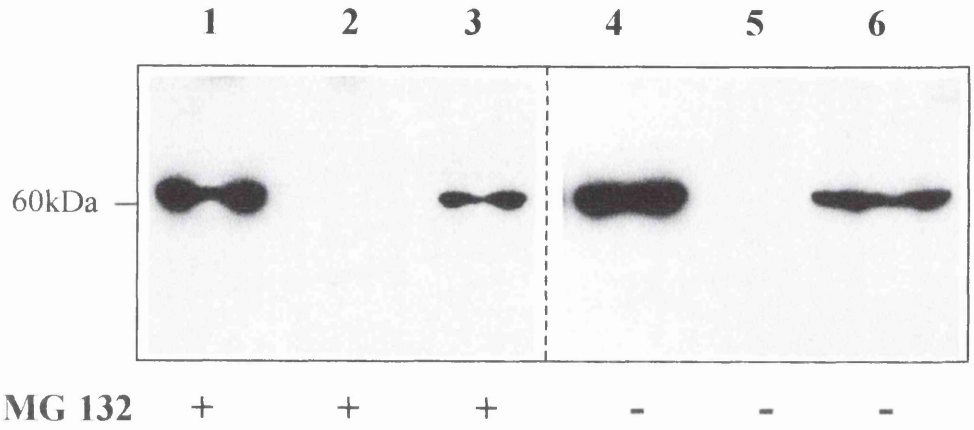
Lane 3: P1S1[pGL269] plus inhibitor

Lane 4: wild type minus inhibitor

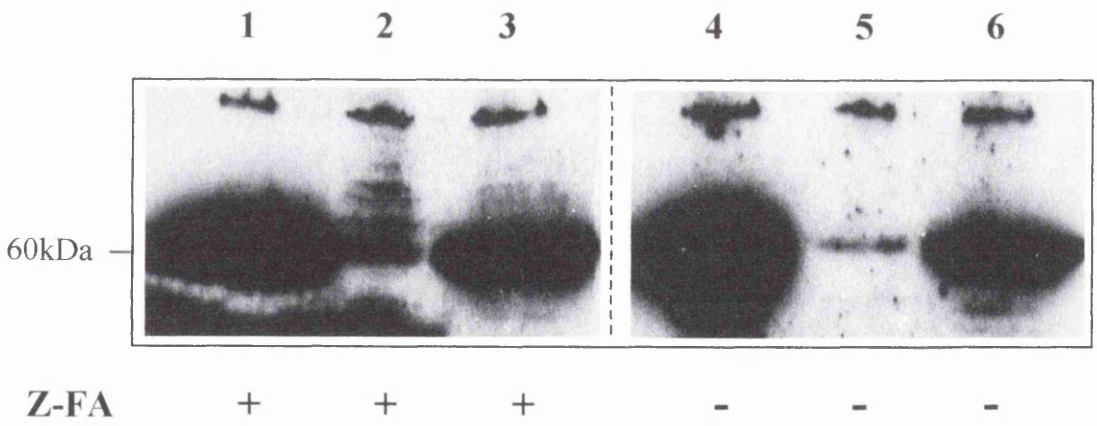
Lane 5: P1S1 minus inhibitor

Lane 6: P1S1[pGL269] minus inhibitor

A



B



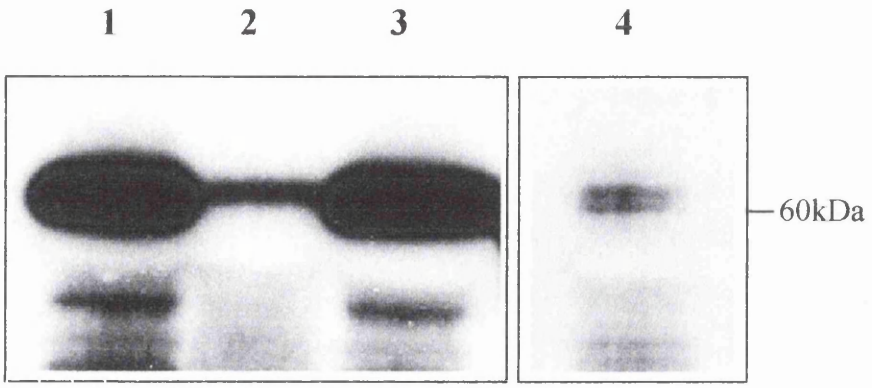


Figure 4.15: Detection of GP63 by Western blotting in cell lysates of P1S1 cells re-expressing a mutated copy of *GPI8*. The cysteine residue at position 216 of LMGPI8 was mutated to glycine and the modified gene was expressed in *GPI8* null mutant (P1S1) cells. Lysates of this cell line were prepared, electrophoresed (SDS-PAGE) and electroblotted, together with wild type, P1S1 and P1S1[pGL269] before being exposed to anti-*L. major* GP63 antibodies at a 1 in 25 dilution.

Lane 1: wild type

Lane 2: P1S1

Lane 3: P1S1[pGL269]

Lane 4: P1S1[pGL403] (C216G mutation)

Molecular weight markers: MultiMark Multi-Colored Standards (Novex)

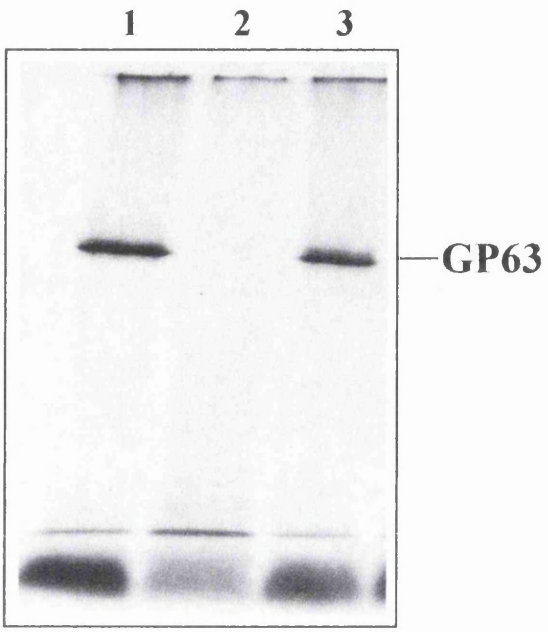


Figure 4.16: [^{3}H] ethanolamine incorporation of GP63. Promastigotes were labelled with [^{3}H]ethanolamine and proteins were fractionated by SDS-PAGE for analysis by fluorography. This work was carried out by Dr. Malcolm McConville and colleagues and is included with his permission.

Lane 1: wild type lysate

Lane 2: P1S1 (*GPI8* null mutant) lysate

Lane 3: P1S1[pGL269] (re-expresser) lysate

CHAPTER 5

PRODUCTION OF RECOMBINANT GPI8 AND FURTHER INVESTIGATIONS OF *L. MEXICANA* GPI8

5.1 INTRODUCTION

GPI8 from yeast has been demonstrated to exist within the lumen of the endoplasmic reticulum (ER) as a type I membrane protein (Benghezal *et al.*, 1996). The yeast protein consists of three domains, as determined by proteinase protection assay: a large ER lumenal domain, the transmembrane region and a short cytoplasmic tail of unknown function. The human homologue also seems to have this overall structure based on the similar hydrophobicity plot to the yeast protein (Benghezal *et al.*, 1996; Yu *et al.*, 1997). By a combination of enzymatic treatment to remove glycans and Western blotting techniques, it was found that yeast Gpi8p is post-translationally modified by the attachment of N-linked glycans at three predicted sites (Benghezal *et al.*, 1996). The human homologue lacks any potential N-glycosylation sites. The *L. mexicana* GPI8 protein, however, contains three potential sites for N-linked glycosylation and part of this section of the thesis was to determine if these are utilised. LMGPI8 (38kDa) is smaller than the *S. cerevisiae* (47.5kDa) and human (45kDa) homologues and it lacks the C-terminal hydrophobic domain which is thought to be involved in attaching these other GPI8 proteins to the lumenal membrane of the ER. Another integral membrane protein may be required to attach LMGPI8 to the ER membrane. Such a protein could be part of the transamidase complex, required for enzymatic activity or could be an accessory protein. One such a protein has already been identified in yeast. Gaa1p is an integral membrane protein

of the ER which appears to be an essential component of the yeast GPI:protein transamidase (Hamburger *et al.*, 1995). This protein has been identified in humans (Hiroi *et al.*, 1998) but has not yet been shown to be present in parasitic protozoa.

The secretory pathway of trypanosomatids, particularly African trypanosomes, has been studied extensively in recent years. The use of fluorescent markers such as green fluorescent protein (GFP) fused to a protein of interest such as the major surface protein of procyclic African trypanosomes (procyclic acidic repetitive protein, or PARP) has allowed the trafficking to the cell surface to be measured in live cells (Bangs, 1998). Eukaryotic proteins that enter the secretory pathway are generally targeted to the ER by means of a hydrophobic, N-terminal signal peptide that is removed upon entry to the ER lumen (Walter *et al.*, 1984). But what governs the retention of ER proteins? There is evidence to suggest that retention of some proteins is facilitated by a C-terminal tetrapeptide; these proteins are still able to exit the ER, however specific receptors in a later compartment in the secretory pathway bind these proteins facilitating return to the ER by retrograde transport (Pelham, 1991; Bangs, 1998). The C-terminal tetrapeptide in LMGPI8 (YDLE) resembles the classical ER-retention motif described in both mammals and yeast (KDEL and HDEL respectively) (Pelham *et al.*, 1988; Munro and Pelham, 1987) and may be significant in keeping soluble GPI8 in the ER, where it can interact with the other membrane-bound components of the transamidase complex. BiP, the ER molecular chaperone protein of *T. brucei* (Bangs *et al.*, 1993) also possesses such a motif (MDDL) and deletion of the BiP tetrapeptide results in secretion of the protein (Bangs *et al.*, 1996).

To begin further analyses of LMGPI8, overexpression of the protein for antibody production was necessary. The most common method for providing large amounts of recombinant protein is to use *E. coli*. *E. coli* often cannot express proteins in their correctly folded, soluble form and produce insoluble inclusion bodies containing the overexpressed protein. For example the cysteine proteinase cathepsin B of mammals (Kuhelj *et al.*, 1995) and one of the cysteine proteinases (CPB2.8) encoded by the *CPB* array of *L. mexicana* (Sanderson *et al.*, 1999) were both expressed in *E. coli* as insoluble inclusion bodies. It may be possible to refold the purified protein artificially during the removal of the denaturants in which the protein was purified, as was done with *L. mexicana* CPB. Modified versions of the protein can be expressed to simplify purification. For example, the inclusion of a 6 x Histidine tag can be used, allowing the protein to be purified on a Nickel column. Ni^{2+} ions coordinate strongly with the histidine residues that act as ligands due to electron donation by the amide components of the amino acid side chains.

It was the aim of this section to:

- Express recombinant GPI8 in *E. coli*
- Produce mono-specific antibodies to GPI8
- Investigate whether LMGPI8 is post-translationally modified by the addition of N-linked glycans.
- Attempt to establish the subcellular location of LMGPI8 by generation of a GPI8-GFP fusion protein with a view to providing further evidence that LMGPI8 is the homologue of yeast Gpi8p and a component of the GPI:protein transamidase.

The fluorescent marker protein, GFP has also been used successfully to determine the subcellular location of particular proteins, by generating and expressing a fusion protein and detecting the GFP fluorescence in particular subcellular compartments (Ha *et al.*, 1996). For example, dolichol phosphate mannose synthase (DPMS), an enzyme associated with GPI-anchor biosynthesis was localised to a novel tubular structure in *L. mexicana*, thought to be a subcompartment of the ER (Ilgoutz *et al.*, 1999b).

5.2 METHODS

5.2.1 Preparation of constructs

5.2.1.1 Generation of pQE-30/GPI8 (pGL184)

The parent plasmid chosen for expression of GPI8 in *E. coli* was pQE-30 (Qiagen), which places a 6 x His tag at the N-terminus of the protein allowing the one-step purification on a Ni-NTA column. It was decided not to include the hydrophobic ER signal sequence of GPI8 as it is probably absent from the mature protein. The position of the ER signal peptidase cleavage site was predicted using a programme called SignalP available via the EXPASY Proteomics tools web site (http://www.expasy.ch/tools/). PCR primers were designed accordingly (Table 5.1), and *GPI8* was amplified and cloned into pQE-30 to produce pGL184 (Figure 5.1).

Table 5.1: Primers used to amplify *GPI8* for cloning into pQE-30

Primer	Sequence (5'-3')	Restriction site
OL437	CGGGATCCGGCAAGGGCCAGAGCAACAAC	***Bam*HI**
OL438	CTAGTCGACCTACTCGAGGTCGTAGCGGACGT	***Sal*I**

Note that the ER signal sequence was omitted from the construct but that the naturally occurring stop codon (shown in bold) was retained. Restriction sites are underlined.

pGL184 was used to transform electrocompetent M15[pRep4], AB1899F` and AB1899F` GroEL/ES bacterial strains by electroporation and these cells were used for protein overexpression.

5.2.1.2 Generation of pXG-*GPI8*-GFP (pGL190)

GPI8 was amplified by PCR using the primers shown in Table 5.2. Primers were designed to engineer a *Bam*HI restriction site on the 3` end of *GPI8*, in place of the stop codon to allow simple cloning into pXG-GFP, such that the GFP component of the fusion protein is located at the C-terminus of GPI8.

Table 5.2: Primers used to amplify *GPI8* for cloning into pXG-GFP

Primer	Sequence (5`-3`)	Restriction site
OL510	TCTAGAAAACGGCAGCTAGTGCA	*Xba*I (natural)
OL511	CGGGATCCCTCGAGGTCGTAGCGGACGTCCA	*Bam*HI

The gene was amplified with VENT polymerase from pGL187 (Section 3.2.1.3) using the same conditions as the pQE-30 cloning (Section 5.2.1.1). Amplified fragments were digested with *Bam*HI directly in the PCR mixture followed by phenol/chloroform extraction, ethanol precipitation and gel-purification. The pXG-GFP vector was digested with *Sma*I and *Bam*HI. The *Sma*I blunt-end was ligated to the 5` end of the amplified fragment and the sticky-ended *Bam*HI site at the 3` of the amplified gene was ligated with the similarly digested end of pXG-GFP yielding pXG-*GPI8*-GFP, designated pGL190 (Figure 5.2).

5.2.2 Expression of recombinant GPI8His

5.2.2.1 Buffers for denaturing purification of GPI8His

Buffer A: 6M guanidine hydrochloride, 0.1M NaH_2PO_4, 0.01M Tris-HCl, pH 8.0. Store at room temperature.

Buffer B: 8M urea, 0.1M NaH_2PO_4, 0.01M Tris-HCl, pH 8.0. Store at room temperature.

Buffer C: Same as buffer B, but pH adjusted to 6.3 with HCl. Store at room temperature.

Buffer D: Same as buffer B, but pH adjusted to 5.9 with HCl. Store at room temperature.

Buffer E: Same as buffer B, but pH adjusted to 4.5 with HCl. Store at room temperature.

Buffer F: 6M guanidine hydrochloride pH7.8, 0.2M acetic acid. Store at room temperature.

5.2.2.2 Small scale purification

500µl of an overnight culture of M15[pRep4] [pGL184] was added to 10mls of fresh L-broth and incubated for about 60-90 minutes at 37°C with shaking until the optical density (600nm) reached 0.7 to 0.9. IPTG (2mM final concentration) was added to induce the expression of GPI8His. IPTG binds to the repressor protein (which is expressed from the pRep4 plasmid), preventing it from binding to pGL184 and

expression occurs. 1ml samples were taken at 0, 1, 2, 3 and 4 hours after the addition of IPTG. Cells were pelleted and lysed in 150µl of the denaturing Buffer A for 1 hour at RT with agitation. Ni-NTA beads were added to bind any GPI8His and after 30 minutes, beads were pelleted, washed with 1ml of Buffer B and 3 x 1ml of Buffer C before elution in 50µl Buffer C + 100mM EDTA.

5.2.2.3 Large scale purification

A 20ml overnight culture of M15[pRep4][pGL184] was added to 1 litre of fresh L-broth with 100µg/ml ampicillin and 25µg/ml kanamycin. After about 3 hours the OD_{600} was about 0.7 and IPTG was added to 2mM final concentration. Cells were incubated for 3 hours at 37°C before being harvested in 4 x 250ml aliquots. Pellets were stored at –70°C until required.

One pellet was lysed in 5ml of Buffer A and sonicated (Microtip 20% amplitude with 5 x 5 second pulses). The sample was centrifuged at 10,000 x g for 15 minutes to remove any unbroken cells and the supernatant was passed through a 0.4µM filter. A 4ml Ni-NTA column was prepared according to manufacturers instructions and equilibriated with Buffer A and the sample loaded onto the column using a Waters 625LC FPLC system (Millipore). Absorbance at 280nm was monitored to detect proteins in the flow through. Fractions were collected throughout the washing and elution of the column. The column was washed with Buffers B and C until no more protein was eluted. Buffers D, E or F would be expected to contain GPI8His. In fact the recorded elution profile indicated that the majority of protein was eluted from the column in Buffer F.

5.2.2.4 Inclusion body purification

A second frozen pellet derived from the large-scale expression of GPI8His (previous section) was lysed and inclusion bodies were purified using a modified version of an existing method (Kuhelj *et al.*, 1995). The cell pellet was resuspended in 3ml 50mM Tris-HCl, pH8.0 containing 5mM EDTA, 5% sucrose and 1mg/ml lysozyme. The suspension was incubated for 20 minutes at 37°C for lysis then the volume was made up to 10ml in same buffer without lysoszyme. The lysate was subjected to two cycles of freezing (in dry ice/ethanol) and thawing (incubation at room temperature) then sonication: 6 x 10s, 200W to fragment nucleic acids. The lysate was centrifuged, (6000 x g, 10mins) to pellet the inclusion bodies and these were washed once in 50mM Tris/HCl, pH8.0, 5mM EDTA, 0.1% Triton X-100 and twice in 50mM Tris/HCl, pH8.0, 5mM EDTA, 2mM urea. The inclusion bodies were washed in water and centrifuged and were solubilised in Buffer B + 10mM DTT (37°C, 1 hour, then 4°C, overnight).

5.2.3 Co-expression of GPI8His with the chaperonins GroEL and GroES

4.5ml of overnight cultures of AB1899F`[pGL184] and the AB1899F`GroEL/ES [pGL184] clones were used to inoculate 2 x 100ml of L-broth each. The flasks were shaken at 37°C for 2 hours (OD_{600} = 0.7). IPTG was added to a final concentration of 500μM. Note that this was less IPTG than was used with M15[pRep4] [pGL184] cells since expression of the chaperonin was also induced by IPTG, presumably allowing sufficient quantities of GroEL/ES production before GPI8His levels get too high to form inclusion bodies. The flasks were incubated as follows:

Flask 1 – AB1899F' [pGL184] – shaken at 37°C for 3 hours

Flask 2 – AB1899F' [pGL184] – shaken at 19°C for 24 hours

Flask 3 – GroEL/ES [pGL184] – shaken at 37°C for 3 hours

Flask 4 – GroEL/ES [pGL184] – shaken at 19°C for 24 hours

1ml samples were taken from all flasks at t=0 hours after addition of IPTG (cells were pelleted and frozen (-20°C) until samples from all four flasks had been collected). 1ml samples were taken from flasks 1 and 3, three hours after IPTG addition – the remainder of each culture was harvested and stored at –70°C until required. 1ml samples were taken from flasks 2 and 4, 24 hours after induction and the cells were pelleted (the rest of each culture was also harvested and stored at -70°C).

The remainder of the 100ml GroEL/ES culture was lysed in 5ml of LS-T buffer. The lysate was passed through a 0.4μM filter and loaded onto the column with LS buffer, which was passed through the column until no more Triton X-100 or protein was detected in the flow through (note that Triton X-100 absorbs UV light at 280nm, masking the protein signal meaning that the elution profile could not be used as a means of detecting protein). The column was washed with high salt buffer (HS) and proteins were eluted in an imidazole gradient of 0-250mM. Finally 500mM imidazole was passed through to remove all traces of protein. Fractions were collected throughout the purification and fractions 2-5, 8, 11 and 14 taken during the imidazole gradient were concentrated by ultrafiltration using Centricon-10 tubes, which have a 10kDa molecular mass cut off.

5.2.4 Antibody generation and affinity purification

Purified recombinant GPI8His was dialysed against PBS for two days with several changes of PBS. The protein, which had precipitated was diluted to 1mg/ml as estimated by spectrophotometry by comparison with known concentrations of bovine serum albumin (BSA). This material was sent as an insoluble suspension to the Scottish Antibody Production Unit (SAPU) for inoculation into rabbits.

The amino acid sequence of LMGPI8 was commercially analysed by Genosys for identification of potential antigenic regions. A 21 amino acid GPI8 peptide (NFNYHPYRVEASRNRSKPAHRD) was synthesised and used to inoculate two rabbits. This resulted in anti-peptide GPI8 antibodies (R771 and R455)

GPI8 peptide was immobilised using the AminoLink Immobilization Kit (Pierce) according to manufacturer's protocol. 1ml of R771 (anti-peptide GPI8) antisera was added directly to the immobilised epitope column and sample was mixed with the column resin/epitope for 1 hour at room temperature. Column was washed with 16ml PBS and bound proteins were eluted with IgG Elution Buffer (Pierce) and collected in 1ml fractions, which were analysed for protein by spectrophotometry (absorbance readings were taken at 280nm wavelength). Fractions containing protein were pooled and dialysed against several changes of PBS before storage at 4°C.

5.2.5 Periodate treatment

The method of Bayne and co-workers was used to periodate-treat Hybond C-blotted lysates (Bayne *et al.*, 1993). Two blots were prepared and one was incubated with 10mM sodium metaperiodate (Sigma) whilst the other was incubated with periodate buffer without periodate for 60 minutes. Both blots were washed in wash buffer

containing glycerol followed by three rinses in the same buffer minus glycerol. Membranes were then blocked and exposed to antibodies as normal.

5.2.6 *Trypanosoma brucei* immunofluorescence with R491

Procyclic *T. b. rhodesiense* were fixed on slides and exposed to two primary antibodies simultaneously using the method of McDowell *et al.*, (1998). One primary antibody was a mouse monoclonal raised against the TY epitope tag (Bastin *et al.*, 1996) present in QM protein and the other was R491 (rabbit anti-GPI8His). Both were used simultaneously at 1 in 100 dilution. Two secondary antibodies were also used simultaneously, anti-mouse IgG-TRITC conjugate and anti-rabbit IgG-FITC conjugate at 1 in 100 dilution, together with DAPI at 20μg/ml. Cells were viewed by immunofluorescence microscopy (a Zeiss Axioplan fluorescence microscope with a Hamamatsu Digital Camera using the OpenLab Software Package (Improvision, University of Warwick)). FITC fluorescence due to QM was observed using Zeiss filter set 09 whilst TRITC was observed using Zeiss filter set 15.

5.3 RESULTS

5.3.1 Expression of recombinant GPI8His in *E. coli*

5.3.1.1 Small-scale expression of GPI8His

To determine the time course of GPI8His expression, a small-scale expression experiment was carried out and hourly time points were taken. These samples were analysed by SDS-PAGE and the protein detected by staining with Coomassie Blue (Figure 5.3). Significant quantities of GPI8His were detected at all timepoints. Note that the purification of GPI8His was carried out under strongly denaturing conditions (6M guanidine hydrochloride and 8M Urea).

To determine if any soluble GPI8His was expressed, the induction was repeated and the purification was repeated under soluble conditions but no purified protein was detected (not shown).

5.3.1.2 Large scale expression and purification of GPI8His

Larger cultures expressing GPI8His were harvested, lysed and purified on a Ni-NTA column. Fractions that coincided with the protein peaks of the elution profile were analysed by SDS-PAGE. Note that the proteins in Buffers A and F were TCA precipitated and resuspended in the same volume of SDS-PAGE loading buffer prior to electrophoresis because 6M guanidine hydrochloride forms a precipitate.

The gel was stained with Coomassie Blue and is shown in Figure 5.4A. A protein of 36-40kDa (approximately the size expected for GPI8His) was detected in Buffer F elution (Figure 5.4A – lane 8) that was not present in the other fractions. Despite this, there were number of bands of higher mobility, which were presumably contaminating *E. coli* proteins or partially degraded GPI8His.

Because GPI8His is insoluble, it forms inclusion bodies in *E. coli*. This was exploited as a simple extra purification step, given that inclusion bodies could be purified directly from *E. coli* using a revised version of the method of Kuhelj *et al.*, (1991) (see Section 5.2.2.4). The solubilised inclusion bodies were passed through a 0.4μM filter and loaded and washed on the column. As before very little protein was detected in the elution profile of Buffers D and E (not shown). Most of the eluted protein was collected in a single Buffer F fraction and this was analysed by SDS-

PAGE (Figure 5.4B, lane 3). Whilst bands other than GPI8His were visible, the sample was much cleaner than before and was used to inoculate rabbits.

5.3.1.3 Generation of antisera against GPI8His

The concentration of GPI8His was estimated by absorbance at 280nm to be approximately 0.75mg/ml, using known concentrations of BSA to produce a standard curve of A_{280} readings. Dialysis of GPI8His to remove Buffer F was attempted to re-fold the protein, however it was found to consistently precipitate out of solution. Even gradually decreasing the amount of the denaturant over time was not effective in re-folding the protein and resulted in precipitation. 6ml (approximately 4.5mg) of GPI8His in Buffer F was dialysed against PBS (with several changes of PBS) to remove as much denaturant as possible. The precipitate was collected by centrifugation (17000xg, 30 minutes) and was resuspended at approximately 1mg/ml in PBS. 8x400µl aliquots were sent to the Scottish Antibody Production Unit (SAPU) to be injected into two rabbits for antibody production (R491 and R492). In addition, a synthetic peptide (peptide sequence) of part of GPI8 was produced and used to produce antibodies in rabbits by Genosys (R771 and R455).

These antibodies were tested by Western blotting with purified recombinant GPI8His. All were found to bind readily to the recombinant protein (not shown).

5.3.1.4 Soluble expression of GPI8His in *E. coli* expressing the chaperonins GroEL/GroES

An *E. coli* strain (AB1899F') containing plasmid which expressed the chaperonins GroEL/GroES on induction with IPTG (AB1899F' GroEL/ES) were a gift from Dr.

John Hyde of UMIST. The chaperonins GroEL and GroES are proteins that provide a vessel for the re-folding of non-native polypeptides (see (Ranson *et al.*, 1998) for review). Co-expression of GPI8His with these chaperonins could provide folded and soluble GPI8. Therefore, the *E. coli* strain AB1899F` and AB1899F`GroEL/ES and were transformed with pGL184.

The four small scale GPI8His expression samples described in Section 5.2.2 were harvested and lysed in 50µl sonication buffer + 1mg/ml lysozyme for 15 minutes at 37oC followed by two freeze/thaw cycles. Lysates were mixed with SDS-PAGE sample buffer, boiled for 5 minutes and analysed by SDS-PAGE and Western blotting using affinity purified R771 peptide antibodies (Figure 5.5). At t=0, GPI8His could not be detected in either AB1899F`[pGL184] (Figure 5.5A – lanes 1 and 2) or AB1899F`GroEL/ES[pGL184] (Figure 5.5B – lanes 1 and 2) cells. GPI8His was detected in abundance in the insoluble fractions of both cell lines incubated at 37°C during expression (Figure 5.5A – lane 4; Figure 5.5B – lane 4) and a very small amount was detectable only in the soluble fraction of the GroEL/ES cells. Interestingly, when GroEL/ES cells were incubated at 19°C during expression, a much larger amount of GPI8His was present in the soluble fraction (Figure5.4B – lane 5) although a band of about the same intensity was detected in the insoluble fraction (lane 6). No protein was detected in the soluble fraction of AB1899F` incubated at 19°C (Figure 5.5– lane 5).

5.3.1.5 Large-scale purification of soluble GPI8His

2ml fractions collected from the Ni-NTA column were concentrated by ultrafiltration and analysed by SDS-PAGE (Coomassie-stained gel) and Western blotting with R771

antibodies (Figure 5.6). From the Coomassie gel (Figure 5.6A) Fraction 11 (lane 6) was found to contain several proteins. As the major band was approximately 60kDa it is likely to correspond to GroEL. There were a few higher mobility proteins, notably one of approximately 36-38kDa, which is the size predicted for GPI8His. Western analysis (Figure 5.6B) identified a single protein of 42kDa (5.6B – lane 6), slightly larger than expected. The more abundant chaperonin at 60kDa was not recognised by this antibody. Because the band of interest was absent in fractions 8 (Figure 5.6B – lane 5) and 14 (Figure 5.6A – lane 7), fractions 9, 10, 12 and 13 were concentrated by ultrafiltration and analysed by Western blotting (Figure 5.7). The 42kDa band could be detected in fractions 10, 12 and, to a lesser extent, 13 (Figure 5.7 – lanes 2, 3 and 4 respectively). There was no detectable protein in fraction 9 (lane 1). The reason for the lower than expected mobility of the putative GPI8His is not clear. To confirm that the antibodies are cross-reacting with GPI8His the protein could be separated by SDS-PAGE, purified and subjected to N-terminal sequencing to confirm its identity.

5.3.2 Western analysis of *L. mexicana* lysates with anti-GPI8 antibodies

Repeated efforts to detect GPI8 in *L. mexicana* whole cell lysates by Western blotting with R491 and R492 (anti-GPI8His) and R771 and R455 (anti-peptide) antibodies have failed. A variety of modifications of the standard technique were used, including:

- Different amounts of material were loaded in each lane, ranging from 10^6 to $2x10^7$ cell-equivalents.

- Primary antibodies were used at different concentrations. R491 and R492 anti-GPI8His were used at between 1 in 200 to 1 in 1500 dilutions of crude serum. The GPI8 peptide was immobilised on Aminolink resin and this was used to affinity purify R771 antibodies (Section 5.2.4). Affinity purified R771 anti-peptide GPI8 antibody was used at between 1in 20 and 1 in 200 dilution. R771 crude serum was used at between 1 in 500 and 1 in 1500 dilution.
- Three different blocking solutions (1, 2 and 3) were used with the different antibody concentration described above (described in Chapter 2).

All of these antibodies detected recombinant GPI8His (not shown). The affinity-purified anti-peptide GPI8 antibodies were the best as they detected relatively low levels of soluble recombinant protein expressed in and purified from *E. coli* (Figure 5.7).

GPI8 should be present in both wt and the re-expresser cell lines, but not in P1S1. Thus the *GPI8* null mutant serves as a negative control in Western analysis. Analysis of lysates of these cell lines did detect proteins, but these were in the region of 60kDa and were also present in *GPI8* null mutants indicating that these proteins were not GPI8. It could be that GPI8 is present in such small amounts that this method of detection is not sensitive enough to detect the protein. To address this, the Biotin/Avidin Labelling Kit (Santa Cruz ABC Reagents) was used in conjunction with Western Blotting according to manufacturer's instructions. This kit provides amplification of antibody binding by the use of a biotinylated secondary antibody (anti-rabbit IgG). Avidin possesses four binding sites for biotin, one of which is filled by the secondary antibody. Biotinylated-HRP when added binds to the remaining

sites on Avidin and produces a much stronger signal with ECL. This was carried out with four separate blots, each of which had lysates from wt, *GPI8* null mutant P1S1 and the re-expresser cell line P1S1[pGL269]. The blots were probed with both anti-GPI8His antibodies (R491 and R492) and neat and affinity purified anti-peptide GPI8 antibody R771. A large number of bands were detected, however none of these were absent in the null mutant but present in both wt and P1S1[pGL269] (Figure 5.8B).

5.3.2.1 Enrichment of ER proteins by purification of microsomes

In order to enrich the ER proteins in an attempt to maximise the chances of detecting GPI8, microsomes were purified from large, *L. mexicana* stationary phase cultures.

Microsomes were prepared from 10^9 wt, P1S1 and P1S1[pGL269] promastigotes by lysing the cells hypotonically in 1ml of Microsome buffer 1 (MB1) and membranes were stabilised by the addition of an equal volume of Microsome buffer 2 (MB2). Unbroken cells and larger organelles (e.g. nuclei) were pelleted by centrifugation (2,100xg, 10 minutes) and the supernatant was further fractionated into medium sized organelles (e.g. kinetoplasts) by centrifugation (15,800xg, 10 minutes). Microsomes were pelleted from the remaining supernatant by ultracentrifugation (240,000xg, 60 minutes). The microsome pellets were re-suspended in 60µl of 0.25M sucrose and analysed by SDS-PAGE Coomassie-stained gels (Figure 5.8B) and Western blotting. Despite the presence of microsome proteins (Figure 5.8B), Western blotting failed to detect GPI8 (not shown). Further experiments to detect ER marker proteins such as BiP, to which antibodies were available (Bangs *et al.*, 1993; Bangs *et al.*, 1996) would be required to confirm the presence of ER markers in this preparation.

5.3.2.2 Inhibition of N-glycan attachment and further attempts to detect GPI8

Tunicamycin is a widely used antibiotic that blocks the attachment of N-glycans by inhibiting glycosyltransferases in the ER and Golgi (Kink and Chang, 1987). This inhibitor has been used successfully with *Leishmania* to block glycosylation of GP63 (Funk *et al.*, 1994). If the failure to detect GPI8 is due to extensive glycosylation, blockage of N-glycan may allow antibodies to bind allowing detection.

Wt promastigotes were grown in the presence of 5μg/ml tunicamycin (Sigma) for 100 hours using the method of Funk and co-workers (1994). After this time, the cells were harvested and analysed by SDS-PAGE and Western Blotting using affinity purified anti-peptide GPI8 antibodies. No GPI8 could be detected.

5.3.2.3 Periodate treatment of glycoproteins prior to Western blotting

As an alternative to blockage of N-glycan attachment, treatment with sodium metaperiodate could be used. Periodate-treatment brings about the oxidation of carbohydrates, causing glycan destruction perhaps making the protein component of the glycoprotein accessible to antibodies.

Lysates of *L. mexicana* wt cells were subjected to SDS-PAGE and electroblotting onto Hybond C. One blot was treated with periodate and the other wasn't. Both blots were then blocked and exposed to the affinity purified anti-peptide GPI8 primary antibody and anti-rabbit IgG-HRP secondary antibody. GPI8 was not detected.

5.3.2.4 GPI8His antibodies can detect a protein of the expected size in *T. brucei* procyclic and bloodstream-form cell lysates

L. mexicana GPI8 was found to share a high degree of homology (>70%) with the *T. brucei* homologue over the region characterised in Chapter 3. Antibodies raised against LMGPI8 could potentially cross-react with TBGPI8. Pellets containing approximately 10^8 procyclic (strain EATRO795) and 10^8 bloodstream-form (laboratory cultured 427 strain) African trypanosomes were supplied by Dr. Jaap van Hellemond and Dr. Patricia Blundell, respectively. Lysates were analysed by SDS-PAGE and Western blotting (Figure 5.9). A band of approximately 38-40kDa was detected in both procyclic (Figure 5.9 – lane 1) and bloodstream form (lane 2) lysates. There were a number of slightly higher mobility bands detectable in procyclics and these could be unglycosylated or partially glycosylated forms of TBGPI8.

5.3.2.5 Immunofluorescence of *T. brucei rhodesiense* procyclic cells with R491 antibodies

This work was carried out with Mr. Simon Lillico (University of Glasgow) and is included with his permission. A gene for a *T. brucei* protein called QM, which is a potential homologue of the human and yeast ribosomal QM (Dowdy *et al.*, 1991; Koller *et al.*, 1996), was identified and isolated. The TY epitope tag (Bastin *et al.*, 1996) was inserted into the QM gene and expressed in procyclic trypanosomes. The aim of this experiment was to investigate the subcellular location of the 38-40kDa band, detected with anti-GPI8His antibodies in the previous section, using immunofluorescence.

Fluorescence due to TY-tagged QM was compared to either putative TBGPI8 fluorescence or BiP (Figure 5.10). QM (Figure 5.10, Panel B) produced a very scattered pattern with a clear nuclear envelope signal, which is indicative of an ER-associated protein (it is envisaged that many ribosomes will be associated with the cytoplasmic face of the ER due to the high level of expression of secretory pathway/GPI-anchored proteins). Panel C shows the fluorescence with anti-GPI8His antibodies. These patterns were strikingly similar and they overlapped almost perfectly (Panel D). However, the ER marker protein, BiP (Panel G), used by Bangs and co-workers (Bangs *et al.*, 1993; Bangs *et al.*, 1996) produced some overlap with QM but the signals were clearly different (Panel H). Anti-BiP antibodies were a gift of Dr. Jay Bangs of the University of Wisconsin.

5.3.3 The use of green fluorescent protein to investigate the subcellular localisation of LMGPI8

Green fluorescent protein (GFP) has been used successfully with *Leishmania* to localise LPG1, one of the proteins involved in LPG biosynthesis (Ha *et al.*, 1996). It was found that if the gene was fused to *GFP* and expressed in LPG1-mutant cells, complementation of the mutant phenotype and restoration LPG surface expression was achieved. Fluorescence due to GFP can be visualised by microscopy in the same way as FITC immunofluorescence staining visualisation. The aim of this section was to clone GPI8 into the pXG-GFP vector (Ha *et al.*, 1996), a gift of Dr. Steve Beverley (Washington University Medical School), to produce a GPI8-GFP fusion protein. Fusion of GFP to the C-terminus of GPI8 would allow the naturally occurring ER-signal sequence of GPI8 to be utilised in trafficking of the fusion

protein. When re-introduced into *L. mexicana GPI8* null mutants, the GPI8-GFP fusion could be tested for the ability to function as a transamidase by analysis of GP63 processing and subsequently attempts to detect GFP fluorescence in fixed cells could be made.

5.3.3.1 Transfection of P1S1 null mutants with pXG-*GPI8*-GFP (pGL190)

P1S1 promastigotes were transfected with 15μg of pGL190. Transfected cells were grown on agar plates with 25μg/ml each of nourseothricin and G418 and 10μg/ml puromycin. Antibiotics were sufficient to prevent growth of the control transfection cells. Four triple-resistant clones (GFP1, GFP2, GFP3 and GFP4) were obtained and these were transferred to liquid medium and grown up for analysis.

5.3.3.2 Western blot analysis of GP63 in P1S1[pGL190] cells

The first question to be addressed was whether the GPI8-GFP fusion could restore GPI:protein transamidase function, as seen with the wild type gene.

As shown previously, the presence of surface GP63 is an excellent marker for the functionality of the fusion protein in terms of its transamidation capability. Two of these clones (GFP2 and GFP3) together with wt, P1S1 and re-expresser cell lines, were analysed by SDS-PAGE and electroblotting using anti-GP63 monoclonal antibody. The blot is shown in Figure 5.11. As found previously (Figure 4.13) GP63 was detected in both wt and the re-expresser cell line but was absent in *GPI8* null mutant cells (Figure 5.11 – lanes 1, 3 and 2 respectively). In the GFP2 and GFP3 lysates several higher mobility proteins (approximately 55-58kDa) were detected (5.11 – lanes 4 and 5). These proteins appeared to be more abundant than the

comparable proteins in wt and re-expresser lysates (lanes 1 and 3 marked by a yellow arrow).

This indicates that GPI8-GFP is partially rescuing the mutant phenotype as some GP63 is detected, however, the fusion protein appears to be insufficient to allow the complete processing of the surface protein. This remains to be tested by immunofluorescence with anti-GP63.

5.4 DISCUSSION

It was the aim of this section of work to express GPI8 in *E. coli* to produce sufficient quantities for antibody production in rabbits. This would allow further characterisation of GPI8 in terms of post-translational modifications such as N-linked glycosylation since GPI8 has three sites for potential N-glycan addition. Following this, experiments to elucidate the subcellular localisation of LMGPI8 were begun.

Recombinant GPI8His was expressed in *E. coli* in insoluble inclusion bodies. Purification by Ni-NTA columns was carried out under denaturing conditions leading to an insoluble preparation of GPI8His that was not completely pure - two other proteins with higher mobility on SDS-PAGE were present. These could just be partially degraded GPI8His but this was not investigated further. Because recombinant GPI8His was expressed as insoluble inclusion bodies and purification of inclusion bodies can result in very pure preparations of recombinant protein (Kuhelj *et al.*, 1995), this step was included as an extra purification step. Ni-NTA purification was used as the second step in purification and resulted in a greater purity preparation of recombinant GPI8His. All but one of the higher mobility proteins detectable in the

first purification attempt were absent. Attempts to refold the protein by the gradual removal of denaturant by dialysis failed and the protein was eventually sent to SAPU (for antibody production) in the form of an insoluble suspension of GPI8His in PBS. Anti-peptide GPI8 antibodies were also raised by the inoculation of rabbits with a synthetic peptide of a fragment of GPI8. Both sets of antibodies were found to detect purified recombinant protein, however it proved difficult to detect GPI8 in *Leishmania* lysates.

It would be desirable to be able to produce milligram quantities of recombinant GPI8 for structural studies. Because the protein is expressed as insoluble protein, it would be necessary to develop a method of refolding the protein. Alternatively it has proved possible to produce low levels of soluble GPI8His by co-expression with the chaperonins GroEL and GroES. The technique requires refinement to maximise yields of GPI8His. Furthermore, a single step purification on Ni-NTA columns is insufficient to remove the chaperonin GroEL, which is the most abundant protein in the fractions eluted from the column. Further purification steps, for example anion exchange, may allow purification of GPI8 to homogeneity if the recombinant protein is capable of remaining soluble without association with the chaperonin. The purified GPI8 could then be used for studies of the activity of the transamidase. A cell-free assay system has recently been developed for *T. brucei* GPI-anchor addition involving the use of purified microsomes to study the release of radiolabelled VSG from membranes with the nucleophile hydrazine replacing the role of the GPI-anchor (Sharma *et al.*, 1999a). These membranes could be subject to alkaline wash to remove soluble components of the ER lumen, as was done previously with yeast (Vidugiriene and Menon, 1995) and the soluble protein, assuming it has folded

correctly, could be added back to this system. This, of course, assumes that the *T. brucei* homologue of GPI8 is not membrane bound, like the human and yeast homologues.

Detection of GPI8 in *L. mexicana* cell lysates with these antibodies has not proved possible to date. Enrichment of ER proteins by microsome purification did not alter this, although the "microsome" sample was not checked for the presence of ER-marker proteins such as BiP, to confirm the nature of the purified fraction. It is possible that all of the potential GPI8 glycosylation sites have been used and that the protein is extensively glycosylated. As a result, the epitopes recognised by the antibodies available may be blocked by glycans. The peptide used to produce R771 anti-peptide antibodies, which was designed against the predicted sequence derived from the gene, contains one of the potential N-glycosylation sites. Attempts to inhibit N-glycosylation, using the antibiotic tunicamycin, were made. This antibiotic has already been used successfully with *Leishmania* in the inhibition of GP63 glycosylation (Funk *et al.*, 1994), however it did not appear to have any effect on allowing the detection of GPI8. It could be that the concentration of tunicamycin used was too low and on this occasion, no positive controls were used in the experiment. When the experiment is repeated, tunicamycin will be used at different concentrations and GP63 glycosylation will be assessed as a positive control for the blockage of glycosylation.

Treatment of blotted lysates with sodium metaperiodate to oxidise carbohydrates also had no effect in allowing the detection of GPI8. A further approach to remove glycans attached to GPI8 is the enzymatic removal of N-glycans with enzymes such as

petide N-glycosidase A, or endoglycosidase F must be performed. Only then, if GPI8 remains undetectable will it be certain that attached glycans do not block the epitopes recognised by GPI8 antibodies. Episomal expression of modified *GPI8* could be attempted. Epitope tags such as the TY tag, which are oligonucleotide sequences encoding a short stretch of 10 amino acids to which monoclonal antibodies are available, could be inserted into the GPI8 ORF and expressed in *Leishmania*. The GPI:protein transamidase activity of the GPI8-TY protein could be assessed by detection of GP63 using Western Blotting and/or immunofluorescence. The detection of GPI8-TY with anti-TY monoclonal antibodies could allow subsequent analysis of its post-translational modifications. Clearly the GPI8 gene must be expressed (it cannot be a pseudogene) because deletion of the gene causes phenotypes associated with a defective GPI:protein transamidase and these are reversed by the episomal re-expression of the gene (see Chapter 4). Both sets of antibodies recognise the recombinant GPI8His therefore it should be possible to detect the native protein. Further studies are required to pursue this.

The anti-leishmanial GPI8 antibody detects a 38-40kDa protein in both procyclic and bloodstream form lysates of *T. brucei*. This is about the size expected for GPI8 if it was not heavily glycosylated; yeast Gpi8p is a 44kDa protein and 46, 48 and 50kDa forms are detected by Western analysis (Benghezal *et al.*, 1996). Treatment of yeast Gpi8p with endoglycosidase H results in detection of just the 44kDa protein indicating the protein is N-glycosylated. Immunofluorescence of *T. brucei rhodesiense* procyclic cells detects a pattern reminiscent of ER and which almost perfectly matches that of a ribosomal protein called QM. We would expect the greatest concentration of ribosomes to be on the rough endoplasmic reticulum. Since

these are the sites of protein synthesis and entry of nascent polypeptides into the ER, we might predict that the transamidase will be closely associated with these regions, but on the other side of the ER membrane. It is strange that the antibodies seem to recognise trypanosome but not leishmanial GPI8. Again, this may be due to differences in glycosylation between the species. Another possibility could be that another antibody found in R471 serum could be cross-reacting with a ribosomal protein because the immunofluorescence pattern differs significantly from that of BiP, a widely used ER marker protein. When the expression of soluble GPI8His has been refined, the GPI8 antibodies could be affinity purified by the immobilisation of GPI8His onto Aminolink resin (Pierce), as was done with the anti-peptide GPI8 antibodies. The affinity-purified antibodies (if they still detect the 38-40kDa band) could then be incubated with an excess of recombinant GPI8 before Western blotting as a control. If the putative TBGPI8 protein is absent then this would be good evidence that the antibodies are detecting the actual *T. brucei* homologue. Attempts are underway to determine if the affinity-purified R771 (peptide) antibodies are able to detect proteins in *T. brucei* lysates. If they do, then this experiment can be carried out using these antibodies.

In an attempt to determine the subcellular location of LMGPI8, the gene was cloned into pXG-GFP (Ha. *et al.*, 1996) such that the expressed protein would be a fusion with GFP fused to the C-terminus of GPI8 allowing detection in the cell by UV fluorescence microscopy. Fusion of GFP has been used successfully by Ilgoutz and co-workers (1999b) to demonstrate the subcellular location of DPMS to a subcompartment of the ER. It was found that, when this protein was expressed in *GPI8* null mutant cell line P1S1, rescue of the loss of GP63 was partially restored.

The mature 63kDa protein corresponding to GPI-anchored GP63 was not present in GFP2 and GFP3 cell lines. A higher mobility protein (56-58kDa) that was absent in P1S1 lysates but was barely detectable in both wt and re-expresser lysates, was recognised by anti-GP63 antibodies in pXG-GPI8-GFP cell lines. Further studies are required to determine what this protein could be. One possibility is that GPI8-GFP has partial activity and has processed the C-terminal GPI signal sequence, but has not added a GPI-anchor. This would affect its forward trafficking – the protein could remain within the ER and possibly associated with the GPI:protein transamidase. If trafficking is disrupted then glycosylation may be blocked resulting in a lower molecular weight form. As a comparison, an *L. major* GP63 (accession CAA68673) would have a molecular weight of approximately 57.5kDa after removal of both the ER and the GPI-anchor signal sequences (prior to glycosylation or GPI-anchor addition (Voth *et al.*, 1998)) as estimated from the primary structure using the protein analysis tools at the Expasy Proteomics Website (http://www.expasy.ch/tools). This size would not increase drastically after anchor addition since the *L. major* GP63 anchor is simple, having no branching side chains (Schneider *et al.*, 1990). PI-PLC treatment and Triton X-114 extraction could determine if the GP63 detected in the GPI8-GFP expressing cell lines has a GPI-anchor attached. If the protein is found in the aqueous phase after enzyme treatment but in the detergent phase without enzyme treatment then it is clear that it has been GPI-anchored (Taguchi *et al.*, 1985; Koch *et al.*, 1986). Alternatively, detection of [^{3}H]ethanolamine incorporation of the high mobility GP63 species would indicate the presence of an anchor. Whatever the case, GPI8-GFP is incapable of completely rescuing the mutant phenotype. This means that this fusion protein cannot be used reliably to determine the subcellular location. GFP is a large protein and may be interfering with the binding of other components of

the transamidase complex (e.g. GAA1). Because the putative ER tetrapeptide at the carboxyl terminus of GPI8 is fused to GFP, retention of the fusion protein in the ER may be blocked. The generation of other GPI8-GFP fusion proteins such as GFP at the N-terminus of GPI8, carboxyl to the ER signal sequence may prove to be active. Until effective GPI8 antibodies are available, the epitope-tagged GPI8 could be expressed in the null mutant and detection of the protein by immunofluorescence, using antibodies that recognise the tag, can be used to determine localisation.

These experiments represent the early stages of the analysis of GPI8, in terms of its structure, location and activity. There are many questions still to be answered and these form the basis of research that will take place in the near future.

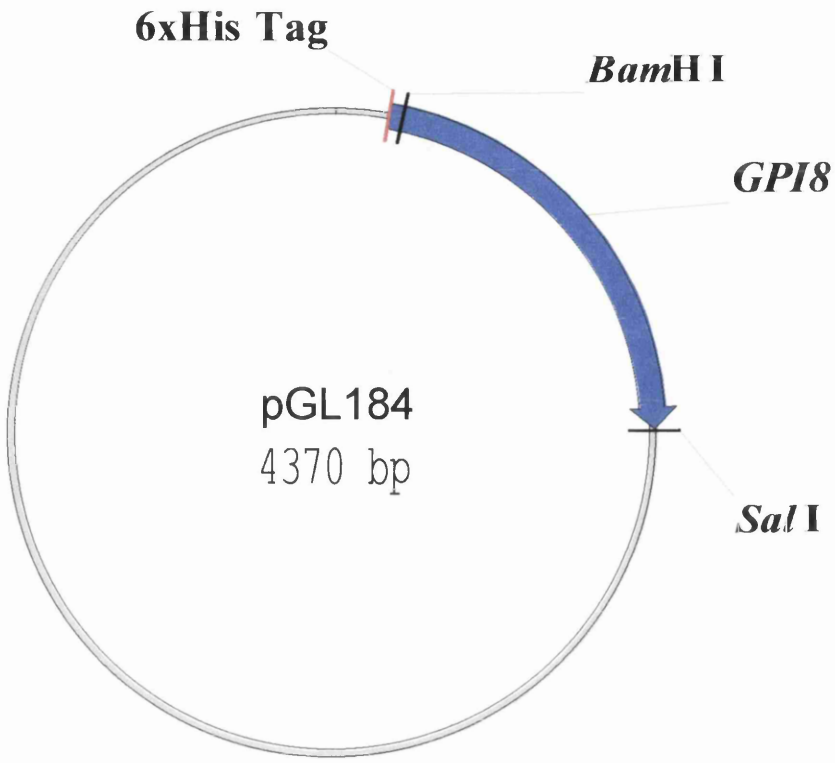


Figure 5.1: Construct for expression of recombinant GPI8 with a 6xHistidine tag. The gene was amplified by PCR using primers with engineered *Bam*HI and *Sal*I sites on the 5` and 3` ends respectively with VENT polymerase.

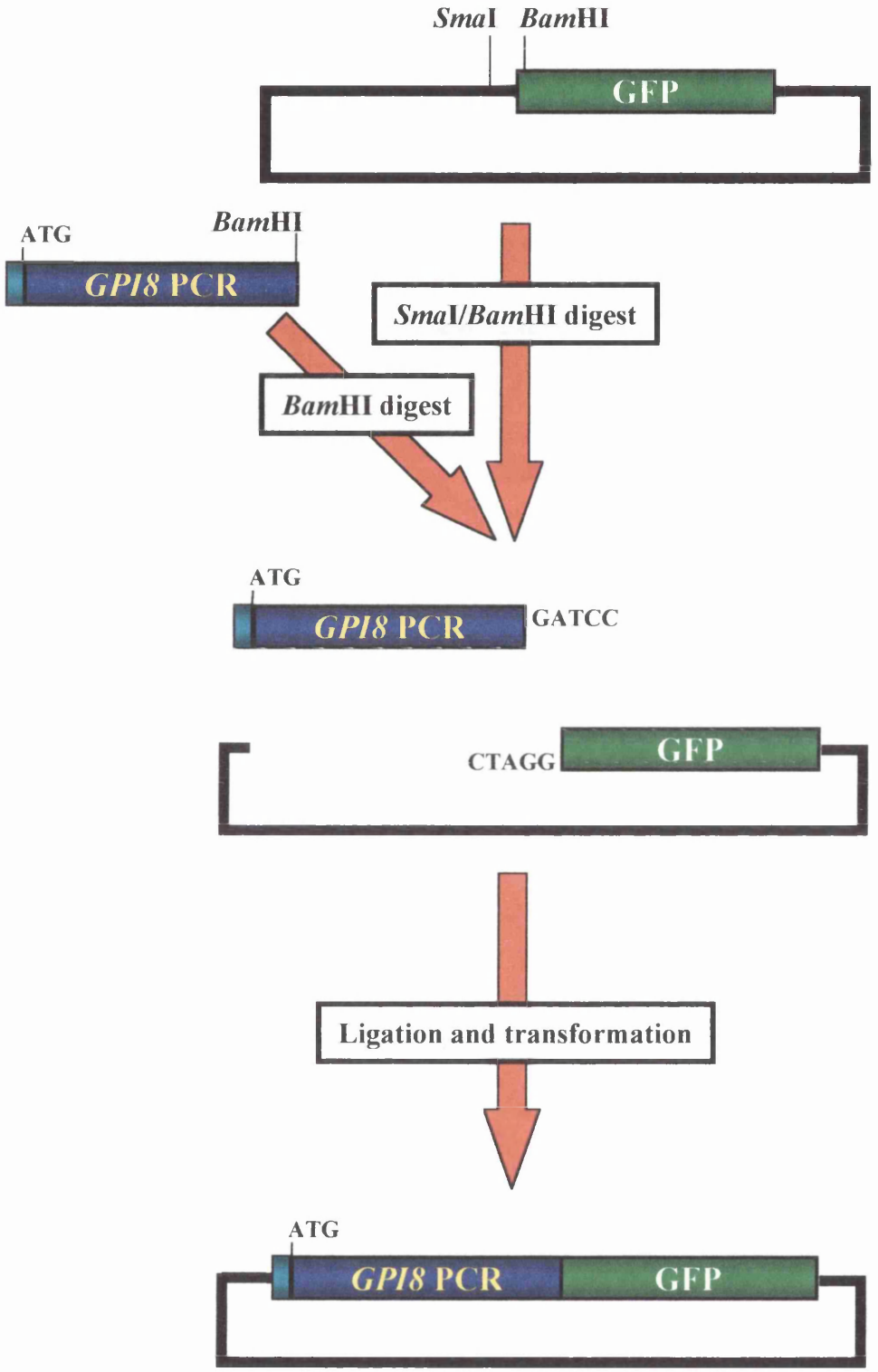
94°C 4 minutes x1

94°C, 30 seconds
55°C, 30 seconds } x15
72°C, 2 minutes

72°C, 4 minutes x1

A-overhangs were added to the PCR product, which was cloned into pGEM-T. The gene was excised from this construct by restriction digest with *Bam*HI and *Sal*I and subcloned into the similarly digested pQE-30 vector (Qiagen). Successful cloning was determined by PCR screening, restriction digest and sequencing (not shown).

Figure 5.2: Ligation of *GPI8* into pXG-GFP vector to allow expression of a GPI8-GFP fusion protein. *GPI8* was amplified by PCR as described in the text. The *Bam*HI digested PCR product was ligated into a *Sma*I/*Bam*HI digested pXG-GFP vector and used to transform *E. coli*. The new vector, designated pGL190 was isolated and sequenced (using primer OL462, see Chapter 3) across the new junction between *GPI8* and *GFP* to confirm that cloning had taken place correctly and that the fusion protein would be expressed correctly.


SmaI
BamHI
GFP
ATG
BamHI
GPI8 PCR
SmaI/BamHI digest
BamHI digest
ATG
GPI8 PCR
GATCC
CTAGG
GFP
Ligation and transformation
ATG
GPI8 PCR
GFP

pGL190

A

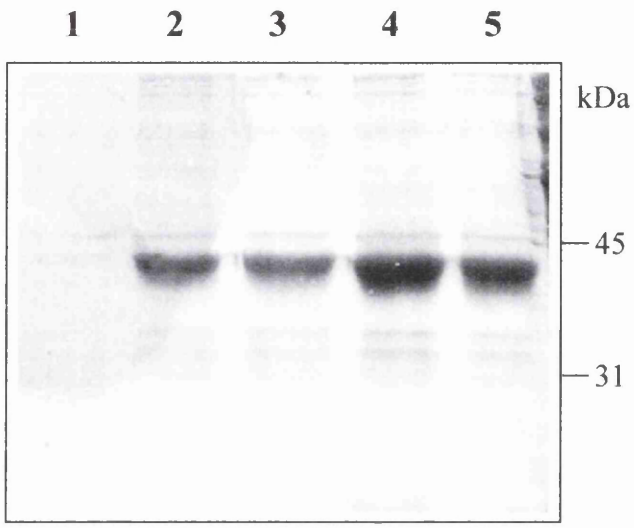


Figure 5.3: Expression of recombinant GPI8His in *E. coli* – SDS-PAGE and Coomassie staining of *E. coli* cell lysates selected with Ni-NTA. Cells were grown up for expression as described in the text and induced with 2mM IPTG for four hours. The time-course of rGPI8 expression was monitored by harvesting 1ml of cells at 0, 1, 2, 3 and 4 hours after IPTG addition. Cells were lysed in 150µl of Buffer A for 1 hour with stirring and were incubated with 35µl of Ni-NTA beads for a further 30 minutes. Beads were washed with 1ml of Buffer B and 3x1ml of Buffer C before elution in 50µl of Buffer C + 100mM EDTA. Samples were mixed with SDS-PAGE loading buffer and electrophoresed as follows:

Lane 1 – 0h

Lane 2 – 1h

Lane 3 – 2h

Lane 4 – 3h

Lane 5 – 4h

Figure 5.4: Large scale purification of GPI8His on a Ni-NTA column.

(A) An *E. coli* pellet was lysed directly in denaturing Buffer A for 1 hour (stirred) and filter-sterilised before passing through a 4ml Ni-NTA (Qiagen) column (pre-equilibriated with Buffer A). The column was washed with buffers B and C until no more protein could be detected (OD_{280}). Bound proteins were then eluted in Buffers D, E and F. Fractions were collected and analysed by SDS-PAGE. Gels were stained with Coomassie-blue.

Lane 1 – Buffer A flow through

Lane 2 – Buffer B wash

Lane 3 – Buffer C wash

Lane 4 – Buffer D elution

Lane 5 – Buffer E elution

Lane 6 – Buffer F elution.

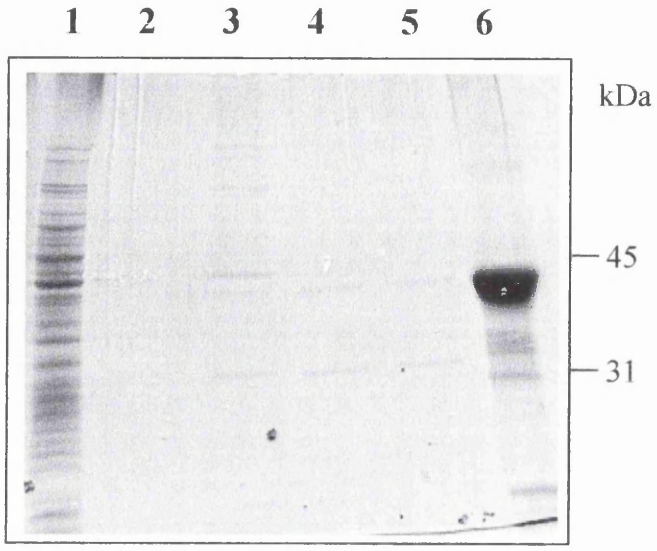
(B) Insoluble inclusion bodies were purified as described in the text and were solubilised in Buffer B and filtered. This was loaded onto a Ni-NTA column (pre-equilibriated with Buffer B) and the column was washed with Buffer B and C as before. Elution was also performed as before (in Buffers D, E and F and fractions were collected. SDS-PAGE and staining with Coomassie-blue was used to analyse the samples.

Lane 1 – Sample loaded onto column

Lane 2 – Buffer B flow through

Lane 3 – Buffer F elution containing GPI8His.

A



B

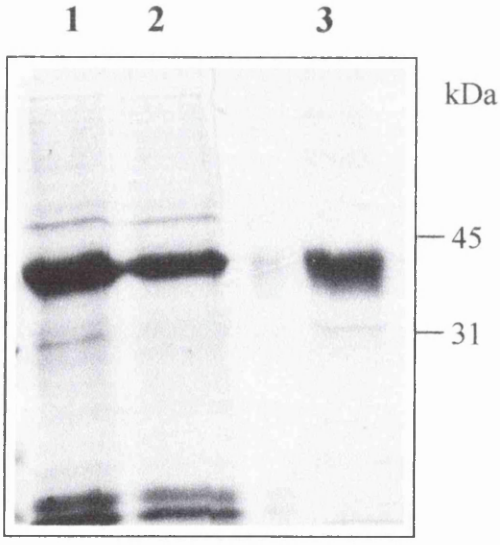


Figure 5.5: Western blot analysis of small-scale expression of GPI8His in *E. coli* expressing the chaperonins GroEL and GroES. *E. coli* AB1899F`[pGL184] and AB1899F`GroEL/ES [pGL184] were grown at 37°C for 3 hours and 19°C for 24 hours after addition of 500mM IPTG. 1ml samples collected at t = 0, t = 3 (for 37°C expression) and t = 24 (for 19°C expression) were lysed in 50µl sonication buffer/1mg/ml lysozyme and incubated at 37°C for 15 minutes before three freeze-thaw cycles. Soluble and insoluble material was separated by centrifugation (15,000xg, 10 minutes). The insoluble fraction was resuspended in 50µl of denaturing Buffer C and was mixed with Laemmli sample buffer. Soluble fractions were also mixed with Laemmli buffer and boiled prior to SDS-PAGE. Gels were electroblotted and exposed to R771 (peptide) anti-GPI8 antibodies at 1 in 50 dilution for 1 hour. Secondary antibody was used at 1 in 5000 dilution and the blot was incubated with ECL reagents (Pierce).

(A) AB1899F`[pGL184]

Lane 1 – t = 0, Soluble

Lane 2 – t = 0, Insoluble

Lane 3 – t = 3, (37°C) soluble

Lane 4 – t = 3, (37°C) insoluble

Lane 5 – t = 24, (19°C) soluble

Lane 6 – t = 24, (19°C) insoluble

(B) AB1899F`GroEL/ES[pGL184]

Lane 1 – t = 0, Soluble

Lane 2 – t = 0, Insoluble

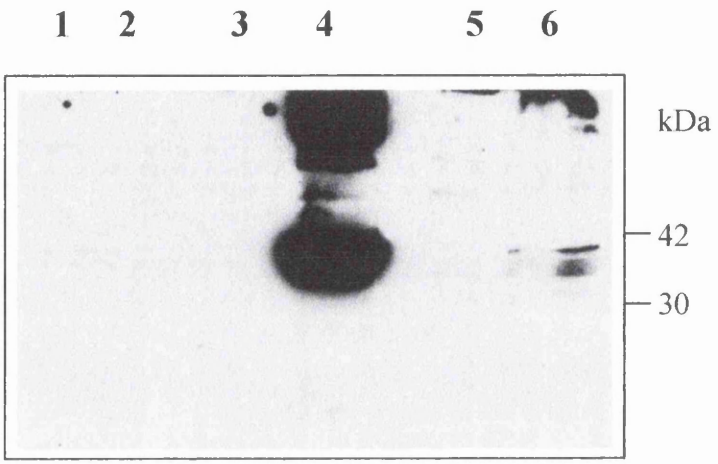
Lane 3 – t = 3, (37°C) soluble

Lane 4 – t = 3, (37°C) insoluble

Lane 5 – t = 24, (19°C) soluble

Lane 6 – t = 24, (19°C) insoluble

A



B

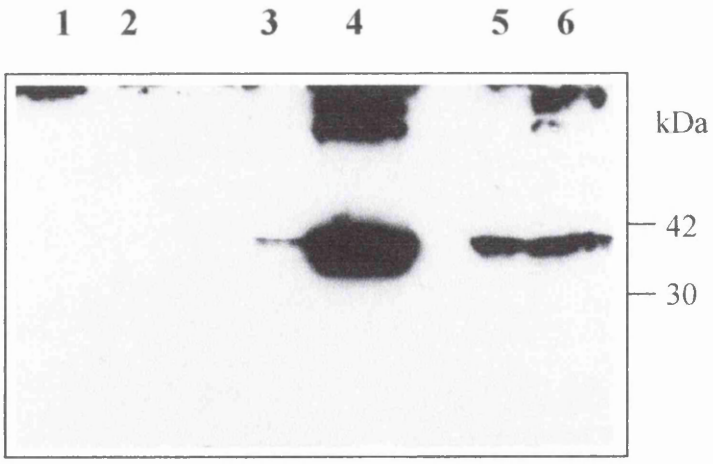


Figure 5.6: Large-scale expression of soluble recombinant GPI8His.

The proteins eluted from the Ni-NTA column were collected and fractions representing a cross-section of the purification were selected and concentrated using Centricon-10 ultrafiltration columns. Fractions were denatured with Laemmli reducing sample buffer, incubated at 37°C for 10 minutes and electrophoresed on a 10% SDS-PAGE gel. One gel was stained for protein with Coomassie-blue (**A**) and the other was electroblotted and exposed to affinity purified anti-peptide GPI8 antibodies (**B**).

Lane 1 – Fraction 2

Lane 2 – Fraction 3

Lane 3 – Fraction 4

Lane 4 – Fraction 5

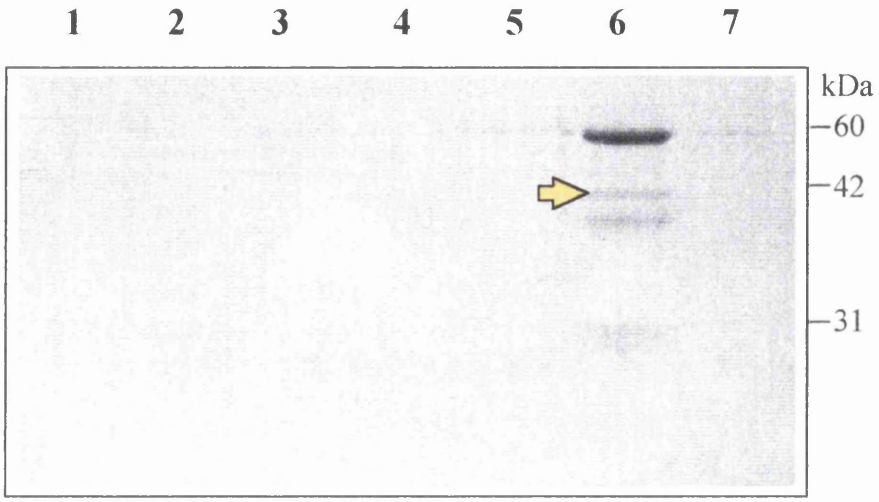
Lane 5 – Fraction 8

Lane 6 – Fraction 11

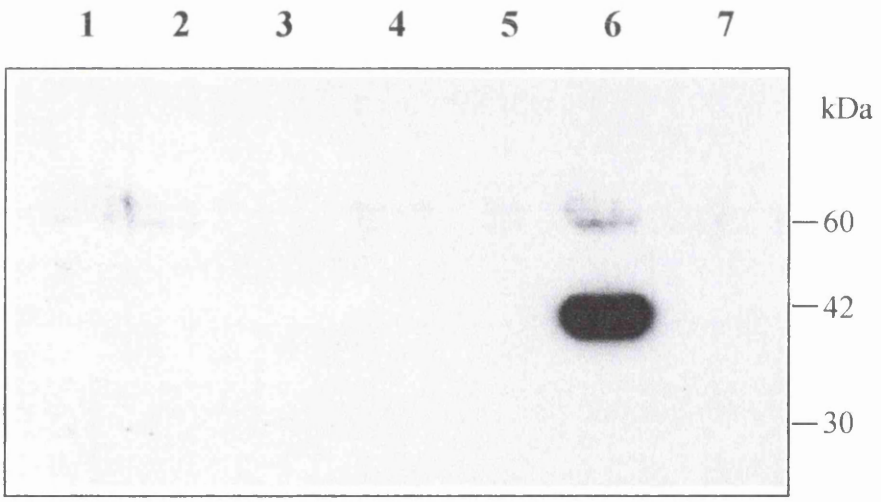
Lane 7 – Fraction 14.

The putative GPI8His is marked by a yellow arrow on the Coomassie gel (**A**) based on the size of the band detected by Western blotting (**B**).

A



B



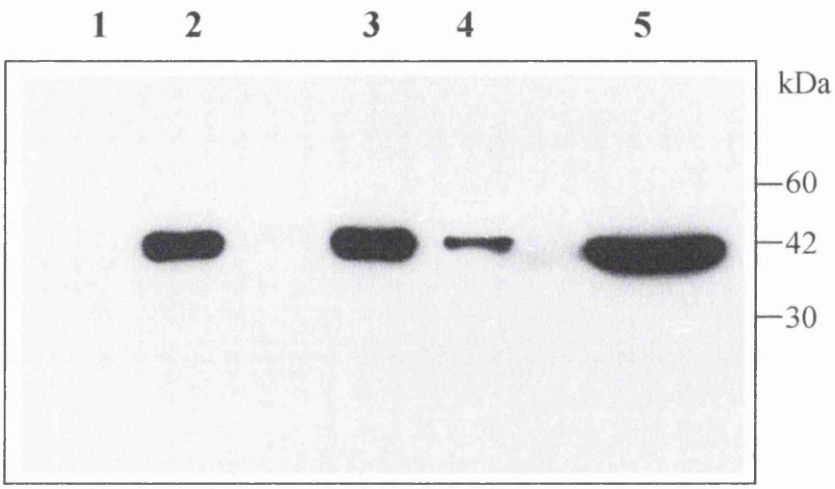


Figure 5.7: Western analysis of purified soluble GPI8His.

Fractions 9, 10, 12 and 13 from the large scale purification of soluble GPI8His were concentrated by ultrafiltration (Centricon-10) before SDS-PAGE and electroblotting. Blots were exposed to R771 antibodies as described in Figure 5.6.

Lane 1 – Fraction 9

Lane 2 – Fraction 10

Lane 3 – Fraction 12

Lane 4 – Fraction 13

Lane 5 – Soluble fraction prior to purification

Figure 5.8: Attempts to detect LMGPI8 in cell lysates using the Biotin/Avidin Labelling kit and enrichment of ER proteins by preparation of microsomes. Cell lysates from wild type, P1S1 and P1S1[pGL269] were electrophoresed and blotted onto Hybond C. This was done four times for each lysate **(A)**.

Blot 1; R491 anti-GPI8His antibodies

Blot 2; R492 anti-GPI8His antibodies

Blot 3; R771 anti-GPI8 peptide antibodies

Blot 4; R771 affinity purified antibodies

The detection of antibody binding was amplified using the Biotin/Avidin Labelling Kit (Santa Cruz ABC Reagents).

Lane 1: wild type

Lane 2: P1S1

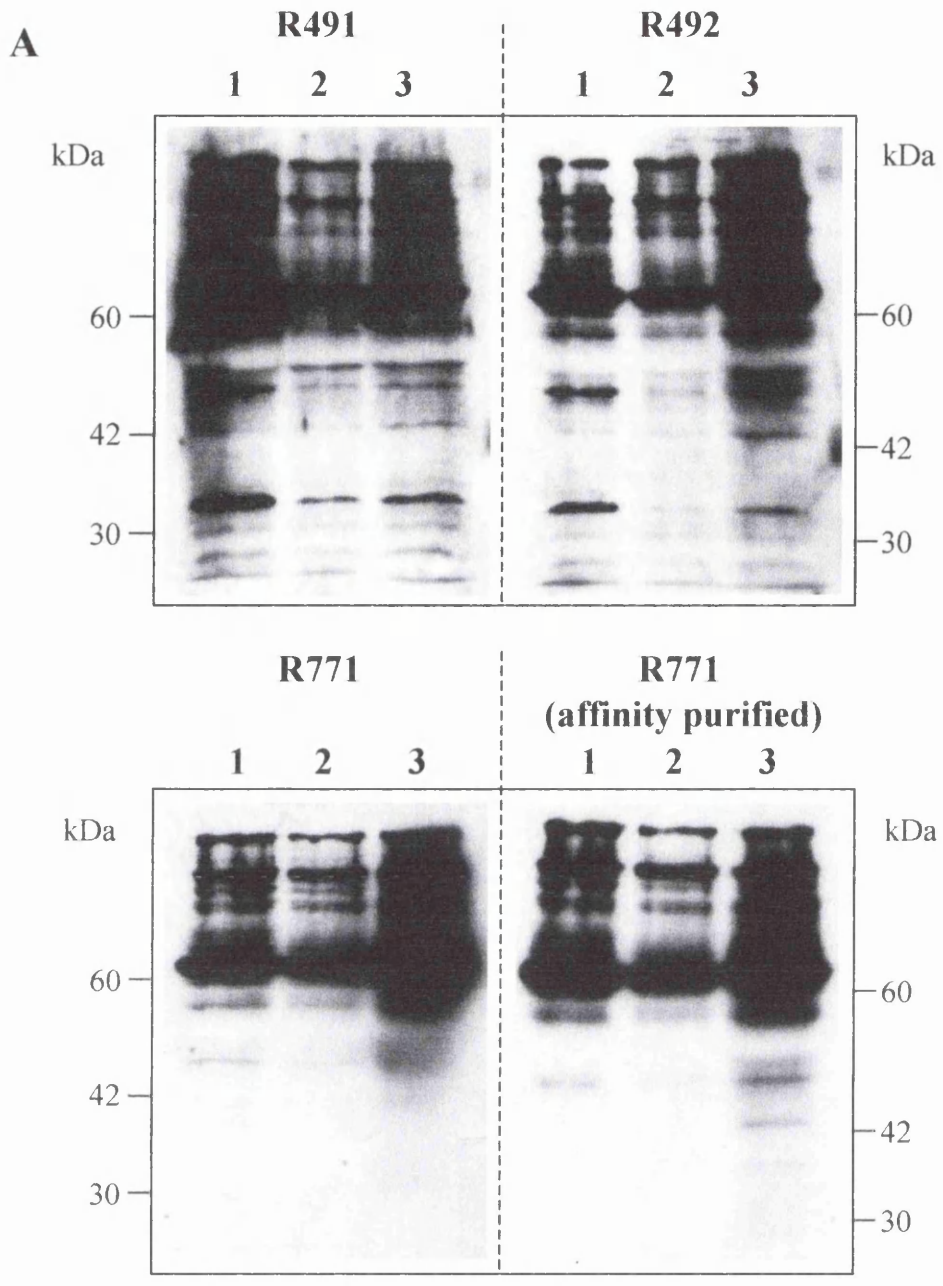
Lane 3: P1S1[pGL269]

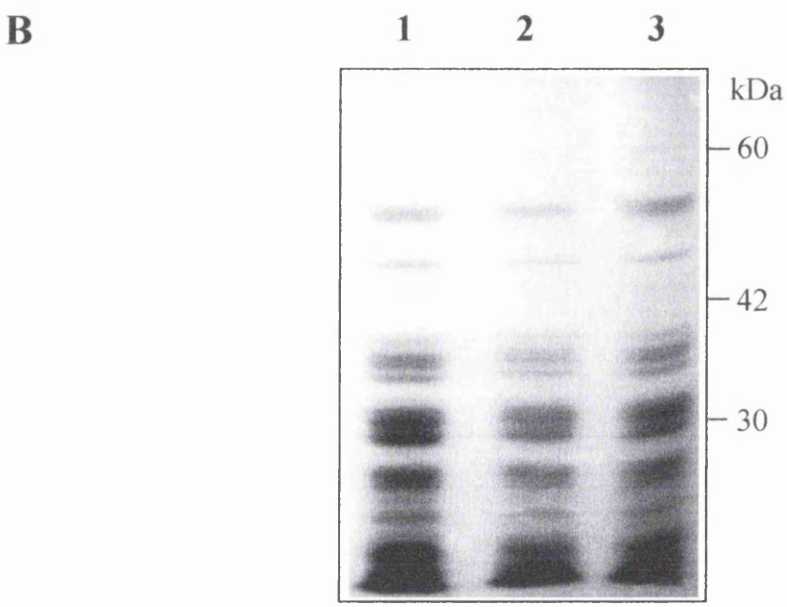
10^9 cells (wild type, P1S1 and P1S1[pGL269] were hypotonically lysed and microsomes were prepared as described in the text. Microsome pellets were resuspended in 60µl of 0.25M sucrose. Boiled samples were electrophoresed on two SDS-PAGE gels. One gel was stained with Coomassie blue **(B)** and the other was electroblotted and affinity purified R771 antibodies were used to attempt to detect GPI8. The latter is not shown, as GPI8 was not detected.

Lane 1: wild type

Lane 2: P1S1

Lane 3: P1S1[pGL269]


A
R491
R492
1 2 3
1 2 3
kDa
kDa
60
42
30
60
42
30
R771
R771
(affinity purified)
1 2 3
1 2 3
kDa
kDa
60
42
30
60
42
30


B
1 2 3
kDa
60
42
30

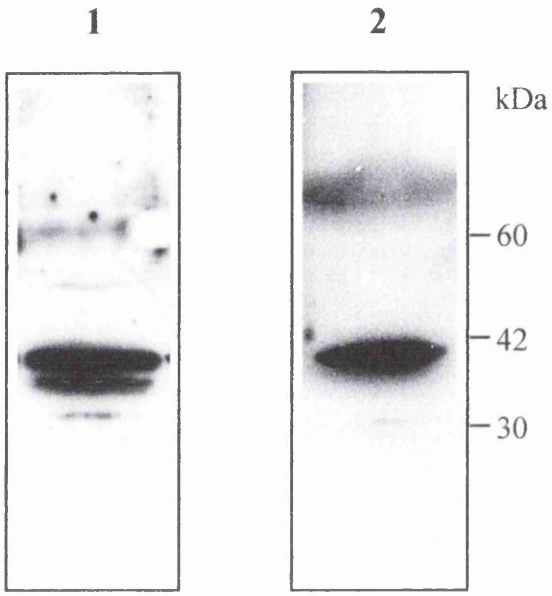


Figure 5.9: Analysis of *T. brucei* procyclic and bloodstream stage cell lysates by Western blotting with anti-GPI8His (R471) antibody.

10^8 cells from each life cycle stage were harvested and lysed in 150µl of 0.25% Triton X-100 in Laemmli sample buffer. 18µl ($1.2x10^7$ cell equivalents) was electrophoresed and electroblotted before exposure to R491 (anti-GPI8) antibodies. Anti-GPI8His antibody R491 was used at 1 in 1000 dilution with anti-rabbit IgG-HRP secondary antibody used at 1 in 5000 dilution

Lane 1: Procyclic cells

Lane 2: Bloodstream cells

Figure 5.10: Detection of QM, BiP and GPI8 in *T. brucei rhodesiense* procyclic cells by immunofluorescence.

Procyclic cells were smeared on slides and fixed by immersion in methanol and acetone using the method of Bangs *et al.*, (1996). Fluorescence was viewed with a Zeiss Axioplan microscope and images recorded and manipulated using the Openlab software (Improvision). Secondary antibodies used were anti-mouse IgG-FITC conjugated and anti-rabbit-TRITC conjugated (Sigma).

(A) Fixed cells were exposed simultaneously to mouse anti-QM (TY-tagged) antibodies (Panel B) and rabbit anti-LMGPI8 (R471) antibodies (Panel C). The images were overlaid (Panel D) and overlapping fluorescence is indicated in yellow.

(B) Cells were incubated with, in addition to the QM (TY) antibodies, rabbit anti-BiP, a gift of Dr. Jay Bangs, University of Wisconsin, USA. QM detection is shown in Panel F and BiP detection, in Panel G. Overlaid images are given in Panel H, again producing yellow fluorescence where the signals overlap.

Note that some cells expressing TY-tagged QM were not labelled with the anti-TY antibodies. A white arrow in (B) indicates such a cell.

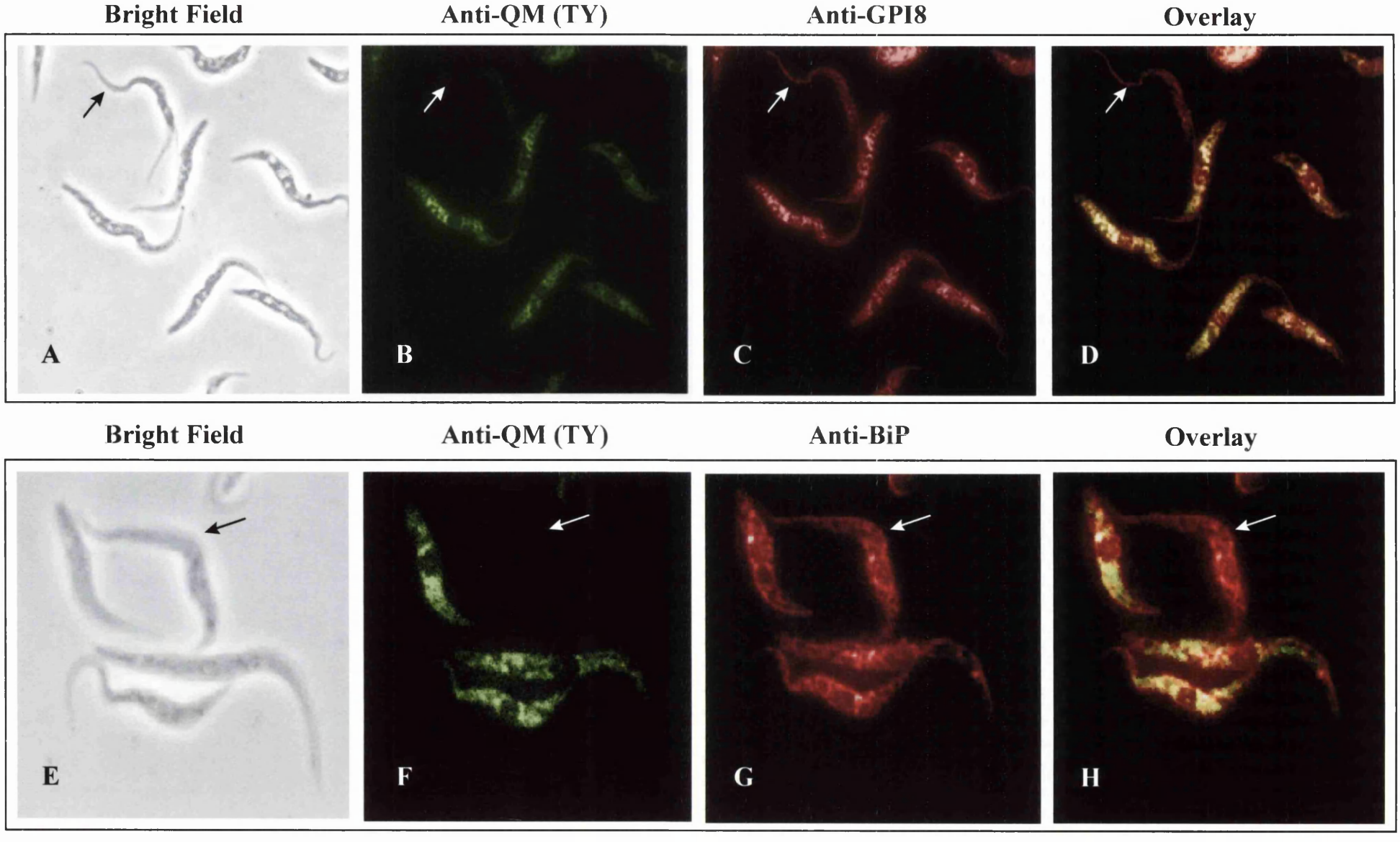

Bright Field
Anti-QM (TY)
Anti-GPI8
Overlay
A
B
C
D
Bright Field
Anti-QM (TY)
Anti-BiP
Overlay
E
F
G
H

Figure 5.11: Analysis of GP63 in *GPI8* null mutant cells expressing a GPI8-GFP fusion protein. Lysates were prepared from two P1S1[pGL190] clones GFP2 and GFP3 as well as wt, P1S1 and P1S1[pGL269]. These were subjected to SDS-PAGE on two gels at $1.5x10^7$ cell equivalents per lane. One was stained with Coomassie blue **(A)** to demonstrate approximately equal lane loadings and the other was electroblotted onto Hybond C **(B)**. Blots were exposed to anti-*L. major* GP63 monoclonal antibodies at 1 in 25 dilution. Higher mobility bands recognised by this antibody in wild type and *GPI8* re-expresser lysates are marked by yellow arrows.

Lane 1: wild type

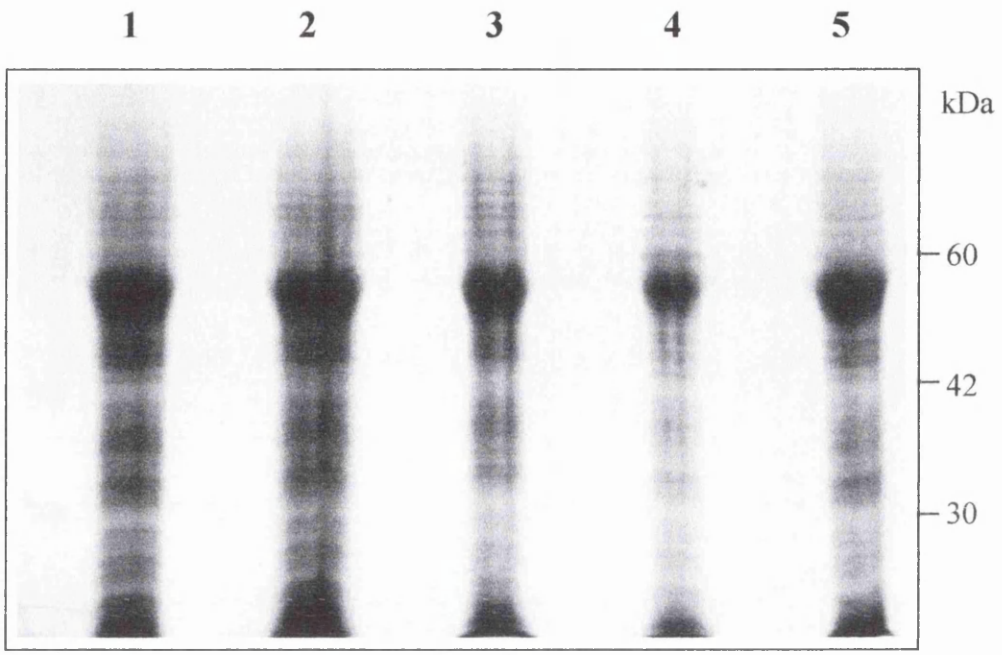
Lane 2: P1S1

Lane 3: P1S1[pGL269]

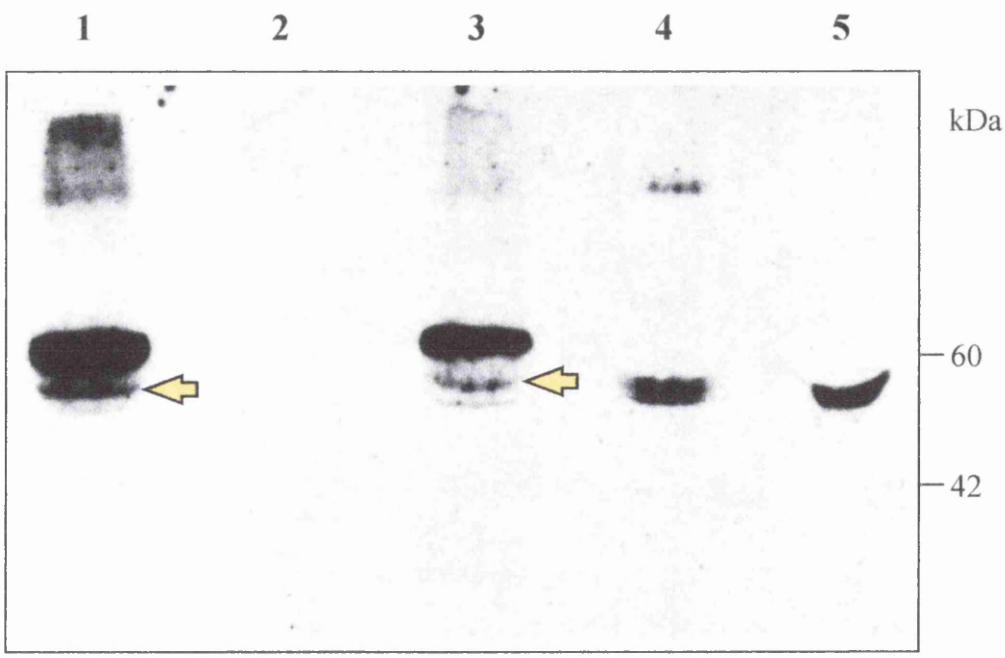
Lane 4: GFP2

Lane 5: GFP3

A



B



CHAPTER 6

CONCLUSIONS

The protein GPI biosynthetic pathway contains potential targets for anti-parasite chemotherapy and, other than dolichol phosphate mannose synthase (Mazhari-Tabrizi *et al.*, 1996; Ilgoutz *et al.*, 1999a), none of the genes involved in this pathway have been identified in the trypanosomatids.

The aim of this work was to isolate *GPI8* from *Leishmania mexicana* and to begin its characterisation. The gene encodes a component of the GPI:protein transamidase, which catalyses the terminal step in the protein GPI biosynthetic pathway, transfer of GPI anchors to nascent polypeptides.

Sequence analysis of the predicted *L. mexicana GPI8* ORF revealed that LMGPI8 protein shares 31% identity with both yeast and human GPI8 proteins, and 19% identity to legumain and schistosome haemoglobinase SM32. Several lines of evidence point to the GPI:protein transamidase having cysteine proteinase activity. The predicted mechanism of the transamidase (Figure 6.1) and evidence that the *T. brucei* reaction is sensitive to sulphydryl alkylating reagents (Sharma *et al.*, 1999a), together with the sequence homology with cysteine proteinases in the C13 family (Barrett and Rawlings, 1996) suggests that enzyme is a novel cysteine proteinase. A number of conserved amino acids, that are potential cysteine proteinase active site residues, have been identified (His^{63}, Cys^{94}, His^{174}, Cys^{216}). Cys^{216} is also conserved in Jack Bean legumain making this residue the most likely candidate for an active site

cysteine (see below). The *L. mexicana* protein lacks a C-terminal ER transmembrane domain found in the yeast and human proteins but does possess a putative ER retention motif (YDLE). Investigations as to whether this is a functional ER retention signal have yet to begin. Three potential N-glycosylation sites are present in the yeast (Benghezal *et al.*, 1996) and *L. mexicana* GPI8s but it is not yet known whether these are utilised. In contrast, the human GPI8 lacks N-linked glycosylation sites (Benghezal *et al.*, 1996; Yu *et al.*, 1997). It also remains to be determined if LMGPI8 is able to replace functionally, the yeast homologue and/or vice versa. Temperature-sensitive yeast *GPI8* mutants defective in GPI attachment (Benghezal *et al.*, 1996) are available to test this (a gift of Andreas Conzelmann).

Direct evidence that LMGPI8 is the GPI:protein transamidase came from targeted gene replacement of *GPI8* with antibiotic resistance markers and subsequent phenotype analysis. Failure to detect *GPI8* open reading frame by Southern analysis, as well as the prediction of the correct size of hybridisation fragments with a number of other probes indicated that the *GPI8* ORF had been lost. The isolation of GPI8-deficient promastigotes showed that the gene was not essential for viability of *L. mexicana* grown in culture and growth rates were comparable to wild type cells. The effect of these knockouts on GP63 was striking with the absence of GP63 on the cell surface, as well as the almost total loss of cell-associated GP63. [^{3}H]ethanolamine incorporation into GP63 in *GPI8* null mutants was undetectable indicating that GP63 did not receive a GPI anchor. Whether this lack of anchoring extends to all GPI-anchored proteins, including PSA2 (or GP46) remains to be tested. The fate of GP63 in *GPI8* null mutants also remains unclear since it does not appear to be secreted and inhibition of the proteasome and lysosomal cysteine proteinases does not significantly

block GP63 loss. The proteasome has been implicated in the quality control of GPI anchor attachment in mammalian cells (Wilbourn *et al.*, 1998). Alternatively, other lysosomal proteinases could be involved in degradation of unprocessed GP63. Further investigation of the fate GP63 will be carried out in the near future.

In vitro infection of macrophages revealed that *GPI8* null mutant promastigotes could infect macrophages, differentiate to amastigote forms and replicate, intracellularly, to wild type levels. This shows that GPI anchored proteins are not required for these processes, which means that this step in the biosynthetic pathway is unlikely to provide a useful anti-leishmanial drug target. Infection of BALB/c mice by *GPI8* null mutants and re-expresser cell lines was significantly reduced, however. This could be due to increased sensitivity to complement-mediated lysis of metacyclic promastigotes since GP63 is thought to be important here (Joshi *et al.*, 1998). The *GPI8* heterozygote used to generate the null mutants also do not infect BALB/c mice as effectively as wild type cells, so it is not clear if this reduced virulence is directly caused by the absence of the transamidase or perhaps by loss of virulence associated with the transfection and selection process. To investigate this, two independent *GPI8* heterozygote clones have been used to infect BALB/c mice but it is still too early to discuss their *in vivo* growth phenotype. The early indications from another ongoing experiment are that *GPI8* null mutants isolated from mouse lesions have increased virulence. Re-expression of *GPI8* in the null mutant background does not restore virulence to wild type levels and the reasons for this are not clear. Loss of the episome is one possible explanation and to address this, amastigotes could be isolated from lesions, differentiated to promastigotes *in vitro*, and analysed for GP63 expression on the cell surface. The levels of LPG and GIPL are unaffected in *GPI8*

null mutants indicating that reduced virulence is not a result of the disruption of the other GPI biosynthetic pathways caused by the observed buildup of complete protein anchors (Hilley, Zawadzki, McConville, Coombs and Mottram, manuscript submitted). Expression of GPI-anchored protein is maximal in promastigotes suggesting importance for survival in the sandfly stages. Whether these *GPI8* mutants are able to infect and survive in the sandfly will be tested in the near future.

Detection of LMGPI8 in *L. mexicana* cell lysates has not been possible to date, using the available antibodies raised against recombinant GPI8His and the anti-GPI8 peptide antibodies. The reasons for this are not clear, although one possibility is extensive glycosylation of the protein preventing epitope binding. Indeed, the GPI8 peptide used to raise antibodies includes one of the potential N-glycosylation attachment sites. Periodate treatment to destroy glycans, and tunicamycin treatment to prevent their attachment had no effect on the detection of GPI8. Enzymatic removal of glycans remains to be tested and production of epitope-tagged versions of GPI8 is under way. If the protein is heavily glycosylated then tagging the protein may not help. Thus, post-translational modifications of LMGPI8 have yet to be demonstrated. Mutant GPI8 in which the cysteine residue at position 216 was mutated to glycine (C216G), when re-expressed in the null mutant background did not restore transamidase activity, as assessed by detection of GP63 in cell lysates by Western blotting. This suggests that the protein is non-functional and that Cys^{216} may be the active site residue. However, the inability to detect GPI8 means that a lack of expression of the mutant gene could be an explanation for this rather than non-functional protein. The anti-GPI8His antibody detects a protein of 38-40kDa in *T. brucei* cell lysates and produces immunofluorescence patterns identical to those of an

epitope-tagged ribosomal protein, QM. However it cannot be ruled out that the anti-GPI8His antibodies are not cross-reacting with another ribosomal protein. It is not clear why these antibodies would recognise GPI8 in *T. brucei* but not in *L. mexicana*, however a 38-40kDa protein suggests little or no glycosylation of the *T. brucei* protein and consequently no blocking of the epitopes recognised by anti-GPI8His antibodies.

If epitope-tagging of LMGPI8 produces detectable, functional enzyme as determined by GP63 cell-surface expression, then this paves the way for the localisation of this enzyme. It will be interesting to determine if, like the other GPI biosynthetic enzymes, the GPI:protein transamidase is associated with the novel dolichol phosphate mannose synthase (DPMS) tubule (Ilgoutz *et al.*, 1999b) or if it remains within the bulk ER.

The yeast and human GPI:protein transamidase has at least two subunits but a homologue of GAA1 remains to be identified in trypanosomatids. Evidence discussed in Chapter 1 suggests that the molecular chaperone BiP may also play a role in the biosynthetic step, perhaps indirectly. The epitope-tag antibodies could lead to the identification of other components of the transamidase using similar methods to those of Kinoshita and coworkers who demonstrated that the mammalian GlcNAc transferase complex that catalyses the first step in GPI biosynthesis requires a complex of at least four proteins (see Watanabe *et al.*, 1996; Watanabe *et al.*, 1998 and Chapter 1).

Although attachment of GPI anchors to proteins does not appear to be an essential process in *L. mexicana*, it is likely to be essential in other trypanosomatids such as *T. brucei* that rely heavily on such proteins in host-parasite interactions. This work has provided unique insights into the function of GPI anchored proteins in *Leishmania* as well as providing access to the GPI:protein transamidase of African trypanosomes for future study.

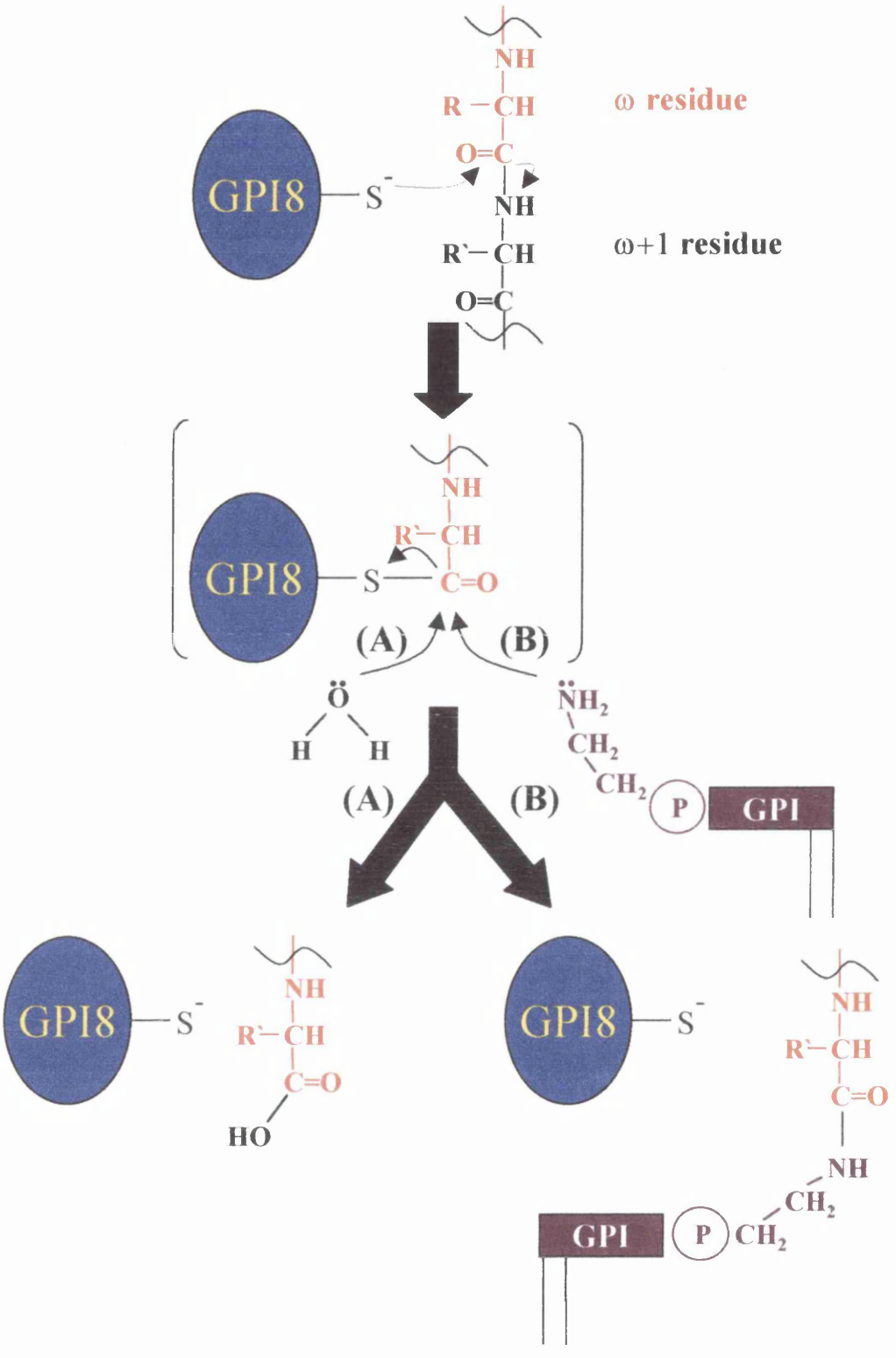


Figure 6.1: Predicted mechanism of the GPI:protein transamidase.

The transamidation mechanism allows GPI8 to behave as a proteinase by using water (A), rather than GPI (B) as the nucleophile in the second step of the reaction. (Adapted from Maxwell *et al.*, 1995a and Sharma *et al.*, 1999a).

REFERENCES

Abe, Y., Shirane, K., Yokosawa, H., Matsushita, H., Mitta, M., Kato, I., and Ishii, S.I. (1993). Asparaginyl endopeptidase of jack bean seeds. J. Biol. Chem. *268*, 3525-3529.

Alexander, J. and Vickerman, K. (1975). Fusion of host cell secondary lysosomes with the parasitophorous vacuoles of *Leishmania mexicana*-infected macrophages. J. Protozool. *22*, 502-508.

Alonso, G., Guerva, P., and Ramirez, J.L. (1992). Trypanosomatidae codon usage and GC distribution. Memorias de Instituto Oswaldo Cruz *87*, 517-532.

Amthauer, R., Kodukula, K., Brink, L., and Udenfriend, S. (1992). Phosphatidylinositol-glycan (PI-G)-anchored membrane proteins: requirement for ATP and GTP for translation-independent COOH-terminal processing. Proc. Natl. Acad. Sci. U.S.A. *89*, 6124-6128.

Amthauer, R., Kodukula, K., Gerber, L., and Udenfriend, S. (1993). Evidence that the putative COOH-terminal signal transamidase involved in glycosylphosphatidylinositol protein synthesis is present in the endoplasmic reticulum. Proc. Natl. Acad. Sci. USA *90*, 3973-3977.

Bahr, V., Stierhof, Y.-D., Ilg, T., Demar, M., Quinten, M., and Overath, P. (1993). Expression of lipophosphoglycan, high-molecular weight phosphoglycan and glycoprotein 63 in promastigotes and amastigotes of *Leishmania mexicana*. Mol. Biochem. Parasitol. *58*, 107-122.

Bailey, C.A., Gerber, L., Howard, A.D., and Udenfriend, S. (1989). Processing at the carboxyl terminus of nascent placental alkaline phosphatase in a cell-free system: evidence for specific cleavage of a signal peptide. Proc. Natl. Acad. Sci. USA *86*, 22-26.

Bailey, C.A., Howard, A., Micanovic, R., Berger, J., Heimer, E., Felix, A., Gerber, L., Brink, L., and Udenfriend, S. (1988). Site-directed antibodies for probing the structure and biogenesis of phosphatidylinositol glycan-linked membrane proteins: application to placental alkaline phosphatase. Anal. Biochem. *170*, 532-541.

Bangs, J.D. (1998). Surface coats and secretory trafficking in African trypanosomes. Current Opinion in Microbiology *1*, 448-454.

Bangs, J.D., Brouch, E.M., Ransom, D.M., and Roggy, J.L. (1996). A soluble secretory reporter system in *Trypanosoma brucei* - Studies on endoplasmic reticulum targeting. J. Biol. Chem. *271*, 18387-18393.

Bangs, J.D., Ransom, D.M., McDowell, M.A., and Brouch, E.M. (1997). Expression of bloodstream variant surface glycoproteins in procyclic stage *Trypanosoma brucei*: role of GPI anchors in secretion. EMBO J. *16*, 4285-4294.

Bangs, J.D., Uyetake, L., Brickman, M.J., Balber, A.E., and Boothroyd, J.C. (1993). Molecular cloning and cellular localization of a BiP homologue in *Trypanosoma brucei*. Divergent ER retention signals in a lower eukaryote. J. Cell Sci. *105*, 1101-1113.

Barrett, A.J. and Rawlings, N.D. (1996). Families and clans of cysteine peptidases. Perspect. Drug Disc. Design *6*, 1-11.

Barrett, M.P., Mottram, J.C., and Coombs, G.H. (1999). Recent advances in identifying and validating drug targets in trypanosomes and leishmanias. Trends in Microbiology *7*, 82-88.

Barry, J.D. (1997). The biology of antigenic variation in African trypanosomes. In Trypanosomiasis and Leishmaniasis: Biology and Control. G. Hide, J.C. Mottram, G.H. Coombs, and P.H. Holmes, eds. (Oxford: CAB International), pp. 89-107.

Bart, G., Coombs, G.H., and Mottram, J.C. (1995). Isolation of *lmcpc*, a gene encoding a *Leishmania mexicana* cathepsin B-like cysteine proteinase. Mol. Biochem. Parasitol. *73*, 271-274.

Bart, G., Frame, M.J., Carter, R., Coombs, G.H., and Mottram, J.C. (1997). Cathepsin B-like cysteine proteinase-deficient mutants of *Leishmania mexicana*. Mol. Biochem. Parasitol. *88*, 53-61.

Bastin, P., Bagherzadeh, A., Matthews, K.R., and Gull, K. (1996). A novel epitope tag system to study protein targeting and organelle biogenesis in *Trypanosoma brucei*. Mol. Biochem. Parasitol. *77*, 235-239.

Bates, P.A., Robertson, C.D., Tetley, L., and Coombs, G.H. (1992). Axenic cultivation and characterization of *Leishmania mexicana* amastigote-like forms. Parasitology *105*, 193-202.

Bayne, R.A.L., Kilbride, E.A., Lainson, F.A., Tetley, L., and Barry, J.D. (1993). A major surface antigen of procyclic stage *Trypanosoma congolense*. Mol. Biochem. Parasitol. *61*, 295-310.

Benghezal, M., Benachour, A., Rusconi, S., Aebi, M., and Conzelmann, A. (1996). Yeast Gpi8p is essential for GPI anchor attachment onto proteins. EMBO J. *15*, 6575-6583.

Berens, R.L., Brun, R., and Krassner, S.M. (1976). A simple monophasic medium for axenic culture of hemoflagellates. J. Parasitol. *62*, 360-365.

Beverley, S.M. and Turco, S.J. (1998). Lipophosphoglycan (LPG) and the identification of virulence genes in the protozoan parasite *Leishmania*. Trends Microbiol. *6*, 35-40.

Bordier, C., Etges, R.J., Ward, J., Turner, M.J., and Dealmeida, M.L.C. (1986). Leishmania and trypanosoma surface glycoproteins have a common glycophospholipid membrane anchor. Proc. Natl. Acad. Sci. USA *83* , 5988-5991.

Brittingham, A., Morrison, C.J., McMaster, W.R., McGwire, B.S., Chang, K.P., and Mosser, D.M. (1995). Role of the *Leishmania* surface protease gp63 in complement fixation, cell adhesion, and resistance to complement-mediated lysis. J. Immunol. *155*, 3102-3111.

Butcher, B.A., Turco, S.J., Hilty, B.A., Pimenta, P.F.P., Panunzio, M., and Sacks, D.L. (1996). Deficiency in β1,3-galactosyltransferase of a *Leishmania major* lipophosphoglycan mutant adversely influences the *Leishmania*-sand fly interaction. J.Biol.Chem. *271*, 20573-20579.

Button, L.L., Reiner, N.E., and McMaster, W.R. (1991). Modification of GP63 genes from diverse species of *Leishmania* for expression of recombinant protein at high levels in *Escherichia coli*. Mol. Biochem. Parasitol. *44*, 213-224.

Button, L.L., Russell, D.G., Klein, H.L., Medina-Acosta, E., Karess, R.E., and McMaster, W.R. (1989). Genes encoding the major surface glycoprotein in *Leishmania* are tandemly linked at a single chromosomal locus and are constitutively transcribed. Mol. Biochem. Parasitol. *32*, 271-283.

Canivenc-Gansel, E., Imhof, I., Reggiori, F., Burda, P., Conzelmann, A., and Benachour, A. (1998). GPI anchor biosynthesis in yeast: phosphoethanolamine is attached to the α1,4-linked mannose of the complete precursor glycophospholipid. Glycobiology *8*, 761-770.

Caras, I.W., Weddell, G.N., Davitz, M.A., Nussenzweig, V., and Martin, D.W. (1987). Signal for attachment of a phospholipid membrane anchor in decay accelerating factor. Science *238*, 1280-1283.

Caras, I.W., Weddell, G.N., and Williams, S.R. (1989). Analysis of the signal for attachment of a glycophospholipid membrane anchor. J. Cell Biol. *108*, 1387-1396.

Carmona, E., Dufour, E., Plouffe, C., Takebe, S., Mason, P., Mort, J.S., and Menard, R. (1996). Potency and selectivity of the cathepsin L propeptide as an inhibitor of cysteine proteases. Biochemistry *35*, 8149-8157.

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. J. Biol. Chem. *272*, 8090-8098.

Chen, J.M., Dando, P.M., Stevens, R.A.E., Fortunato, M., and Barrett, A.J. (1998). Cloning and expression of mouse legumain, a lysosomal endopeptidase. Biochem. J. *335*, 111-117.

Chen, R., Udenfriend, S., Prince, G.M., Maxwell, S.E., Ramalingham, S., Gerber, L.D., Knez, J., and Medof, M.E. (1996). A defect in glycosylphosphatidylinositol (GPI) transamidase activity in mutant K cells is responsible for their inability to display GPI surface proteins. Proc. Natl. Acad. Sci. USA *93*, 2280-2284.

Chen, R., Walter, E.I., Parker, G., Lapurga, J.P., Millan, J.L., Ikehara, Y., Udenfriend, S., and Medof, M.E. (1998). Mammalian glycophosphatidylinositol anchor transfer to proteins and posttransfer deacylation. Proc. Natl. Acad. Sci. USA *95*, 9512-9517.

Clayton, C., Adams, M.A.R., Baltz, T., Barrett, M., Bastien, P., Belli, S., Beverley, S., Biteau, N., Blackwell, J., Blaineau, C.B.M., Bringaud, F., Cross, G., Cruz, A., Degrave, W., Donelson, J., El-Sayed, N., Fu, G., Ersfeld, K., Gibson, W., Gull, K., Ivens, A., Kelly, J., Lawson, D., Lebowitz, J., Majiwa, P., Mathews, K., Melville, S., Merlin, G., Michels, P.M.P., Norrish, A., Opperdoes, F., Papadopoulou, B., Parsons, M., Seebeck, T., Smith, D., Stuart, K., Turner, M., Ullu, E., and Vanhamme, L. (1998). Genetic nomenclature for *Trypanosoma* and *Leishmania*. Mol. Biochem. Parasitol. *97*, 221-224.

Clayton, C.E. and Mowatt, M.R. (1989). The procyclic acidic repetitive proteins of *Trypanosoma brucei* - purification and post-translational modification. J. Biol. Chem. *264*, 15088-15093.

Coburn, C.M., Otteman, K.M., Mcneely, T., Turco, S.J., and Beverley, S.M. (1991). Stable DNA transfection of a wide range of trypanosomatids. Mol. Biochem. Parasitol. *46*, 169-180.

Colussi, P.A., Taron, C.H., Mack, J.C., and Orlean, P. (1997). Human and *Saccharomyces cerevisiae* dolichol phosphate mannose synthases represent two classes of the enzyme, but both function in *Schizosaccharomyces pombe*. Proc. Natl. Acad. Sci. USA *94*, 7873-7878.

Coombs, G.H. and Mottram, J.C. (1997). Proteinases of trypanosomes and *Leishmania*. In Trypanosomiasis and leishmaniasis: Biology and control. G. Hide, J.C. Mottram, G.H. Coombs, and P.H. Holmes, eds. (Oxford, UK: CAB International), pp. 177-197.

Costello, L.C. and Orlean, P. (1992). Inositol acylation of a potential glycosyl phosphatidylinositol anchor precursor from yeast requires acyl coenzyme A. J. Biol. Chem. *267*, 8599-8603.

Coulson, R.M.R., Conner, V., Chen, J.C., and Ajioka, J.W. (1996). Differential expression of *Leishmania major* β-tubulin genes during the acquisition of promastigote infectivity. Mol. Biochem. Parasitol. *82*, 227-236.

Croft, S.L., Urbina, J.A., and Brun, R. (1997). Chemotherapy of human Leishmaniasis and Trypanosomiasis. In Trypanosomiasis and Leishmaniasis. G. Hide, J.C. Mottram, G.H. Coombs, and P.H. Holmes, eds. (Oxford: CAB International), pp. 245-257.

Cross, G.A.M. (1990). Glycolipid anchoring of plasma membrane proteins. Ann. Rev. Cell Biol. *6*, 1-39.

Cruz, A.K. and Beverley, S.M. (1990). Gene replacement in parasitic protozoa. Nature *348*, 171-173.

Cruz, A.K., Coburn, C.M., and Beverley, S.M. (1991). Double targeted gene replacement for creating null mutants. Proc. Natl. Acad. Sci. USA *88*, 7170-7174.

Cruz, A.K., Titus, R., and Beverley, S.M. (1993). Plasticity in chromosome number and testing of essential genes in *Leishmania* by targeting. Proc. Natl. Acad. Sci. USA *90*, 1599-1603.

Curotto de Lafaille, M.A., Laban, A., and Wirth, D.F. (1992). Gene expression in *Leishmania* - analysis of essential 5' DNA sequences. Proc. Natl. Acad. Sci. USA *89*, 2703-2707.

Curotto de Lafaille, M.A. and Wirth, D.F. (1992). Creation of Null/+ mutants of the alpha-tubulin gene in *Leishmania enriettii* by gene cluster deletion. J. Biol. Chem. *267*, 23839-23846.

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

Dalton, J.P., Smith, A.M., Clough, K.A., and Brindley, P.J. (1995b). Digestion of haemoglobin by Schistosomes: 35 years on. Parasitol. Today *11*, 299-303.

Doering, T.L., Masterson, W.J., Englund, P.T., and Hart, G.W. (1989). Biosynthesis of the glycosyl phosphatidylinositol membrane anchor of the trypanosome variant surface glycoprotein. Origin of the non-acetylated glucosamine. J. Biol. Chem. *264*, 11168-11173.

Doerrler, W.T., Ye, J., Falck, J.R., and Lehrman, M.A. (1996). Acylation of glucosaminyl phosphatidylinositol revisited. J. Biol. Chem. *271*, 27031-27038.

Dowdy, S.F., Lai, K.M., Weissman, B.E., Matsui, Y., Hogan, B.L.M., and Stanbridge, E.J. (1991). The isolation and characterization of a novel cDNA demonstrating an altered mRNS level in nontumorigenic Wilms' microcell hybrid cells. Nucleic Acids Res. *19*, 5763-5769.

Dumas, C., Ouellette, M., Tovar, J., Cunningham, M.L., Fairlamb, A.H., Tamar, S., Olivier, M., and Papadopoulou, B. (1997). Disruption of the trypanothione reductase gene of *Leishmania* decreases its ability to survive oxidative stress in macrophages. EMBO J. *16*, 2590-2598.

Eakin, A.E., Bouvier, J., Sakanari, J.A., Craik, C.S., and McKerrow, J.H. (1990). Amplification and sequencing of genomic DNA fragments encoding cysteine proteases from protozoan parasites. Mol. Biochem. Parasitol. *39*, 1-8.

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

El-Sayed, N.M.A. and Donelson, J.E. (1997). African trypanosomes have differentially expressed genes encoding homologues of the *Leishmania* GP63 surface protease. J. Biol. Chem. *272*, 26742-26748.

Engel, J.C., Doyle, P.S., Palmer, J., Hsieh, I., Bainton, D.F., and Mckerrow, J.H. (1998). Cysteine protease inhibitors alter Golgi complex ultrastructure and function in *Trypanosoma cruzi*. J. Cell Sci. *111*, 597-606.

Englund, P.T. (1993). The structure and biosynthesis of glycosyl phosphatidylinositol protein anchors. Annu. Rev. Biochem. *62*, 121-138.

Ferguson, M.A.J., Homans, S.W., Dwek, R.A., and Rademacher, T.W. (1988). Glycosyl-phosphatidylinositol moiety that anchors *Trypanosoma brucei* variant surface glycoprotein to the membrane. Science *239*, 753-759.

Ferguson, M.A.J., Low, M.G., and Cross, G.A.M. (1985). Glycosyl-sn-1,2-dimyristylphosphatidylinositol is covalently linked to *Trypanosoma brucei* variant surface glycoprotein. J. Biol. Chem. *260*, 4547-4555.

Ferguson, M.A.J. and Williams, A.F. (1988). Cell-surface anchoring of proteins via glycosylphosphatidylinositol structures. Annu. Rev. Biochem. *57*, 285-320.

Field, M.C., Menon, A.K., and Cross, G.A.M. (1991). A glycosylphosphatidylinositol protein anchor from procyclic stage *Trypanosoma brucei* - lipid structure and biosynthesis. EMBO J. *10*, 2731-2739.

Field, M.C., Menon, A.K., and Cross, G.A.M. (1992). Developmental variation of glycosylphosphatidylinositol membrane anchors in *Trypanosoma brucei*. *In vitro* biosynthesis of intermediates in the construction of the GPI anchor of the major procyclic surface glycoprotein. J. Biol. Chem. *267*, 5324-5329.

Frame, M. J. (1997). Studies on the cysteine proteinases of *Leishmania mexicana*. University of Glasgow. Thesis submitted for the degree of Doctor of Philosophy.

Freedman, D.J. and Beverley, S.M. (1993). Two more independent selectable markers for stable transfection of *Leishmania*. Mol. Biochem. Parasitol. *62*, 37-44.

Frommel, T.O., Button, L.L., Fujikura, Y., and McMaster, W.R. (1990). The major surface glycoprotein (gp63) is present in both life stages of *Leishmania*. Mol. Biochem. Parasitol. *38*, 25-32.

Funk, V.A., Jardim, A., and Olafson, R.W. (1994). An investigation into the significance of the *N*-linked oligosaccharides of *Leishmania* gp63. Mol. Biochem. Parasitol. *63*, 23-35.

Garg, N., Tarleton, R.L., and Mensa-Wilmot, K. (1997). Proteins with glycosylphosphatidylinositol (GPI) signal sequences have divergent fates during a GPI deficiency. J. Biol. Chem. *272*, 12482-12491.

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*. Biochem. J. *290*, 801-806.

Grant, K.M., Hassan, P., Anderson, J.S., and Mottram, J.C. (1998). The *crk3* gene of *Leishmania mexicana* encodes a stage-regulated cdc2-related histone H1 kinase that associates with $p12^{cks1}$. J. Biol. Chem. *273*, 10153-10159.

Gueiros-Filho, F.J. and Beverley, S.M. (1994). On the introduction of genetically modified *Leishmania* outside the laboratory. Exp. Parasitol. *78*, 425-428.

Guther, M.L. and Ferguson, M.A. (1995). The role of inositol acylation and inositol deacylation in GPI biosynthesis in *Trypanosoma brucei*. EMBO J. *14*, 3080-3093.

Guther, M.L., Masterson, W.J., and Ferguson, M.A. (1994). The effects of phenylmethylsulfonyl fluoride on inositol-acylation and fatty acid remodelling in African trypanosomes. J. Biol. Chem. *269*, 18694-18701.

Ha, D.S., Schwartz, J.K., Turco, S.J., and Beverley, S.M. (1996). Use of the green fluorescent protein as a marker in transfected *Leishmania*. Mol. Biochem. Parasitol. *77*, 57-64.

Hamburger, D., Egerton, M., and Riezman, H. (1995). Yeast Gaa1p is required for attachment of a completed GPI anchor onto proteins. J. Cell Biol. *129*, 629-639.

Handman, E., Osborn, A.H., Symons, F., Van Driel, R., and Cappai, R. (1995). The *Leishmania* promastigote surface antigen 2 complex is differentially expressed during the parasite life cycle. Mol. Biochem. Parasitol. *74*, 189-200.

Hara-Nishimura, I., Shimada, T., Hiraiwa, N., and Nishimura, M. (1995). Vacuolar processing enzyme responsible for maturation of seed proteins. Journal of Plant Physiology *145*, 632-640.

Hilley, J.D., Zawadzki, J.L., McConville, M.J., Coombs, G.H., and Mottram, J.C. (1999). *Leishmania mexicana* mutants lacking GPI:protein transamidase provide insights into the biosynthesis and function of GPI-anchored proteins. Submitted

Hiroi, Y., Komuro, I., Chen, R., Hosoda, T., Mizuno, T., Kudoh, S., Georgescu, S.P., Medof, M.E., and Yazaki, Y. (1998). Molecular cloning of human homolog of yeast *GAA1* which is required for attachment of glycosylphosphatidylinositols to proteins. FEBS Lett. *421*, 252-258.

Hirose, S., Prince, G.M., Sevlever, D., Ravi, L., Rosenberry, T.L., Ueda, E., and Medof, M.E. (1992). Characterization of putative glycoinositol phospholipid anchor precursors in mammalian cells. Localization of phosphoethanolamine. J. Biol. Chem. *267*, 16968-16974.

Hong, Y., Ohishi, K., Watanabe, R., Endo, Y., Maeda, Y., and Kinoshita, T. (1999). GPI1 stabilizes an enzyme essential in the first step of glycosylphosphatidylinositol biosynthesis. J. Biol. Chem. *274*, 18582-18588.

Ilg, T., Harbecke, D., and Overath, P. (1993) The lysosomal gp63-related protein in *Leishmania mexicana* amastigotes is a soluble metalloproteinase with an acidic pH optimum. FEBS Letters 327, 103-107.

Ilgoutz, S.C., Mullin, K.A., Southwell, B.R., and McConville, M.J. (1999b). Glycosylphosphatidylinositol biosynthetic enzymes are localized to a stable tubular subcompartment of the endoplasmic reticulum in *Leishmania mexicana*. EMBO J. *18*, 3643-3654.

Ilgoutz, S.C., Zawadzki, J.L., Ralton, J.E., and McConville, M.J. (1999a). Evidence that free GPI glycolipids are essential for growth of *Leishmania mexicana*. EMBO J. *18*, 2746-2755.

Inoue, N., Kinoshita, T., Orii, T., and Takeda, J. (1993). Cloning of a human gene, *PIG-F*, a component of glycosylphosphatidylinositol anchor biosynthesis, by a novel expression cloning strategy. J. Biol. Chem. *268*, 6882-6885.

Inoue, N., Watanabe, R., Takeda, J., and Kinoshita, T. (1996). *PIG-C*, one of the three human genes involved in the first step of glycosylphosphatidylinositol biosynthesis is a homologue of *Saccharomyces cerevisiae GPI2*. Biochem. Biophys. Res. Commun. *226*, 193-199.

Ishii, S. (1994). Legumain - asparaginyl endopeptidase. Methods Enzymol. *244*, 604-615.

Joshi, P.B., Sacks, D.L., Modi, G., and McMaster, W.R. (1998). Targeted gene deletion of *Leishmania major* genes encoding developmental stage-specific leishmanolysin (GP63). Mol. Microbiol. *27*, 519-530.

Joshi, P.B., Webb, J.R., Davies, J.E., and McMaster, W.R. (1995). The gene encoding streptothricin acetyltransferase (*sat*) as a selectable marker for *Leishmania* expression vectors. Gene *156*, 145-149.

Kamitani, T., Chang, H.M., Rollins, C., Waneck, G.L., and Yeh, E.T. (1993). Correction of the class H defect in glycosylphosphatidylinositol anchor biosynthesis in Ltk- cells by a human cDNA clone. J. Biol. Chem. *268*, 20733-20736.

Kamitani, T., Menon, A.K., Hallaq, Y., Warren, C.D., and Yeh, E.T. (1992). Complexity of ethanolamine phosphate addition in the biosynthesis of glycosylphosphatidylinositol anchors in mammalian cells. J. Biol. Chem. *267*, 24611-24619.

Kelleher, M., Bacic, A., and Handman, E. (1992). Identification of a macrophage-binding determinant on lipophosphoglycan from *Leishmania major* promastigotes. Proc. Natl. Acad. Sci. USA *89*, 6-10.

Kink, J.A. and Chang, K.P. (1987). Tunicamycin-resistant *Leishmania mexicana amazonensis*: expression of virulence associated with an increased activity of *N*-acetylglucosaminyltransferase and amplification of its presumptive gene. Proc. Natl. Acad. Sci. U.S.A. *84*, 1253-1257.

Kinoshita, T., Ohishi, K., and Takeda, J. (1997). GPI-anchor synthesis in mammalian cells: Genes, their products, and a deficiency. J. Biochem.(Tokyo) *122*, 251-257.

Kirschke, H. and Barrett, A.J. (1987). Chemistry of lysosomal proteases. In Lysosomes: Their Role in Protein Breakdown. H. Glaumann and F.J. Ballard, eds. (London: Academic Press), pp. 193-238.

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

Koch, F., Thiele, H.G., and Low, M.G. (1986). Release of the rat T cell alloantigen RT-6.2 from cell membranes by phosphatidylinositol-specific phospholipase C. J. Exp. Med. *164*, 1338-1343.

Kodukula, K., Cines, D., Amthauer, R., Gerber, L., and Udenfriend, S. (1992). Biosynthesis of phosphatidylinositol-glycan (PI-G)-anchored membrane proteins in cell-free systems:

cleavage of the nascent protein and addition of the PI-G moiety depend on the size of the COOH-terminal signal peptide. Proc. Natl. Acad. Sci. U.S.A. *89*, 1350-1353.

Kodukula, K., Gerber, L.D., Amthauer, R., Brink, L., and Udenfriend, S. (1993). Biosynthesis of glycosylphosphatidylinositol (GPI)-anchored membrane proteins in intact cells: specific amino acid requirements adjacent to the site of cleavage and GPI attachment. J. Cell Biol. *120*, 657-664.

Kodukula, K., Micanovic, R., Gerber, L., Tamburrini, M., Brink, L., and Udenfriend, S. (1991). Biosynthesis of phosphatidylinositol glycan-anchored membrane proteins. J. Biol. Chem. *266*, 4464-4470.

Koller, H.T., Klade, T., Ellinger, A., and Breitenbach, M. (1996). The yeast grwoth control gene *GRC5* is highly homologous to the mammalian putative tumor suppressor gene *QM*. Yeast *12*, 53-65.

Krakow, J.L., Doering, T.L., Masterson, W.J., Hart, G.W., and Englund, P.T. (1989). A glycolipid from *Trypanosoma brucei* related to the variant surface glycoprotein membrane anchor. Mol. Biochem. Parasitol. *36*, 263-270.

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

Laban, A., Tobin, J.F., Curotto de Lafaille, M.A., and Wirth, D.F. (1990). Stable expression of the bacterial neo^r gene in *Leishmania enriettii*. Nature *343*, 572-574.

Langford, C.K., Ullman, B., and Landfear, S.M. (1992). *Leishmania* - Codon Utilization of Nuclear Genes. Exp. Parasitol. *74*, 360-361.

LeBowitz, J.H. (1994). Transfection experiments with *Leishmania*. Methods in Cell Biology *45*, 65-78.

LeBowitz, J.H., Smith, J., Rusche, L., and Beverley, S.M. (1994). Coupling of poly(A) site selection and *trans*-splicing in *Leishmania*. Genes & Dev. *7*, 996-1007.

Lee, M.G.S. and Van der Ploeg, L.H.T. (1991). The hygromycin b-resistance-encoding gene as a selectable marker for stable transformation of *Trypanosoma brucei*. Gene *105*, 255-257.

Leidich, S.D., Drapp, D.A., and Orlean, P. (1994). A conditionally lethal yeast mutant blocked at the first step in glycosyl phosphatidylinositol anchor synthesis. J. Biol. Chem. *269*, 10193-10196.

Leidich, S.D., Kostova, Z., Latek, R.R., Costello, L.C., Drapp, D.A., Gray, W., Fassler, J.S., and Orlean, P. (1995). Temperature-sensitive yeast GPI anchoring mutants *gpi2* and *gpi3* are defective in the synthesis of N-acetylglucosaminyl phosphatidylinositol. Cloning of the *GPI2* gene. J. Biol. Chem. *270*, 13029-13035.

Leidich, S.D. and Orlean, P. (1996). Gpi1, a *Saccharomyces cerevisiae* protein that participates in the first step in glycosylphosphatidylinositol anchor synthesis. J. Biol. Chem. *271*, 27829-27837.

Lisanti, M.P., Field, M.C., Caras, I.W., Menon, A.K., and Rodriguezboulan, E. (1991). Mannosamine, a novel inhibitor of glycosyl-phosphatidylinositol incorporation into proteins. EMBO J. *10*, 1969-1977.

Lohman, K.L., Langer, P.J., and McMahon-Pratt, D. (1990). Molecular cloning and characterization of the immunologically protective surface glycoprotein gp46/m-2 of *Leishmania amazonensis*. Proc. Natl. Acad. Sci. USA *87*, 8393-8397.

Mason, R.W. (1991). Proteinases of mammals: an overview. In Biochemical Protozoology. G.H. Coombs and M.J. North, eds. (London: Taylor & Francis Ltd.), pp. 168-179.

Masterson, W.J., Doering, T.L., Hart, G.W., and Englund, P.T. (1989). A novel pathway for glycan assembly: biosynthesis of the glycosyl-phosphatidylinositol anchor of the trypanosome variant surfce glycoprotien. Cell *56*, 793-800.

Masterson, W.J., Raper, J., Doering, T.L., Hart, G.W., and Englund, P.T. (1990). Fatty acid remodelling: a novel reaction sequence in the biosynthesis of trypanosome glycosyl phosphatidylinositol membrane anchors. Cell *62*, 73-80.

Mauel, J. (1996). Intracellular survival of protozoan parasites with special reference to *Leishmania* spp., *Toxoplasma gondii* and *Trypanosoma cruzi*. Advances in Parasitology *38*, 1-51.

Maxwell, S., Ramalingham, S., Gerber, L.D., Brink, L., and Udenfriend, S. (1995b). An active carbonyl formed during glycosylphosphatidylinositol addition to a protein is evidence of catalysis by a transamidase. J. Biol. Chem. *270*, 19576-19582.

Maxwell, S., Ramalingham, S., Gerber, L.D., and Udenfriend, S. (1995a). Cleavage without anchor addition accompanies the processing of a nascent protein to its glycosylphosphatidylinositol-anchored form. Proc. Natl. Acad. Sci. USA *92*, 1550-1554.

Mayor, S., Menon, A.K., and Cross, G.A.M. (1990b). Glycolipid precursors for the membrane anchor of *Trypanosoma brucei* variant surface glycoproteins II. J. Biol. Chem. *265*, 6174-6181.

Mayor, S., Menon, A.K., and Cross, G.A.M. (1991). Transfer of glycosyl-phosphatidylinositol membrane anchors to polypeptide acceptors in a cell-free system. J. Cell Biol. *114*, 61-71.

Mayor, S., Menon, A.K., Cross, G.A.M., Ferguson, M.A.J., Dwek, R.A., and Rademacher, T.W. (1990a). Glycolipid precursors for the membrane anchor of *Trypanosoma brucei* variant surface glycoproteins I. J. Biol. Chem. *265*, 6164-6173.

Mazhari-Tabrizi, R., Eckert, V., Blank, M., Müller, R., Mumberg, D., Funk, M., and Schwarz, R.T. (1996). Cloning and functional expression of glycosyltransferases from parasitic protozoans by heterologous complementation in yeast: The dolichol phosphate mannose synthase from *Trypanosoma brucei brucei*. Biochem. J. *316*, 853-858.

McConville, M.J. (1997). Glycosylphosphatidylinositols and the surface architecture of parasitic protozoa. In Molecular Biology of Parasitic Protozoa. D.F. Smith and M. Parsons, eds. (pp. 205-228.

McConville, M.J. and Blackwell, J.M. (1991). Developmental changes in the glycosylated phosphatidylinositols of *Leishmania donovani*. Characterization of the promastigote and amastigote glycolipids. J. Biol. Chem. *266*, 15170-15179.

McConville, M.J. and Ferguson, M.A.J. (1993). The structure, biosynthesis and function of glycosylated phosphatidylinositols in the parasitic protozoa and higher eukaryotes. Biochem. J. *294*, 305-324.

McConville, M.J., Turco, S.J., Ferguson, M.A.J., and Sacks, D.L. (1992). Developmental modification of lipophosphoglycan during the differentiation of *Leishmania major* promastigotes to an infectious stage. EMBO J. *11*, 3593-3600.

McGwire, B.S. and Chang, K.P. (1996). Posttranslational regulation of a *Leishmania* HEXXH metalloprotease (gp63) - The effects of site-specific mutagenesis of catalytic, zinc binding, *N*-glycosylation, and glycosyl phosphatidylinositol addition sites on N-terminal end cleavage, intracellular stability, and extracellular exit. J. Biol. Chem. *271*, 7903-7909.

McKerrow, J.H. (1989). Parasite proteases. Exp. Parasitol. *68*, 111-115.

McNeely, T.B. and Turco, S.J. (1987). Inhibition of protein kinase C activity by the *Leishmania donovani* lipophosphoglycan. Biochem. Biophys. Res. Commun. *148*, 653-657.

Medina-Acosta, E. and Cross, G.A.M. (1993). Rapid isolation of DNA from trypanosomatid protozoa using a simple 'mini-prep' procedure. Mol. Biochem. Parasitol. *59*, 327-330.

Medina-Acosta, E., Karess, R.E., and Russell, D.G. (1993). Structurally distinct genes for the surface protease of *Leishmania mexicana* are developmentally regulated. Mol. Biochem. Parasitol. *57*, 31-46.

Medina-Acosta, E., Karess, R.E., Schwartz, H., and Russell, D.G. (1989). The promastigote surface protease (gp63) of *Leishmania* is expressed but differentially processed and localized in the amastigote stage. Mol. Biochem. Parasitol. *37*, 263-273.

Menon, A.K., Eppinger, M., Mayor, S., and Schwarz, R.T. (1993). Phosphatidylethanolamine is the donor of the terminal phosphoethanolamine group in trypanosome glycosylphosphatidylinositols. EMBO J. *12*, 1907-1914.

Mensa-Wilmot, K., Garg, N., McGwire, B.S., Lu, H.G., Zhong, L., Armah, D.A., LeBowitz, J.H., and Chang, K.P. (1999). Roles of free GPIs in amastigotes of *Leishmania*. Mol. Biochem. Parasitol. *99*, 103-116.

Mensa-Wilmot, K., LeBowitz, J.H., Chang, K.-P., Al-Qahtani, A., McGwire, B.S., Tucker, S., and Morris, J.C. (1994). A glycosylphosphatidylinositol (GPI)-negative phenotype produced in *Leishmania major* by GPI phospholipase C from *Trypanosoma brucei*: Topography of two GPI pathways. J. Cell Biol. *124*, 935-947.

Micanovic, R., Bailey, C.A., Brink, L., Gerber, L., Pan, Y.C., Hulmes, J.D., and Udenfriend, S. (1988). Aspartic acid-484 of nascent placental alkaline phosphatase condenses with a phosphatidylinositol glycan to become the carboxyl terminus of the mature enzyme. Proc.Natl.Acad.Sci.U.S.A. *85*, 1398-1402.

Micanovic, R., Gerber, L.D., Berger, J., Kodukula, K., and Udenfriend, S. (1990). Selectivity of the cleavage/attachment site of phosphatidylinositol-glycan anchored membrane proteins determined by site directed mutagenesis at Asp484 of placental alkaline phosphatase. Proc.Natl.Acad.Sci.U.S.A. *87*, 157-161.

Milne, K.G., Field, R.A., Masterson, W.J., Cottaz, S., Brimacombe, J.S., and Ferguson, M.A.J. (1994). Partial purification and characterization of the *N*-acetylglucosaminyl-phosphatidylinositol de-*N*-acetylase of glycosylphosphatidylinositol anchor biosynthesis in African trypanosomes. J. Biol. Chem. *269*, 16403-16408.

Miyata, T., Takeda, J., Iida, Y., Yamada, N., Inoue, N., Takahashi, M., Maeda, K., Kitani, T., and Kinoshita, T. (1993). The cloning of *PIG-A*, a component in the early step of GPI-anchor biosynthesis. Science *259*, 1318-1320.

Mohney, R.P., Knez, J.J., Ravi, L., Sevlever, D., Rosenberry, T.L., Hirose, S., and Medof, M.E. (1994). Glycoinositol phospholipid anchor-defective K562 mutants with biochemical

lesions distinct from those in Thy-1-murine lymphoma mutants. J. Biol. Chem. *269*, 6536-6542.

Molyneux, D.H. (1997). Current public health status of the trypanosomiases and leishmaniases. In Trypanosomiasis and leishmaniasis: Biology and control. G. Hide, J.C. Mottram, G.H. Coombs, and P.H. Holmes, eds. (Oxford, UK: CAB International), pp. 39-50.

Moran, P. and Caras, I.W. (1991a). A nonfunctional sequence converted to a signal for glycophosphatidylinositol membrane anchor attachment. J. Cell Biol. *115*, 329-336.

Moran, P. and Caras, I.W. (1991b). Fusion of sequence elements from non-anchored proteins to generate a fully functional signal for glycophosphatidylinositol membrane anchor attachment. J. Cell Biol. *115*, 1595-1600.

Moran, P. and Caras, I.W. (1992). Proteins containing an uncleaved signal for glycophosphatidylinositol membrane anchor attachment are retained in a post-ER compartment. J. Cell Biol. *119*, 763-772.

Moran, P. and Caras, I.W. (1994). Requirements for glycosylphosphatidylinositol attachment are similar but not identical in mammalian cells and parasitic protozoa. J. Cell Biol. *125*, 333-343.

Moran, P., Raab, H., Kohr, W.J., and Caras, I.W. (1991). Glycophospholipid membrane anchor attachment. J. Biol. Sci. *266*, 1250-1257.

Mottram, J.C., Frame, M.J., Brooks, D.R., Tetley, L., Hutchison, J.E., Souza, A.E., and Coombs, G.H. (1997). The multiple *cpb* cysteine proteinase genes of *Leishmania mexicana* encode isoenzymes that differ in their stage regulation and substrate preferences. J. Biol. Chem. *272*, 14285-14293.

Mottram, J.C., McCready, B.P., Brown, K.P., and Grant, K.M. (1996b). Gene disruptions indicate an essential function for the LmmCRK1 cdc2-related kinase of *Leishmania mexicana*. Mol. Microbiol. *22*, 573-582.

Mottram, J.C., North, M.J., Barry, J.D., and Coombs, G.H. (1989). A cysteine proteinase cDNA from *Trypanosoma brucei* predicts an enzyme with an unusual C-terminal extension. FEBS Lett. *258*, 211-215.

Mottram, J.C., Robertson, C.D., Coombs, G.H., and Barry, J.D. (1992). A developmentally regulated cysteine proteinase gene of *Leishmania mexicana*. Mol. Microbiol. *6*, 1925-1932.

Mottram, J.C., Souza, A.E., Hutchison, J.E., Carter, R., Frame, M.J., and Coombs, G.H. (1996a). Evidence from disruption of the *lmcpb* gene array of *Leishmania mexicana* that cysteine proteinases are virulence factors. Proc. Natl. Acad. Sci. USA *93*, 6008-6013.

Munro, S. and Pelham, H.R. (1987). A C-terminal signal prevents secretion of luminal ER proteins. Cell *48*, 899-907.

Murray, P.J., Spithill, T.W., and Handman, E. (1989). The PSA-2 glycoprotein complex of *Leishmania major* is a glycosylphosphatidylinositol-linked promastigote surface antigen. J.Immunol. *143*, 4221-4226.

Musto, H., Rodriguez-Maseda, H., and Bernardi, G. (1995). Compositional properties of nuclear genes from *Plasmodium falciparum*. Gene *152*, 127-132.

Musto, H., Rodríguez-Maseda, H., and Bernardi, G. (1994). The nuclear genomes of African and American trypanosomes are strikingly different. Gene *141*, 63-69.

Mutomba, M.C., To, W.Y., Hyun, W.C., and Wang, C.C. (1997). Inhibition of proteasome activity blocks cell cycle progression at specific phase boundaries in African trypanosomes. Mol. Biochem. Parasitol. *90* , 491-504.

Nakamura, N., Inoue, N., Watanabe, R., Takahashi, M., Takeda, J., Stevens, V.L., and Kinoshita, T. (1997). Expression cloning of PIG-L, a candidate N-acetylglucosaminyl-phosphatidylinositol deacetylase. J. Biol. Chem. *272*, 15834-15840.

Nobrega, O.T., Silva, M.A., Teixeira, A.R., and Santana, J.M. (1998). Cloning and sequencing of *tccb*, a gene encoding a *Trypanosoma cruzi* cathepsin B-like protease. Mol. Biochem. Parasitol. *97*, 235-240.

North, M.J. (1991). Proteinases of parasitic protozoa: an overview. In Biochemical Protozoology. G.H. Coombs and M.J. North, eds. (London: Taylor & Francis Ltd), pp. 180-185.

North, M.J., Mottram, J.C., and Coombs, G.H. (1990b). Cysteine proteinases of parasitic protozoa. Parasitol.Today *6*, 270-275.

North, M.J., Robertson, C.D., and Coombs, G.H. (1990a). The specificity of trichomonad cysteine proteinases analyzed using fluorogenic substrates and specific inhibitors. Mol. Biochem. Parasitol. *39*, 183-194.

Nuoffer, C., Horvath, A., and Riezman, H. (1993). Analysis of the sequence requirements for glycosylphosphatidylinositol anchoring of *Saccharomyces cerevisiae* Gas1 protein. J. Biol. Chem. *268*, 10558-10563.

Ogata, S., Hayashi, Y., Takami, N., and Ikehara, Y. (1988). Chemical characterization of the membrane-anchoring domain of human placental alkaline phosphatase. J. Biol. Chem. *263*, 10489-10494.

Orlean, P. (1990). Dolichol phosphate mannose synthase is required *in vivo* for glycosyl phosphatidylinositol membrane anchoring, O mannosylation and N glycosyltaion of protein in *Saccharomyces cerevisiae*. Mol.Cell.Biol. *10*, 5796-5805.

Orlean, P., Albright, C., and Robbins, P.W. (1988). Cloning and sequencing of the yeast gene for dolichol phosphate mannose synthase, an essential protein. J. Biol. Chem. *263*, 17499-17507.

Papadopoulou, B. and Dumas, C. (1997). Parameters controlling the rate of gene targeting frequency in the protozoan parasite *Leishmania*. Nucleic Acids Res. *25*, 4278-4286.

Pays, E. and Nolan, D.P. (1998). Expression and function of surface proteins in *Trypanosoma brucei*. Mol.Biochem.Parasitol. *91*, 3-36.

Pelham, H.R., Hardwick, K.G., and Lewis, M.J. (1988). Sorting of soluble ER proteins in yeast. EMBO J. *7*, 1757-1762.

Pelham, H.R.B. (1991). Recycling of proteins between the endoplasmic reticulum and the Golgi complex. Curr.Opin.Cell Biol. *3*, 585-591.

Pimenta, P.F.P., Saraiva, E.M.B., Rowton, E., Modi, G.B., Garraway, L.A., Beverley, S.M., Turco, S.J., and Sacks, D.L. (1994). Evidence that the vectorial competence of phlebotomine sand flies for different species of *Leishmania* is controlled by structural polymorphisms in the surface lipophosphoglycan. Proc. Natl. Acad. Sci. USA *91*, 9155-9159.

Pimenta, P.F.P., Turco, S.J., McConville, M.J., Lawyer, P.G., Perkins, P.V., and Sacks, D.L. (1992). Stage-Specific Adhesion of Leishmania Promastigotes to the Sandfly Midgut. Science *256*, 1812-1815.

Puentes, S.M., da Silva, R.P., Sacks, D.L., Hammer, C.H., and Joiner, K.A. (1990). Serum resistance of metacyclic stage *Leishmania major* promastigotes is due to release of C5b-9. J.Immunol. *145*, 4311-4316.

Ralton, J.E. and McConville, M.J. (1998). Delineation of three pathways of glycosylphosphatidylinositol biosynthesis in *Leishmania mexicana* - Precursors from different pathways are assembled on distinct pools of phosphatidylinositol and undergo fatty acid remodeling. J. Biol. Chem. *273*, 4245-4257.

Ralton, J.E., Milne, K.G., Guther, M.L., Field, R.A., and Ferguson, M.A. (1993). The mechanism of inhibition of glycosylphosphatidylinositol anchor biosynthesis in *Trypanosom brucei* by mannosaminė. J. Biol. Chem. *268*, 24183-24189.

Ramalingam, S., Maxwell, S.E., Medof, M.E., Chen, R., Gerber, L.D., and Udenfriend, S. (1996). COOH-terminal processing of nascent polypeptides by the glycosylphosphatidylinositol transamidase in the presence of hydrazine is governed by the same parameters as glycosylphosphatidylinositol addition. Proc. Natl. Acad. Sci. USA *93*, 7528-7533.

Ranson, N.A., White, H.E., and Saibil, H.R. (1998). Chaperonins. Biochem. J. *333*, 233-242.

Reggiori, F., Canivenc-Gansel, E., and Conzelmann, A. (1997). Lipid remodelling leads to the introduction and exchange of defined ceramides on GPI proteins in the ER and Golgi of *Saccharomyces cerevisiae*. EMBO J. *16*, 3506-3518.

Riezman, H. and Conzelmann, A. (1998). Glycosylphosphatidylinositol:protein transamidase. In Handbook of Proteolytic Enzymes. A.J. Barrett, N.D. Rawlings, and J.F. Woessner, eds. (London: Academic Press), pp. 756-759.

Robertson, C.D. and Coombs, G.H. (1993). Cathepsin B-like cysteine proteases of *Leishmania mexicana*. Mol. Biochem. Parasitol. *62*, 271-280.

Sacks, D.L., Pimenta, P.F.P., McConville, M.J., Schneider, P., and Turco, S.J. (1991). Stage-specific binding of *Leishmania donovani* to the sandfly vector midgut is regulated by conformational changes in the abundant surface lipophosphoglycan. J.Exp.Med. *181*, 685-697.

Sakanari, J.A., Staunton, C.E., Eakin, A.E., Craik, C.S., and McKerrow, J.H. (1989). Serine proteases from nematode and protozoan parasites: isolation of sequence homologs using generic molecular probes. Proc. Natl. Acad. Sci. USA *86*, 4863-4867.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular cloning: A laboratory manual (Cold Spring Harbor, NY.: Cold Spring Laboratory Press).

Sanderson, S.J., Pollock, K.G.J., Hilley, J.D., Mottram, J.C., Meldal, M., Juliano, L., and Coombs, G.H. (1999). Analysis of the multi-step processing of the pro-region during activation of a recombinant cysteine proteinase of the protozoan parasite *Leishmania mexicana*. Submitted

Schneider, P., Ferguson, M.A.J., McConville, M.J., Mehlert, A., Homans, S.W., and Bordier, C. (1990). Structure of the glycosyl-phosphatidylinositol membrane anchor of the *Leishmania major* promastigote surface protease. J. Biol. Chem. *265*, 16955-16964.

Scory, S., Caffrey, C.R., Stierhof, Y.D., Ruppel, A., and Steverding, D. (1999). *Trypanosoma brucei*: killing of bloodstream forms *in vitro* and *in vivo* by the cysteine proteinase inhibitor Z-phe-ala-CHN_2. Exp.Parasitol. *91*, 327-333.

Selzer, P.M., Chen, X.W., Chan, V.J., Cheng, M.S., Kenyon, G.L., Kuntz, I.D., Sakanari, J.A., Cohen, F.E., and McKerrow, J.H. (1997). *Leishmania major*: Molecular modelling of cysteine proteases and prediction of new nonpeptide inhibitors. Exp.Parasitol. *87*, 212-221.

Sharma, D.K., Smith, T.K., Weller, C.T., Crossman, A., Brimacombe, J.S., and Ferguson, M.A. (1999b). Differences between the trypanosomal and human GlcNAc-PI de-*N*-acetylases of glycosylphosphatidylinositol membrane anchor biosynthesis. Glycobiology *9*, 415-422.

Sharma, D.K., Vidugiriene, J., Bangs, J.D., and Menon, A.K. (1999a). A cell-free assay for glycosylphosphatidylinositol anchoring in African trypanosomes. J. Biol. Chem. *275*, 16479-16486.

Sipos, G., Reggiori, F., Vionnet, C., and Conzelmann, A. (1997). Alternative lipid remodelling pathways for glycosylphosphatidylinositol membrane anchors in *Saccharomyces cerevisiae*. EMBO J. *16*, 3494-3505.

Smith, T.K., Milne, F.C., Sharma, D.K., Crossman, A., Brimacombe, J.S., and Ferguson, M.A.J. (1997). Early steps in glycosylphosphatidylinositol biosynthesis in *Leishmania major*. Biochem. J. *326*, 393-400.

Souza, A.E., Bates, P.A., Coombs, G.H., and Mottram, J.C. (1994). Null mutants for the *lmcpa* cysteine proteinase gene in *Leishmania mexicana*. Mol. Biochem. Parasitol. *63*, 213-220.

Souza, A.E., Waugh, S., Coombs, G.H., and Mottram, J.C. (1992). Characterization of a multicopy gene for a major stage-specific cysteine proteinase of *Leishmania mexicana*. FEBS Lett. *311*, 124-127.

Stevens, V.L. (1995). Biosynthesis of glycosylphosphatidylinositol membrane anchors. Biochem. J. *310*, 361-370.

Stevens, V.L. and Raetz, C.R. (1991). Defective glycosyl phosphatidylinositol biosynthesis in extracts of three Thy-1 negative lymphona cell mutants. J. Biol. Chem. *266*, 10039-10042.

Stevens, V.L., Zhang, H., and Harreman, M. (1996). Isolation and characterization of a Chines hamster ovary (CHO) mutant defective in the second step of glycosylphosphatidylinositol biosynthesis. Biochem. J. *313*, 253-258.

Sugiyama, E., DeGasperi, R., Urakaze, M., Chang, H.M., Thomas, L.J., Hyman, R., Warren, C.D., and Yeh, E.T.H. (1991). Identification of defects in glycosylphosphatidylinositol anchor biosynthesis in the Thy-1 expression mutants. J. Biol. Chem. *266*, 12119-12122.

Sutterlin, C., Escribano, M.V., Gerold, P., Maeda, Y., Mazon, M.J., Kinoshita, T., Schwarz, R.T., and Riezman, H. (1998). *Saccharomyces cerevisiae GPI10*, the functional homologue of human *PIG-B*, is required for glycosylphosphatidylinositol-anchor synthesis. Biochem. J. *332*, 153-159.

Symons, F.M., Murray, P.J., Ji, H., Simpson, R.J., Osborn, A.H., Cappai, R., and Handman, E. (1994). Characterization of a polymorphic family of integral membrane proteins in promastigotes of different *Leishmania* species. Mol. Biochem. Parasitol. *67*, 103-113.

Taguchi, R., Asahi, Y., and Ikezawa, H. (1985). Ectoenzyme release from rat liver and kidney by phosphatidylinositol-specific phospholipase C. J.Biochem.(Tokyo) *97*, 911-921.

Takahashi, M., Inoue, N., Ohishi, K., Maeda, Y., Nakamura, N., Endo, Y., Fujita, T., Takeda, J., and Kinoshita, T. (1996). PIG-B, a membrane protein of the endoplasmic reticulum with a large lumenal domain, is involved in transferring the third mannose of the GPI anchor. EMBO J. *15*, 4254-4261.

Takeda, O., Miura, Y., Mitta, M., Matsushita, H., Kato, I., Abe, Y., Yokosawa, H., and Ishii, S.-I. (1994). Isolation and analysis of cDNA encoding a precursor of *Canavalia ensiformis* asparaginyl endopeptidase (Legumain). J.Biochem.(Tokyo) *116*, 541-546.

Takeda, J., Miyata, T., Kawagoe, K., Iida, Y., Endo, Y., Fujita, T., Takahashi, M., Kitani, T., and Kinoshita, T. (1993). Deficiency of the GPI anchor caused by a somatic mutation of the *PIG-A* gene in paroxysmal nocturnal hemoglobinuria. Cell *73*, 703-711.

Tiede, A., Schubert, J., Nischan, C., Jensen, I., Westfall, B., Taron, C.H., Orlean, P., and Schmidt, R.E. (1998). Human and mouse Gpi1p homologues restore glycosylphosphatidylinositol membrane anchor biosynthesis in yeast mutants. Biochem. J. *334*, 609-616.

Tobin, J.F., Laban, A., and Wirth, D.F. (1991). Homologous recombination in *Leishmania enriettii*. Proc. Natl. Acad. Sci. USA *88*, 864-868.

Tobin, J.F. and Wirth, D.F. (1992). A sequence insertion targeting vector for *Leishmania enriettii*. J. Biol. Chem. *267*, 4752-4758.

Urakaze, M., Kamitani, T., DeGasperi, R., Sugiyama, E., Chang, H.M., Warren, C.D., and Yeh, E.T. (1992). Identification of a missing link in glycosylphosphatidylinositol anchor biosynthesis in mammalian cells. J. Biol. Chem. *267*, 6459-6462.

Vanhamme, L. and Pays, E. (1995). Control of gene expression in trypanosomes. Microbiol.Rev. *59*, 223-240.

Vidugiriene, J. and Menon, A.K. (1993). Early lipid intermediates in glycosyl-phosphatidylinositol anchor assembly are synthesized in the ER and located in the cytoplasmic leaflet of the ER membrane bilayer. J. Cell Biol. *121*, 987-996.

Vidugiriene, J. and Menon, A.K. (1994). The GPI anchor of cell-surface proteins is synthesized on the cytoplasmic face of the endoplasmic reticulum. J. Cell Biol. *127*, 333-341.

Vidugiriene, J. and Menon, A.K. (1995). Soluble constituents of the ER lumen are required for GPI anchoring of a model protein. EMBO J. *14*, 4686-4694.

Vossen, J.H., Ram, A.F., and Klis, F.M. (1995). Identification of SPT14/CWH6 as the yeast homologue of hPIG-A, a gene involved in the biosynthesis of GPI anchors. Biochim. Biophys. Acta *1243*, 549-551.

Voth, B.R., Kelly, B.L., Joshi, P.B., Ivens, A.C., and McMaster, W.R. (1998). Differentially expressed *Leishmania major gp63* genes encode cell surface leishmanolysin with distinct signals for glycosylphosphatidylinosilol attachment. Mol. Biochem. Parasitol. *93*, 31-41.

Wainwright, L.J. and Field, M.C. (1997). Quality control of glycosylphosphatidylinositol anchor attachment in mammalian cells: a biochemical study. Biochem. J. *321*, 655-664.

Walter, P., Gilmore, R., and Blobel, G. (1984). Protein translocation across the endoplasmic reticulum. Cell *38*, 5-8.

Watanabe, R., Inoue, N., Westfall, B., Taron, C.H., Orlean, P., Takeda, J., and Kinoshita, T. (1998). The first step of glycosylphosphatidylinositol biosynthesis is mediated by a complex of PIG-A, PIG-H, PIG-C and GPI1. EMBO J. *17*, 877-885.

Watanabe, R., Kinoshita, T., Masaki, R., Yamamoto, A., Takeda, J., and Inoue, N. (1996). PIG-A and PIG-H, which participate in glycosylphosphatidylinositol anchor biosynthesis form a protein complex in the endoplasmic reticulum. J. Biol. Chem. *271*, 26868-26875.

Watanabe, R., Ohishi, K., Maeda, Y., Nakamura, N., and Kinoshita, T. (1999). Mammalian PIG-L and its yeast homologue Gpi12p are *N*-acetylglucosaminylphosphatidylinositol de-N-acetylases essential in glycosylphosphatidylinositol biosynthesis. Biochem. J. *339*, 185-192.

Werbovetz, K.A. and Englund, P.T. (1996). Lipid metabolism in *Trypanosoma brucei*: utilization of myristate and myristoyllysophosphatidylcholine for myristoylation of glycosyl phosphatidylinositols. Biochem. J. *318*, 575-581.

Wilbourn, B., Nesbeth, D.N., Wainwright, L.J., and Field, M.C. (1998). Proteasome and thiol involvement in quality control of glycosylphosphatidylinositol anchor addition. Biochem. J. *332*, 111-118.

Winter, G., Fuchs, M., McConville, M.J., Stierhof, Y.-D., and Overath, P. (1994). Surface antigens of *Leishmania mexicana* amastigotes: Characterization of glycoinositol phospholipids and a macrophage- derived glycosphingolipid. J. Cell Sci. *107*, 2471-2482.

Yu, J.L., Nagarajan, S., Knez, J.J., Udenfriend, S., Chen, R., and Medof, M.E. (1997). The affected gene underlying the class K glycosylphosphatidylinositol (GPI) surface protein defect codes for the GPI transamidase. Proc. Natl. Acad. Sci. USA *94*, 12580-12585.

APPENDIX

Cell Line	Description	WCMP Designation
P1	*GPI8* heterozygote (puromycin-resistant) clone 1	1139/1238
P2	*GPI8* heterozygote (puromycin-resistant) clone 2	1140
P3	*GPI8* heterozygote (puromycin-resistant) clone 3	1141
S1	*GPI8* heterozygote (nourseothricin-resistant) clone 1	1142
P1S1	*GPI8* null mutant clone 1	1234
P1S3	*GPI8* null mutant clone 2	1236
P1S1[pGL269]	*GPI8* null mutant clone 1 re-expressing *GPI8* from an episome	1235
P1S1[pGL403]	*GPI8* null mutant clone 1 expressing C216G mutant *GPI8* from an episome	1237
P1S3[pGL269]	*GPI8* null mutant clone 2 re-expressing *GPI8* from an episome	TBA
P1S1[pGL190]	*GPI8* null mutant clone 1 expressing GPI8-GFP fusion protein from an episome	TBA

TBA = To be assigned