LIPOIC ACID METABOLISM IN *Leishmania major*

RYAN EUGENE BISSETT BSC

SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

MAY 2009

UNIVERSITY OF GLASGOW

INSTITUTE OF BIOMEDICAL LIFE SCIENCES

DIVISION OF INFECTION AND IMMUNITY

Abstract

Protozoan parasites of the genus *Leishmania* are the causative agents of a complex of diseases referred to as leishmaniasis. *Leishmania* have a digenetic life cycle that involves a sand fly vector (promastigote stage) and a mammalian host (amastigote stage). The parasites reside within very different environmental niches in the two different hosts, and therefore must be able to adapt their energy metabolism to the available carbon and nitrogen sources.

Lipoic acid (LA) is a multifaceted molecule, and plays an important role as a water- and fat-soluble antioxidant. LA is also an essential cofactor of the α-ketoacid dehydrogenase complexes (α-KADHs) and of the glycine cleavage complex (GCC). The α-KADHs include the pyruvate dehydrogenase (PDH), branched-chain α-ketoacid dehydrogenase (BCKDH) and α-ketoglutarate dehydrogenase (α-KGDH), each of which is integral to cellular energy metabolism. In some organisms, LA can be acquired through salvage and biosynthesis pathways, and yet others only encode enzymes that permit one of the two pathways. Lipoylation of the PDH has been demonstrated in a parasite related to *Leishmania* called *Trypanosoma brucei*; however there have not been any investigations into the enzymes involved in LA metabolism in either *Leishmania* or *Trypanosoma brucei*.

In silico analyses identified genes encoding for proteins involved in both LA biosynthesis and salvage (lipoic acid synthase (LIPA), octanoyl-[acyl carrier protein]: protein *N*-octanoyltransferase (LIPB) and lipoate protein ligase (LPLA), respectively), and it was predicted that all three proteins possess mitochondrial targeting peptides. Targeting of these proteins to the mitochondrion was verified by a green fluorescence protein (GFP) reporter system, and by subcellular pre-fractionation using digitonin followed by western blotting. Functionality of *L. major* putative *LIPA*, *LIPB* and *LPLA* genes was determined by showing that the genes complemented the no-growth phenotype of bacteria deficient in either *lipA* or *lipB* genes on minimal medium.

Bioinformatics analyses also showed that *L. major* possesses genes encoding all of the subunits comprising the different α-KADHs and the GCC, and the subunits were predicted to possess mitochondrial targeting peptides. Western blotting of promastigote protein with an antibody recognising protein-bound LA (α-LA

antibody) identified four proteins, which based upon predicted molecular sizes, correspond to the lipoylated transacylase subunits of the three α-KADHs and the H-protein of the GCC. Interestingly, the lipoylation pattern changes throughout promastigote growth *in vitro*, with α-KGDH being lipoylated throughout promastigote life while PDH and BCKDH are not lipoylated and presumably not active in metacyclic promastigotes. These findings indicate that modification of α-KADHs and the GCC by lipoylation is a dynamic process, possibly reflecting adaptations in the parasite's energy metabolism during their developmental cycle.

Three approaches were taken to study the relative importance of the LA biosynthesis and salvage pathways in *L. major* promastigotes. First, LA analogues 8' bromooctanoic acid (8-BOA) and octanoic acid (OA) were tested for their effects on growth in *L. major* maintained in lipid-depleted medium. The IC_{50} for 8-BOA was relatively high when compared to that determined in other organisms, suggesting that LA biosynthesis can compensate for a decrease in LA salvage in medium deficient in LA. Second, attempts to replace either *LIPA* or *LPLA* genes with selectable markers were unsuccessful. *LPLA* could however, be knocked-out when an extra copy of the gene was introduced into the parasite's genome. These data suggest that both LA acquisition pathways might be essential for promastigote growth and development. Third, overexpression of *C*-terminal His-tagged versions of LIPB (LIPB-His), LPLA (LPLA-His) and a LPLA active site mutant, $LPLA^{H118A}$ ($LPLA^{H118A}$-His), resulted in slow-growth phenotypes. Overexpression of LIPB-His and $LPLA^{H118A}$-His resulted in lipoylation of the PDH and BCKDH in metacyclic promastigotes, which is not observed in wild-type metacyclic promastigotes. It is hypothesised that LA biosynthesis and salvage enzymes could have differential substrate-specificities in *L. major*.

A number of avenues require further investigation, including the mechanism that permits a relatively rapid turnover of lipoylated protein, and whether lipoylation patterns differ depending upon the carbon sources that are provided in the growth medium. Also, it will be interesting to determine whether LIPB and LPLA have intrinsic substrate-specificities, and whether this is sufficient to explain the fact that both LIPA and LPLA are essential in the promastigote stage *in vitro*.

Table of Contents

List of Tables

List of Figures

Acknowledgements

Firstly, I would like to say a big thank you to my supervisors, Prof. Sylke Müller and Prof. Graham H. Coombs. Sylke, I very much appreciate your support and guidance throughout, and I have learned some very valuable lessons under your supervision, for which I am truly grateful. Graham, I have benefited from your focus and objectivity.

I'd like to thank all members of the Müller Lab, past and present, for providing a great working environment: Paul McMillan, Lynsey Wallace, Svenja Günther, Janet Storm, Lesley McCaig and Eva Patzewitz. Also, I would like to give a special 'dank je' to Janet Storm; you have provided me with invaluable advice and guidance, and I will undoubtedly miss our 'metabolism chats'!

A big thank you goes out to the L5 & 6 'Kinetoplastida crew'. From lab meetings to 'TCR chat', I have both enjoyed and benefitted massively from the discussions that we have had. In particular, thank you to Jim Hilley and Elmarie Myburgh for training me in FACS analysis, to Lesley Morrison for her WT[uncut pGL1137] line and to Denise Candlish for providing me with purified *L. major* amastigotes. I would like to thank Rod Williams for allowing me to use the α-cysteine synthase antibody that he created. Also, I would like to thank my assessor, Prof. Jeremy Mottram, for his sound advice, especially with regards to Southern blot analysis. Nick Bland; you've been a powerful mate, and I appreciate all of your technical advice and for helping me to get through the troughs, thank you! I'd like to thank Aidan, Cherie and Kate in Glasgow, for their support and for the good times. I would like to give a special mention to my best mate Wes Pancho; you've been an absolute rock and I hope this thesis DHV somewhat!

A ginormous thank you to Lindsey! You're a wee gem, and thank you so much for the years of patience, understanding and encouragement which has helped me get this far!

Finally, from sport to studies, from tennis in London to *Drosophila* in Lyon and *Leishmania* in Glasgow, my parents Maria and Eugene and my sister Lauren have always been there for me. Thank you guys for your love and support; you're the best!!!

Author's Declaration

The research reported in this thesis is the result of my own original work, with the following exceptions:

Chapter 3

- LC-tandem mass spectrometry analysis was performed by Dr Richard Burchmore of the Sir Henry Wellcome Functional Genomics Facility, University of Glasgow.

- Isolation of *L. major* amastigotes from mouse back lesions was carried out by Mrs Denise Candlish of the Division of Infection and Immunity, University of Glasgow.

Chapters 4 and 5

- The line WT[uncut pGL1137] was created by Dr Lesley Morrison of the Wellcome Centre for Molecular Parasitology, University of Glasgow.

Ryan Eugene Bissett

May 2009

List of Abbreviations

ACC	acetyl-CoA carboxylase
ACP	acyl carrier protein
ACSM1	medium-chain acetyl-CoA synthetase 1
ADP	adenosine diphosphate
ALAT	alanine aminotransferase
Amp	ampicillin-resistance gene
Apo-E2	non-lipoylated E2 subunit
Apo-H-protein	non-lipoylated H-protein
APS	ammonium persulphate
ASCT	acetate: succinate CoA transferase
ATc	anhydrotetracycline
ATP	adenosine triphosphate
BCAA	branched-chain amino acid
BCAT	branched-chain aminotransferase
BCKDH	branched-chain α-ketoacid dehydrogenase
BCKDK	BCKDH kinase
BCKDP	BCKDH phosphatase
BioB	biotin synthase
8-BOA	8'bromooctanoic acid
bp	base pairs
BPL	biotin protein ligase
BSA	bovine serum albumin
BSF	bloodstream form
Bt	*Bos Taurus*
℃	degree Celsius
5,10-CH_2-THF	N_5,N_{10}-methylene tetrahydrofolate
CL	cutaneous leishmaniasis
CO_2	carbon dioxide
CoA	coenzyme A
CPB 2.8	cysteine peptidase B 2.8
C-terminus	carboxy terminus of a protein
CV	column-volume
CS	cysteine synthase
Da	Daltons
DA	decanoic acid
DAPI	4' -6-diamidino-2-phenylindole dihydrochloride
DCL	diffuse cutaneous leishmaniasis
DHLA	dihydrolipoic acid
DHTDK1	dehydrogenase E1 and transketolase domain-containing 1
DIC	differential interference contrast
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DTT	dithiothreitol
E1	α-ketoacid decarboxylase subunit of α-KADH
E2	acyltransferase subunit of α-KADH
E3	dihydrolipoamide dehydrogenase subunit of α- KADH
E1b	E1 subunit of the BCKDH
E1k	E1 subunit of the α-KGDH
E1p	E1 subunit of the PDH

E2b	E2 subunit of the BCKDH
E2k	E2 subunit of the α-KGDH
E2p	E2 subunit of the PDH
E2p-L	E2p-like subunit
E3BP	E3 binding protein
Ec	*E. coli*
ECL	enhanced chemiluminescence
EDTA	ethylenediamine tetraacetic acid
ELO	elongase
EPB	electroporation buffer
ER	endoplasmic reticulum
f	femto
FA	fatty acid
FACS	fluorescence activated cell sorting
FAD^+	flavin adenine dinucleotide oxidised form
$FADH_2$	flavin adenine dinucleotide reduced form
FAS	fatty acid synthase
FBP	fructose 1,6-bisphosphatase
FCS	foetal calf serum
[Fe-S]	iron-sulphur cluster
FL	full-length version of gene/protein
fPPG	filamentous proteophosphoglycan
5' FR	five prime flanking region of a gene
3' FR	three prime flanking region of a gene
GCC	glycine cleavage complex
gDNA	genomic DNA
GFP	green-fluorescent protein
GSH	glutathione
GSPS	glutathionyl-spermidine synthetase
GSSG	glutathione disulphide
GTP	guanosine triphosphate
H-protein	carrier protein subunit of the GCC
Hs	*Homo sapiens*
HXK	hexokinase
HYG	hygromycin B phosphotransferase
IC_{50}	concentration of compound that inhibits 50 % of growth/activity
ICL	isocitrate lyase
IgG	immunoglobulin G
IPTG	isopropyl-β-D-thiogalactopyranoside
K	kilo
α-KADHs	α-KADH complexes
Kan	kanamycin-resistance gene
kDNA	kinetoplast DNA
α-KGDH	α-ketoglutarate dehydrogenase complex
kpsi	kilo pound-force per square inch
l	litre
L1	first lipoyl-domain
L2	second lipoyl-domain
LA	lipoic acid
LB	Luria-Bertani
Lb	*Leishmania braziliensis*
Li	*Leishmania infantum*
LIPA	lipoic acid synthase

LIPB	octanoyl-[acyl carrier protein]: protein *N*-octanoyltransferase
LipDH	dihydrolipoamide dehydrogenase subunit of α-KADHs and the GCC
Lmex	*Leishmania mexicana*
Lmj	*Leishmania major*
LPG	lipophosphoglycan
LPLA	lipoate protein ligase
L-protein	dihydrolipoamide dehydrogenase subunit of GCC
LT	lipoyltransferase
µ	micro
M	Molar
m	milli
MCL	mucocutaneous leishmaniasis
MCS	multiple cloning site
min	minutes
MLS	malate synthase
MOPS	3-(N-morpholino) propanesulphonic acid
5 % MPBS	5 % (w/v) milk in PBS
2 % MPBST	2 % (w/v) MPBS with 0.1 % (v/v) Tween-20
mRNA	messenger RNA
Ms	*Mus musculus*
Mt	*Mycobacterium tuberculosis*
n	nano
NAD^+	nicotinamide adenine dinucleotide oxidised form
NADH	nicotinamide adenine dinucleotide reduced form
NEO	neomycin phosphotransferase
N-terminus	amino terminus of a protein
O_2	oxygen
OA	octanoic acid
OD	optical density
ORF	open reading frame
p	pico
PAC	puromycin *N*-acetyltransferase
PBS	phosphate buffered saline
PBST	PBS containing 0.1 % (v/v) Tween-20
PCF	procyclic form
PCR	polymerase chain reaction
PDH	pyruvate dehydrogenase complex
PDK	PDH kinase
PDP	PDH phosphatase
PDPc	PDH phosphatase catalytic subunit
PDPr	PDH phosphatase regulatory subunit
PFK	phosphofructokinase
PG	phosphoglycan
PGK	phosphoglycerate kinase
PHK	protein histidine kinase
PLP	pyridoxal 5'-phosphate
PMN	polymorphonuclear neutrophil granulocytes
PMSF	phenylmethylsulfonylfluoride
PPG	proteophosphoglycan
PPi	pyrophosphate
PPP	pentose phosphate pathway
P-protein	decarboxylase subunit of the GCC
Ps	*Pisum satisvum*

PSG	promastigote secretory gel
PUFA	polyunsaturated fatty acid
PV	parasitophorous vacuole
Rf	relative mobility of proteins in SDS-PAGE
ROS	reactive oxygen species
Rn	*Rattus norvegicus*
RNA	ribonucleic acid
RNAi	RNA interference
Rpm	revolutions per minute
rRNA	ribosomal RNA
SAM	S-adenosyl-methionine
SAS	splice acceptor site
SAT	streptothricin-acetyl-transferase
SbV	pentavalent antimonials
SCS	succinyl-CoA synthetase
SD	standard deviation
SDH	succinate dehydrogenase
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
SeLA	selenolipoic acid
SHMT	serine hydroxymethyltransferase
SP	signal peptide
T1	truncated version 1 of gene/protein
T2	truncated version 2 of gene/protein
T3	truncated version 3 of gene/protein
Ta	*Thermoplasma acidophilum*
TAE	Tris-acetate-containing EDTA
Tb	*Trypanosoma brucei*
TCA	tricarboxylic acid
Tc	*Trypanosoma cruzi*
TEMED	tetramethylethylenediamine
THF	tetrahydrofolate
TPP	thiamine pyrophosphate
T-protein	aminomethyltransferase subunit of the GCC
Tris	Tris [hydroxmethyl] aminomethane
tRNA	transfer RNA
U	unit
UV	ultraviolet
V	Volt
VL	visceral leishmaniasis
VSG	variable surface glycoprotein
v/v	volume per volume
WHO	world health organisation
WT	wild-type
w/v	weight per volume

1 Introduction

1.1 *Leishmania*

1.1.1 Aetiology

The leishmaniases describe a spectrum of zoonotic diseases that are caused by obligate intracellular protozoa of the genus *Leishmania* (family Trypanosomatidae, order Kinetoplastida), which are transmitted to a vertebrate host through the bite of an infected sandfly (Herwaldt, 1999). Leishmaniasis is a complex of diseases, which are classified into four clinical forms:

i. Cutaneous leishmaniasis (CL) is the most common form of leishmaniasis, and results in cutaneous sores, which are localised to the vector bite site. It is caused mainly by *Leishmania major*, *Leishmania tropica* and *Leishmania aethiopica* in the Old World, and by *Leishmania mexicana* complex species, *Leishmania guyanensis* and *Leishmania panamensis* in the New World. The lesions caused by infection with these species generally self-heal within a few months after the onset of symptoms, without the need for drug intervention (Desjeux, 2004).

ii. Diffuse cutaneous leishmaniasis (DCL) is a more extensive and chronic form of CL; the sores that form are non-ulcerative (Herwaldt, 1999) and do not self-heal, due to an inefficient cell-mediated immune response (Desjeux, 2004). It is caused by *L. aethiopica* in the Old World and by *L. mexicana* complex species (most notably, by *Leishmania amazonensis*) in the New World.

iii. Mucocutaneous leishmaniasis (MCL) is another type of CL defined by the development of sores on the naso-oropharyngeal mucosal membranes, which result in severe disfigurement (Herwaldt, 1999). It is caused by *Leishmania braziliensis* subspecies.

iv. Visceral leishmaniasis (VL), commonly referred to as kala azar, is the most severe form of the disease; symptoms include fever, extreme weight loss, anaemia, hepatosplenomegaly, pancytopenia and hypergammaglobulinaemia. The primary species responsible for VL in the Old Word is *Leishmania donovani*,

particularly on the Indian subcontinent and Eastern Africa. In the Mediterranean and the New World, *Leishmania infantum* is the principal cause of VL.

Humans are susceptible to infection from approximately 20 species of *Leishmania*, and leishmaniasis is endemic to 88 countries on four continents (http://www.who.int/leishmaniasis/burden/en/), as illustrated in Figure 1.1.

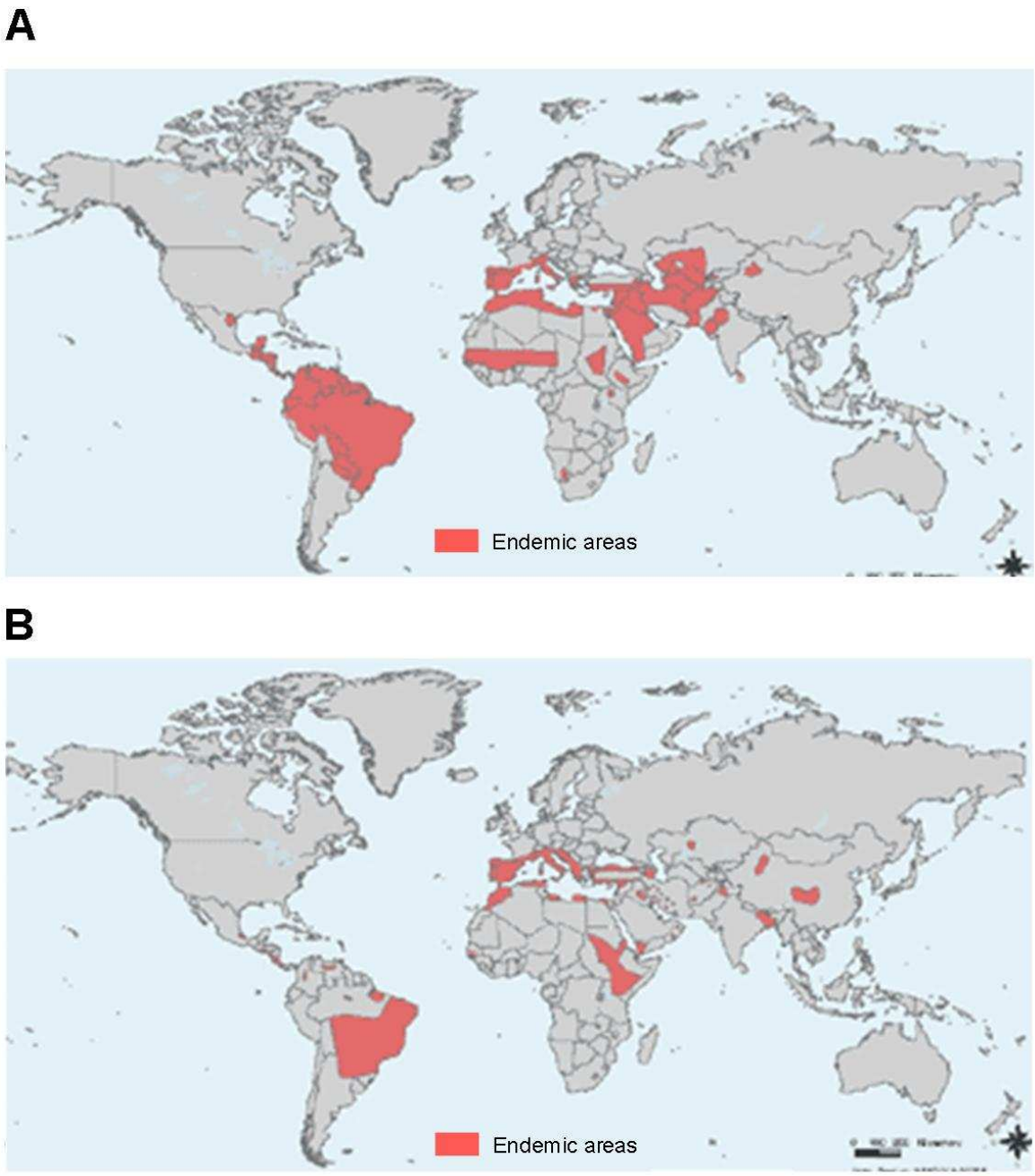


Figure 1.1 Global distribution of cutaneous and visceral leishmaniasis
Maps showing the global distribution of cutaneous (*A*) and visceral (*B*) leishmaniasis. Maps were taken from http://www.who.int/leishmaniasis/leishmaniasis_maps/en/index.html.

1.1.2 Life cycle

Leishmania require two hosts to complete their life cycle; the definitive female, haematophagous sandfly vector (subfamily *Phlebotominae*) and the mammalian host (including dogs and humans).

When a sandfly bites an infected mammal it acquires a bloodmeal containing *Leishmania* amastigotes. In order to survive within the sandfly midgut, these amastigotes must differentiate into the promastigote form. Species of the subgenus *Leishmania* are suprapylarian parasites, restricted to the midgut of the sandfly, whereas species belonging to the *Viannia* subgenus from the New World are peripylarian parasites, entering the hindgut before migrating towards the midgut (Kamhawi, 2006). Studies of *L. mexicana* in its sandfly host *Lutzomyia longipalpis* have illustrated the existence of at least six different promastigote sub-types (Rogers *et al.*, 2002). Firstly, amastigotes ingested within the bloodmeal differentiate into procyclic promastigotes, which then develop into nectomonad promastigotes, then into leptomonad promastigotes and finally into metacyclic promastigotes, which are the mammal-infective form (see Figure 1.2). The two other promastigote sub-types observed are haptomonad promastigotes and paramastigotes, yet the precursors of these have not been determined (Kamhawi, 2006; Rogers *et al.*, 2002). Of the six promastigote sub-types, only the procyclic promastigotes (in the abdominal midgut) and leptomonad promastigotes (in the thoracic midgut and foregut) are replicative (Gossage *et al.*, 2003).

In the sandfly vector, the time taken for ingested amastigotes to develop into metacyclic promastigotes ranges from six to nine days, depending upon the species of *Leishmania* (Kamhawi, 2006). The mode of transmission of metacyclic promastigotes into the mammalian host is via regurgitation. The cause of this regurgitation is the promastigote secretory gel (PSG), which is secreted by the parasites into the thoracic midgut to form a PSG plug, which in turn forces open the stomodeal valve where metacyclic promastigotes accumulate (Rogers *et al.*, 2002; Stierhof *et al.*, 1999) (see Figure 1.2). Upon regurgitation by the sandfly, approximately 1000 metacyclic promastigotes are inoculated into the mammalian host, most of which were derived from just behind the stomodeal valve. Interestingly, the PSG plug is also transmitted into the mammalian host, and the main component of this plug, filamentous proteophosphoglycan (fPPG), has been shown to aid disease progression (Rogers *et al.*, 2002; Stierhof *et al.*, 1999).

A lot of focus has recently been on the events that proceed as metacyclic promastigotes enter the mammalian host. These experiments have illustrated that even though amastigote development is restricted to the phagolysosomal compartment of macrophages, polymorphonuclear neutrophil granulocytes (PMN) appear to be key in establishing infection (Nauseef, 2007; Peters *et al.*, 2008; Segal, 2005; van Zandbergen *et al.*, 2004). An interesting theory has been proposed, whereby PMNs act as a 'Trojan horse' vehicle for *Leishmania* infection of macrophages. It was demonstrated that metacyclic promastigotes are phagocytosed by PMNs, and that PMNs do not kill the internalised parasites by the oxidative burst (Laufs *et al.*, 2002; Muller *et al.*, 2001; van Zandbergen *et al.*, 2004). Additionally, the phagocytosed metacyclic promastigotes do not multiply or differentiate into amastigotes inside PMNs (Laufs *et al.*, 2002; Muller *et al.*, 2001; van Zandbergen *et al.*, 2004). The 'Trojan horse' theory was proposed because it was observed *in vitro* that infected PMNs are phagocytosed by macrophages (van Zandbergen *et al.*, 2004). However, recent *in situ* mouse data show that PMNs undergoing apoptosis in fact release their parasite load, which is then in turn phagocytosed by macrophages (Peters *et al.*, 2008). Thus, the current hypothesis is concurrent with the idea that PMNs are essential to establishment of *Leishmania* infection, however the results do not indicate that PMNs act as 'Trojan horses'; instead, parasites exposed to the internal milieu of PMNs are conferred subsequent resistance somehow after being internalised by macrophages (Peters *et al.*, 2008).

After being phagocytosed by macrophages, *Leishmania* metacyclic promastigotes reside within the acidic phagolysosomal compartment (or parasitophorous vacuole (PV)), and prevent the fusion of the phagolysosome with the lysosome (Bogdan & Rollinghoff, 1999). It is in the PV that metacyclic promastigotes differentiate into the amastigote form, and the nature of the PV niche varies, depending upon the species of *Leishmania* (Handman & Bullen, 2002; Rittig & Bogdan, 2000). Amastigotes are aflagellate and round-oval in shape with a diameter of 4 μm (Herwaldt, 1999) (see Figure 1.2). Amastigotes multiply within the PV and are released and subsequently infect further macrophages, although the exact mechanism of amastigote escape from macrophages is not fully understood.

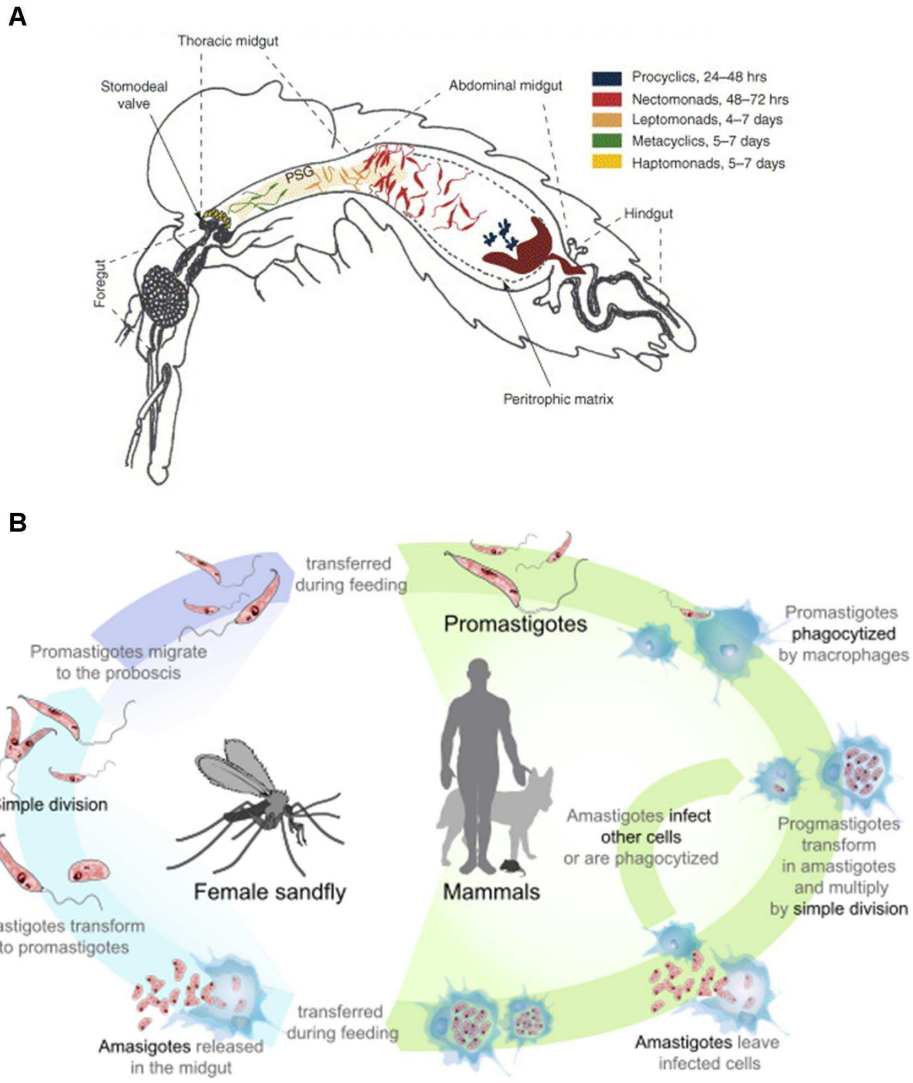


Figure 1.2 Digenetic life cycle of *Leishmania*
A, Schematic representation of a *Leishmania*-infected sandfly, indicating time-dependent appearance of distinct morphological forms. The promastigote secretory gel (PSG) fills the thoracic midgut, and in doing so holds open the stomodeal valve. (Image taken from (Kamhawi, 2006) with permission from Elsevier.) *B,* The life cycle of *Leishmania* is initiated by the uptake of amastigotes in a bloodmeal from the mammalian host by the sandfly vector. Amastigotes differentiate into promastigotes, which migrate to the proboscis, where the promastigotes are inoculated into the mammalian host upon taking another bloodmeal. The promastigotes are ultimately phagocytosed by macrophages, where they multiply within the phagolysosomal compartment. Amastigotes are released and then the free amastigotes are taken up by another macrophage. (Image taken from http://commons.wikimedia.org/wiki/File:Leishmaniasis_life_cycle_diagram_en.svg.)

1.1.3 Ultrastructure of *Leishmania*

Leishmania belong to the order Kinetoplastida, the members of which each possess a unique organelle called the kinetoplast, comprising the mitochondrial DNA. Similar to *Leishmania*, other kinetoplastids such as species within the genus *Trypanosoma*, have complex digenetic life cycles. Kinetoplastid species have adapted in various ways to the very different environments found within the fly and mammalian hosts. *Trypanosoma brucei* for example, has evolved to derive its

nutrition from mammalian blood, and as such bears a flagellum and has adapted to the problem of immune recognition by varying the antigen type (variable surface glycoprotein (VSG)) presented on its plasma membrane (Donelson, 2003). *Leishmania* however, have evolved to survive and replicate within the acidic environment of the macrophage phagolysosome. Amastigotes are small in length (4 μm) and amotile (aflagellate). By contrast, promastigotes have long thin cell bodies that are 5-20 μm in length, and possess a flagellum up to 20 μm in length. As such, a large amount of cellular remodelling is required in the process of promastigote-to-amastigote differentiation (Besteiro *et al.*, 2006; Clayton *et al.*, 1995) (see Figure 1.3). This involves rearrangement of the array of cross-linked subpellicular microtubules that are located underneath the plasma membrane (McConville *et al.*, 2002), which govern the characteristic shapes of *Leishmania* promastigotes and amastigotes.

1.1.3.1Plasma membrane

Leishmania synthesise a family of glycoconjugates that are either attached to the extracellular face of the plasma membrane or secreted from the cell. The base molecule of these glycolipids is phosphoglycan (PG), and important members of this family include membrane-bound lipophosphoglycan (LPG) and proteophosphoglycan (PPG) (Beverley & Turco, 1998; Ilg *et al.*, 1999; McConville & Ferguson, 1993). These molecules make up a thick glycocalyx in procyclic promastigotes, however metacyclic promastigotes and amastigotes possess a much thinner layer. It was demonstrated that LPG is essential for survival of *L. major* within the sandfly (Sacks *et al.*, 2000) and for infection of macrophages and mice (Spath *et al.*, 2000). Interestingly, LPG does not seem to be essential for *L. mexicana* infection of macrophages or mice (Ilg *et al.*, 2001). It is not apparent whether the contrasting results observed for *L. major* and *L. mexicana* are intrinsic to using mouse as a model, or whether the two species indeed do use different surface glycoconjugates to survive within the mammalian host (Turco *et al.*, 2001).

1.1.3.2Flagellar pocket

In kinetoplastids, endocytosis and exocytosis are restricted to a single portion of the plasma membrane at the anterior of the cell called the flagellar pocket. The flagellar pocket is a specialised invagination of the plasma membrane, and is the

only region in the cell that is not subtended by microtubules which make up the cytoskeletal network (Clayton *et al.*, 1995; Overath *et al.*, 1997) (see Figure 1.3).

1.1.3.3Flagellum

Leishmania promastigotes possess a single flagellum, which is derived from one of the two basal bodies located within the kinetoplast (see Figure 1.3). The flagellum consists of a classical 9+2 microtubule double axoneme and the trypanosomatid-specific paraflagellar rod (Bastin *et al.*, 2000) (see Figure 1.3).

In *Leishmania* promastigotes, the flagellum has several roles: the first and most obvious is to enable promastigotes to migrate from the sandfly midgut towards the stomodeal valve. *L. mexicana* promastigotes were shown to swim at a velocity of 29.5 ± 5.9 μm sec^{-1} in culture conditions (Santrich *et al.*, 1997). The flagellum length varies between different promastigote sub-types, and metacyclics possess the greatest flagellum-to-cell body ratio and are highly motile. It has been shown that *Leishmania* are able to use positive chemotaxic cues including sugar and saliva (Barros *et al.*, 2006; Charlab *et al.*, 1995). The flagellum has also been shown to be involved in anchoring promastigotes to the sandfly gut epithelium (Wakid & Bates, 2004). Interestingly, the promastigote plasma membrane glycolipid, LPG, is important for interactions of promastigotes with the gut in certain sandfly species, but not others. It was discovered that LPG is essential for the attachment of *L. major* promastigotes in the specific vector *Phlebotomus papatasi*, however LPG was not essential for adherence of promastigotes in the permissive vectors *Phlebotomus arabicus* or *Lutzomyia longipalpis* (Myskova *et al.*, 2007).

1.1.3.4Kinetoplast

The kinetoplast is a disc-shaped sub-structure located within the mitochondrion, which contains kinetoplast DNA (kDNA). kDNA accounts for 10-20 % of the total cellular DNA content, and can be observed by microscopy using DAPI stain (Bastin *et al.*, 2000). kDNA is found in the form of a few dozen maxicircles and thousand of minicircles. The maxicircles have some functions of higher eukaryote mitochondrial DNA. The minicircles encode guide RNAs, which carry out the post-transcriptional editing of maxicircle RNA species by extensive uridylate insertions (Liu *et al.*, 2005).

1.1.3.5Mitochondrion

Leishmania possess a single mitochondrion, which is large and stretches throughout most of the cell body (Bienen *et al.*, 1981; Coombs *et al.*, 1986). In terms of structure, the mitochondrion in *Leishmania* is similar to that found in other eukaryotes; it consists of outer and inner membranes and tubular cristae. The *Leishmania* mitochondrion is structurally similar in both promastigote and amastigote forms, whereas the *T. brucei* mitochondrion lacks extensive tubular cristae in the mammalian blood stream form stage, and has greatly reduced activity (Bienen *et al.*, 1981).

1.1.3.6Glycosomes

Glycosomes are spherical shaped peroxisome-like organelles with a diameter of approximately 0.7 μm (Michels *et al.*, 2006). A key feature of these organelles is the compartmentalisation of the first seven enzymes involved in glycolysis; interestingly, the enzymes involved in net production of adenosine triphosphate (ATP) are located in the cytosol (see Figure 1.4). In other organisms in which all glycolytic enzymes are localised to the cytosol, two ATP-dependent kinases hexokinase (HXK) and phosphofructokinase (PFK) are negatively regulated by intermediate metabolites of glycolysis. This has been shown to be essential in yeast, since removal of inhibition results in uncontrolled glycolysis and build-up of unacceptably high levels of hexose phosphate glycolytic intermediates (Thevelein & Hohmann, 1995). In *T. brucei*, regulation of HXK and PFK has not been practically demonstrated. Indeed, systems biology modelling and practical data have illustrated that the reason for glycosomal compartmentalisation is to prevent a 'lethal turbo charge of glycolysis'; it seems therefore that *T. brucei*, and possibly other kinetoplastids, have evolved a glycolytic system that has no need for regulation of HXK and PFK enzymes (Haanstra *et al.*, 2008).

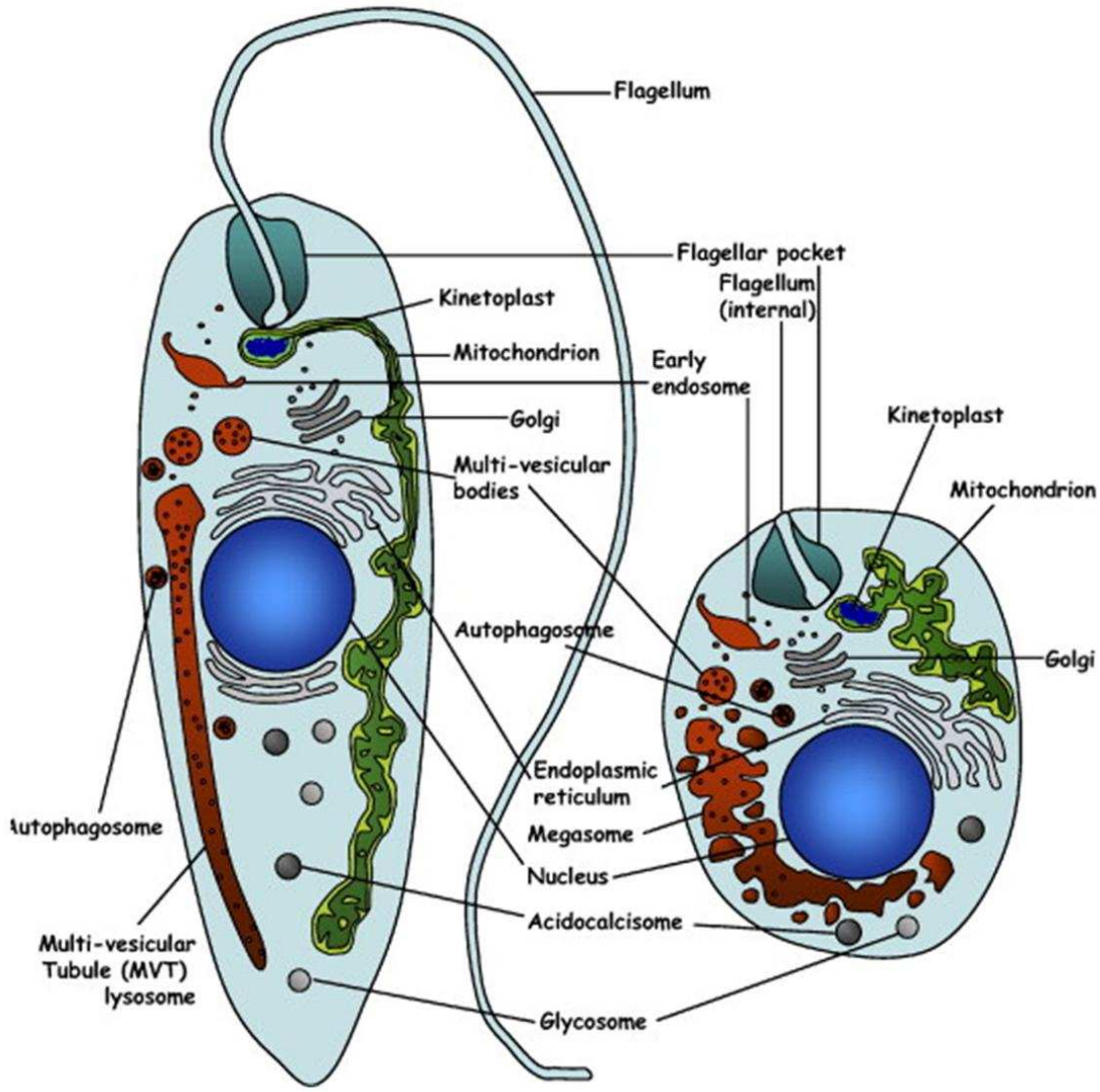


Figure 1.3 Schematic representation of the ultrastructure of *Leishmania* promastigotes and amastigotes
Promastigote on left and amastigote on right. Organelles are labelled. (Image taken from (Besteiro *et al.*, 2007) with permission from Elsevier.)

1.1.4 Chemotherapies

There is currently no vaccine against leishmaniasis and the drugs currently available to treat leishmaniasis are the pentavalent antimonials, that is, sodium stibogluconate (Pentostam) and meglumine antimonate (Glucantime); amphotericin B; pentamidine amd miltefosine (Croft and Coombs, 2003). Aside from miltefosine, all of the other drugs require parenteral administration, which results in a long period of hospitalisation. The drugs also have a range of side effects and drug resistance is increasing (Croft *et al.*, 2005; Murray *et al.*, 2005; Sundar *et al.*, 2000).

In the treatment of visceral leishmaniasis, the pentavalent antimonials have been the first-line treatment for over 60 years (Croft and Coombs, 2003). Antimonials can give rise to severe toxic side effects, including cardiac arrhythmia and pancreatitis. Levels of resistance to these drugs have been reported to be very high in endemic regions, particularly in Bihar, India (Sundar and Rai, 2002). Pentamidine is highly toxic, and as such has generally only been used as a second-line therapy after ineffective antimonial treatment, and resistance to pentamidine treatment has been reported in India (Sundar and Rai, 2002). In some areas of Bihar, Amphotericin B has replaced antimonials as the first-line treatment Amphotericin B is highly effective at treating antimony-resistant cases, but is toxic, with nephrotoxicity and first-dose anaphylaxis reported. The lipid formulation, AmBisome, has much reduced toxicity, but its high cost has limited its use in developing countries until recently (Chappuis *et al.,* 2007).

Miltefosine has been used to treat visceral leishmaniasis in India since 2002 (Davies *et al.*, 2003). Miltefosine has been shown to be safe for both adults and children in India and although gastrointestinal side effects are common, these do not generally prevent completion of treatment (Sundar *et al.,* 2002). It has not however, been established whether miltefosine will be as effective against *L. infantum/chagasi* as it is against *L. donovani* (Croft and Coombs, 2003).

As discussed in Section 1.1.1, cutaneous leishmaniasis often self-heals; as such, drug treatment is not always necessary, and is normally only used if there is a risk of mucocutaneous leishmaniasis. For patients with disfiguring lesions or diffuse or mucocutaneous leishmaniasis, treatment with pentavalent antimonials is the most common route of action (Davies *et al.,* 2003). It is known that there is a difference in the susceptibility of different *Leishmania* species to pentavalent antimonials; *L. donovani* and *L. braziliensis* are more sensitive compared to *L. major*, *L. tropica* and *L. mexicana* (Croft *et al.,* 2002). Pentamidine is used to treat cutaneous and mucocutaneous leishmaniasis (Sundar and Rai, 2002), and amphotericin B is often used as a first-line treatment for travellers (Schwartz *et al.,* 2006).

Miltefosine is effective against *L. panamensis*, less so against *L. mexicana*, and shows a large variation in its efficacy against *L. braziliensis* (Soto *et al.,* 2004; Soto e*t al.,* 2007). Miltefosine has also been tested against Old World cutaneous species and found to give similar cure levels as antimonials (Mohebali *et al.,* 2007).

1.2 Energy metabolism in *Leishmania* and other trypanosomatids

1.2.1 Introduction to energy metabolism

The production of the energy storage compound ATP is essential for active cellular function. There are two methods by which organisms generally produce ATP; substrate-level phosphorylation and oxidative phosphorylation. Substrate-level phosphorylation is defined as the catalysis of phosphoryl transfer from a reactive intermediate to adenosine diphosphate (ADP), resulting in ATP (Nelson & Cox, 2000). For example, as illustrated in Figure 1.4, succinate dehydrogenase (SDH) catalyses the breakage of the high-energy thioester bond between succiante and Coenzyme A (CoA), and in the process itself becomes phosphorylated. SDH then transfers the phosphoryl-group to ADP to produce ATP and succinate (Nelson & Cox, 2000). Oxidative phosphorylation describes an aerobic process (respiration) whereby certain nutrients (for example glucose) are completely oxidised to CO_2 and H_2O (Nelson & Cox, 2000). In the process, electrons of oxidised intermediates are transferred to electron acceptors such as NAD^+ and FAD^+. In eukaryotes, electrons are fed through the electron transport chain, which is located in the inner mitochondrial membrane; the result is the generation of a transmembrane potential and ultimately the production of ATP from the phosphorylation of ADP, and the generation of H_2O from the reduction of O_2 (Nelson & Cox, 2000).

Organisms have adapted their energy metabolisms in different ways, depending upon the nutrients that are available to them as well as the aerobic conditions. Parasites, on the other hand, especially those with multiple hosts within their life cycle, may have also been restricted to certain microenvironments within each host, based on the metabolic machinery that they encode. In terms of energy metabolism in trypanosomatids, most work has been carried out on *T. brucei* (for review see (Bringaud *et al.*, 2006)), and although there are some studies with regards to *Leishmania* promastigote energy metabolism, there are still many questions remaining, especially in terms of amastigote energy metabolism (for review see (Opperdoes & Coombs, 2007)). The aim of this section is to summarise what is known about the nutritional requirements of *Leishmania* with reference to other trypanosomatids (mainly *T. brucei*), and to give an appreciation of how these

organisms have adapted to cope with their digenetic life cycles on a metabolic level.

1.2.2 Nutrition in the fly vector

Trypanosomatids are exposed for part of their life cycles to different compartments within their permissive haematophagous fly vectors. Whereas *Leishmania* promastigotes are restricted to the fly midgut (with the exception of the *braziliensis* complex of *Leishmania*), trypanosomes also invade the fly hindgut and salivary glands. In conjunction with differential geographical locations within the fly vector, the sandfly differs from the vectors of trypanosomes in its feeding habits. For example, it has been shown that the tsetse fly (the vector of *T. brucei*) relies upon amino acids for energy generation during flight (Bursell, 1960; Bursell, 1963). Notably, it was shown that proline is the principal amino acid used as a carbon source, and that proline is partially oxidised via glutamate to alanine (Bursell, 1967). However, sandflies have been reported to feed primarily on nectar and as such *Leishmania* promastigotes in the midgut are thought to have a sugar-rich diet (Schlein, 1986), and *Leishmania* encode glycosidases to digest complex polysaccharides (Jacobson & Schlein, 2001; Jacobson *et al.*, 2001). Interestingly however, *in vitro* studies indicate that *T. brucei* procyclic form (PCF) preferentially use glucose as an energy source when grown in the presence of both glucose and proline (Hellemond *et al.*, 2005; Lamour *et al.*, 2005; van Weelden *et al.*, 2003; van Weelden *et al.*, 2005), yet *Leishmania* promastigotes preferentially use amino acids (Cazzulo *et al.*, 1985). Nevertheless, in the absence of glucose in the medium, proline has been shown to be the principle carbon source for *T. brucei* PCF (Lamour *et al.*, 2005). It is also generally accepted that proline is an important carbon source for energy generation in *Leishmania* in the sandfly (for review see (Bringaud *et al.*, 2006).

1.2.3 Parasite energy metabolism in the fly vector

Although there are likely to be differences between energy metabolism in *T. brucei* PCF and *Leishmania* promastigotes, *T. brucei* has been used as a model for trypanosomatid energy metabolism in the fly stage, primarily due to the fact that this organism is amenable to RNA interference (RNAi), whereas this technique has not been successfully applied in *Leishmania*. As will be discussed in Section 6.3.3.3, gene replacement in *Leishmania* can be a lengthy and unsuccessful tactic

to study certain genes, and RNAi sometimes permits a more subtle dissection of molecular pathways. As such, this section will have a bias towards *T. brucei*.

In *Leishmania* promastigotes and trypanosomatid PCF, energy generation involves both glycolysis and mitochondrial metabolism. As described in Section 1.1.3.6, the ATP-expending reactions of glycolysis occur in peroxisome-like organelles called glycosomes (see Figure 1.3), and the later ATP-yielding reactions of glycolysis occur in the cytosol (see Figure 1.4) (Michels *et al.*, 2006). The end-product of glycolysis – pyruvate – can be decarboxylated to produce acetyl-CoA via the mitochondrial pyruvate dehydrogenase complex (PDH). In yeast, mammals and bacteria, under aerobic conditions acetyl-CoA is completely oxidised through the tricarboxylic acid (TCA) cycle to CO_2, H_2O and ATP. The TCA cycle generates energy through substrate-level phosphorylation (via succinate dehydrogenase; SDH) and through oxidative phosphorylation, feeding reducing equivalents in the form of NADH and $FADH_2$ into the electron transport chain. Trypanosomatids seem to be an exception, since acetyl-CoA is not used in the TCA cycle (Coustou *et al.*, 2008; van Weelden *et al.*, 2003), but is instead converted to acetate by a two-enzyme cycle involving acetate: succinate CoA transferase (ASCT) and succinyl-CoA synthetase (SCS), with concomitant ATP production (see Figure 1.4) (Van Hellemond *et al.*, 1998). Such partial fermentation occurs even in the presence of oxygen, and as such is referred to as aerobic fermentation (for review see (Bringaud *et al.*, 2006)). Acetate is not the only end-product of aerobic fermentation; others include succinate and alanine, and to a lesser extent pyruvate, lactate and ethanol (see Figure 1.4) (for review see (Bringaud *et al.*, 2006)).

It has been shown that in the absence of high concentrations of glucose, *T. brucei* PCF cultured *in vitro* rely upon proline as a carbon source. When cultured in glucose-depleted/proline-containing conditions, the major end products excreted by PCF are alanine (64.2 %) and glutamate (28.2 %) (Coustou *et al.*, 2008). Coustou *et al.* (2008) propose a metabolic pathway, whereby proline is oxidised via glutamate to α-ketoglutarate, which is then fed into the TCA cycle through the action of the α-ketoglutarate dehydrogenase complex (α-KGDH) to produce succinyl-CoA, and ultimately malate. Malate is converted into pyruvate, which is subsequently transaminated with glutamate by alanine aminotransferase (ALAT) to produce alanine. The authors make the potentially intriguing observation that proline metabolism in *T. brucei* PCF mimics that which occurs in the tsetse fly

flight muscle, in that alanine is the principal end-product (Bursell, 1967; Coustou *et al.*, 2008). Interestingly, under glucose-depleted/proline-containing conditions, the electron transport chain was shown to be essential. *Leishmania* promastigotes, however, are semi-tolerant to anoxic conditions, and although they do not proliferate in the absence of oxygen, they do persist. Indeed, after 40 h of anoxia, replenishing the oxygen supply permits a release of the metabolic arrest such that *Leishmania* promastigotes proliferate again (Van Hellemond & Tielens, 1997; Van Hellemond *et al.*, 1997). The slowing of promastigote motility during hypoxic conditions in the sandfly is hypothesised to be important for the simple reason that a reduction in energy expenditure will allow the parasites to persist for longer (Van Hellemond & Tielens, 1997). *T. brucei* PCF are completely intolerant of anaerobic conditions, and start dying after a few days of anoxia (van Weelden *et al.*, 2003).

In comparison to *T. brucei*, little is known in *Leishmania* promastigotes with regards to the relative levels of contribution of glucose and proline in energy metabolism. However, one study illustrates that a *L. mexicana* promastigote mutant line, which lacks three genes that are integral to glucose transport, is able to establish an infection in *L. longipalpis* sandflies, albeit less efficiently (Burchmore *et al.*, 2003). However, the mutant line is unable to replicate within macrophages in the mammalian host (see Section 1.2.5) (Rodriguez-Contreras *et al.*, 2007). Another study illustrated that gluconeogenesis is essential for *L. major* promastigotes in the absence of glucose, since a mutant line lacking the gene encoding the important gluconeogenesis enzyme fructose 1,6-bisphosphatase (FBP) is unable to proliferate on glucose-depleted/glycerol-containing medium (see Figure 1.4) (Naderer *et al.*, 2006). It was also demonstrated that acetyl-CoA derived from β-oxidation of fatty acids (FAs) is not a gluconeogenic substrate in promastigotes (Naderer *et al.*, 2006). This was expected to be the case, since analysis of the *L. major* genome indicates that this parasite lacks the glyoxylate cycle genes encoding isocitrate lyase (ICL) and malate synthase (MLS), and as such is unable to convert acetyl-CoA into malate, and subsequently oxaloacetate, which is used in gluconeogenesis (Ivens *et al.*, 2005). Instead, in conditions of low glucose concentration, amino acids have been shown to be used for gluconeogenesis (Burchmore *et al.*, 2003). Overall, in promastigotes it is likely that glucose plays an important role as a substrate for the pentose phosphate pathway (PPP) to generate precursors for DNA and RNA biosynthesis (Maugeri *et al.*, 2003), and is also likely to be used for the production of carbohydrate, such as

β1,2-mannan reserve oligosaccharides (see Figure 1.4) (Ralton *et al.*, 2003). However, in terms of *Leishmania* energy metabolism, proline is thought to be more important than glucose (see Figure 1.4) (Opperdoes & Coombs, 2007).

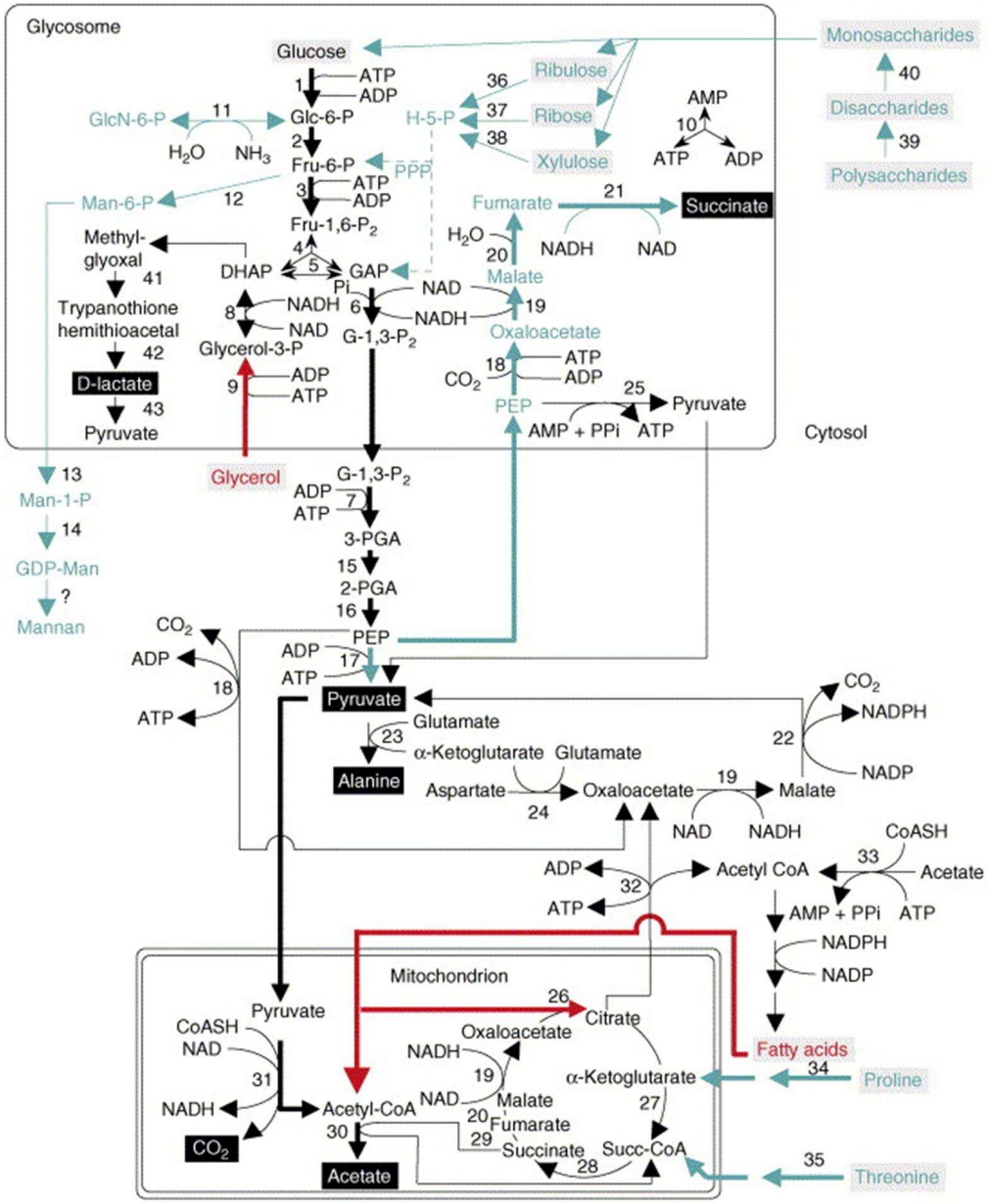


Figure 1.4 The pathways of core metabolism in *L. major*
Figure shows reactions taking place in the glycosome, that are involved in carbohydrate metabolism, and in the mitochondrion, with its tricarboxylic acid cycle, and the flux of metabolites between these two organelles. Boxed metabolites are substrates (grey) or end products (black) of metabolism. Thick arrows represent major metabolite fluxes. Pathways in blue are thought to be more important in promastigotes and pathways in red are thought to be more important in the amastigote. Abbreviations: Fru, fructose; GAP, glyceraldehyde 3 phosphate; Glc, glucose; H-5-P, hexose 5-phosphate; Man, mannose; PEP, phosphoenolpyruvate; PGA, phosphoglyceric acid; PPP, pentose-phosphate pathway. Enzymes: 1, hexokinase; 2, phosphoglucose isomerase; 3, phosphofructokinase;4, fructosebisphosphate aldolase; 5, triosephosphate isomerase; 6, glyceraldehyde-3-phosphate dehydrogenase; 7, phosphoglycerate kinase; 8, glycerol-3-phosphate dehydrogenase; 9, glycerol kinase; 10, adenylate kinase; 11, glucosamine-6-phosphate deaminase; 12, mannose-6-phosphate isomerase; 13, phosphomannomutase; 14, GDP-mannose pyrophosphorylase; 15, phosphoglycerate mutase; 16, enolase; 17, pyruvate kinase; 18, phosphoenolpyruvate carboxykinase; 19, malate dehydrogenase; 20, fumarate hydratase; 21, NADH-dependent fumarate reductase; 22, malic enzyme; 23, alanine aminotransferase; 24, aspartate aminotransferase; 25, pyruvate phosphate dikinase; 26, citrate synthase; 27, 2-ketoglutarate dehydrogenase; 28, succinyl-CoA ligase; 29, succinate dehydrogenase; 30, acetate–succinate CoA transferase; 31, pyruvate dehydrogenase; 32, citrate lyase; 33, acetyl-CoA synthetase; 34, proline oxidation pathway; 35, threonine oxidation pathway; 36, ribulokinase; 37, ribokinase; 38, xylulokinase; 39, amylase-like protein; 40, sucrase-like protein. (Image taken from (Opperdoes & Coombs, 2007) with permission from Elsevier.)

1.2.4 Nutrition in the mammalian host

In the mammalian host (human), *Leishmania* amastigotes and *Trypanosoma cruzi* amastigotes have an intracellular life style, whereas *T. brucei* BSF lives in the blood. *T. cruzi* trypomastigotes infect a variety of cell types, including cardiomyocytes, and once within the cell, trypomastigotes differentiate into amastigotes, which multiply in the cytosol. Both *T. cruzi* and *T. brucei* thus inhabit hexose sugar-rich environments.

Leishmania amastigotes reside within the acidic phagolysomal compartment of macrophages, which is also referred to as the PV. In fact, different species of *Leishmania* are known to infect different populations of macrophages, and the morphology of the phagolysosome differs depending upon the species of *Leishmania* (Antoine *et al.*, 2004). For example, *L. mexicana* induce spacious vacuoles, rather than tight-fitting phagolysosomes found in *L. major* infections (Russell *et al.*, 1992). Interestingly, aside from the Gram-negative bacterium *Coxiella burnetii* (Voth & Heinzen, 2007), no other microbial pathogen is known to replicate within the phagolysosomal compartment. Unlike the hexose-rich cytosolic compartment that *T. cruzi* inhabits, the phagolysosome lumen is hexose-poor, and the main carbon sources for energy generation are thought to come from host amino acids and FAs (Bringaud *et al.*, 2006; McConville *et al.*, 2007; Naderer & McConville, 2008; Opperdoes & Coombs, 2007). Lastly, there is some evidence to suggest that the metabolic and activation status of the infected macrophage has a significant influence upon amastigote growth (Gordon, 2003; Sacks & Anderson, 2004).

1.2.5 Parasite energy metabolism in the mammalian host

Given the very different tissue-specificities of the different trypanosomatids, it is not surprising that their energy metabolism differs much more drastically than in the fly host. *T. brucei* BSF lacks most of the mitochondrial enzymes found in the PCF and completely relies upon glycolysis for energy generation (Bakker *et al.*, 2000; Fairlamb & Opperdoes, 1986).

Leishmania amastigotes acquire energy via β-oxidation of FAs (see Figure 1.4) (Hart & Coombs, 1982). It is thought that the acetyl-CoA produced from β-oxidation is fed into a fully functional TCA cycle (Hart & Coombs, 1982), however

this needs reanalysis given that trypanosomatid promastigotes and PCF do not fully oxidise acetyl-CoA via the TCA cycle (Opperdoes & Coombs, 2007). A recent proteomic study highlighted the fact that during the *in vitro* differentiation of *L. donovani* metacyclic promastigotes into amastigotes, enzymes involved in the glycosomal-restricted portion of glycolysis are down-regulated, whereas key enzymes involved in β-oxidation of FAs were significantly up-regulated (Rosenzweig *et al.*, 2008). The authors also reported increases in the levels of TCA cycle enzymes and of enzymes involved in the catabolism of amino acids (Rosenzweig *et al.*, 2008).

Despite the fact that the phagolysosomal compartment is hexose-poor, an *L. mexicana* mutant lacking three glucose transporters is unable to proliferate when inoculated into mice (Burchmore *et al.*, 2003). It was shown that the mutant promastigotes are taken up by host macrophages and differentiate into amastigotes, yet do not persist (Rodriguez-Contreras *et al.*, 2007). To further strengthen the case that glucose uptake is essential for *L. mexicana* amastigote development, a reversion of this mutant line has been reported, in which an alternative hexose transporter has been amplified on a circular extrachromosomal element (Feng *et al.*, 2009b). This naturally occurring selection for an up-regulation of an alternative hexose transporter reverted the avirulent phenotype of the glucose transporter mutant in mice, thus highlighting the importance of hexoses like glucose for amastigote survival and replication (Feng *et al.*, 2009b). On the other hand, gluconeogenesis has been reported to be essential for the survival of *L. major* within macrophage phagolysosomes (Naderer *et al.*, 2006). If the uptake of hexoses is also essential to *L. major* amastigote fitness, it would appear that gluconeogenesis is not sufficient to provide enough hexose for PPP or for carbohydrate biosynthesis (or for other purposes), and that uptake of even small concentrations of hexoses in the hexose-poor PV niche is critical for the survival of the amastigote stage.

1.3 Production of fatty acids

Leishmania have a very robust energy metabolism, with the capacity to partially oxidise sugars, amino acids and FAs in order to generate energy (see Section 1.2). Key to several of the important enzymes within the TCA cycle is the molecule lipoic acid (LA), which is the topic of this thesis. LA can be synthesised from FAs

(see Section 1.4.3), and as such the aim of this section is to give an appreciation of the different systems that exist in trypanosomatids to produce FAs.

FAs are carboxylic acids with unbranched aliphatic tails that are either saturated or unsaturated. The process of FA synthesis involves the extension of a ‘primer’ FA building block by addition of a 2 carbon (2C) FA unit. FAs are bound to either acyl carrier protein (ACP) or to CoA. ACP and CoA are covalently modified with a 4'-phosphopantetheine group (Magnuson *et al.*, 1993). FAs are bound to ACP and CoA through a thioester linkage with the 4'-phosphopantetheine group. One cycle of 2C addition requires four steps; condensation, reduction, dehydration and reduction. These steps are catalysed by either type I or type II fatty acid synthase (FAS). The type I FAS is found in mammals and yeast, is localised to the cytosol, and arranges its different catalytic activities on separate domains spanning either one or two polypeptide chains (Smith, 1994). The type II FAS occurs in plants and bacteria, and each catalytic activity is associated with an individual polypeptide chain (Rock & Jackowski, 2002; White *et al.*, 2005). Recent data indicate that mammals (Witkowski *et al.*, 2007) and yeast (Hiltunen *et al.*, 2005) also contain a mitochondrial type II FAS, which is not merely a relict of the endosymbiont, but is essential to respiration and cellular fitness in mammals (Feng *et al.*, 2009a) and yeast (Brody *et al.*, 1997; Schonauer *et al.*, 2008; Sulo & Martin, 1993). In both type I and type II FAS, the growing acyl chain is esterified to ACP (White *et al.*, 2005).

In addition to *de novo* biosynthesis of FAs via type I or type II FAS, a set of elongase enzymes also exist, whose role in mammals and yeast is to elongate long-chain FAs (C14-18) to make very long-chain FAs (C20 or longer).

1.3.1 Endoplasmic reticulum elongase system

Elongase enzymes carry out the same four condensation, reduction, dehydration and reduction reactions that are catalysed by FAS (Leonard *et al.*, 2004). Unlike FAS, elongases (ELOs) are localised to the endoplasmic reticulum (ER) and are membrane-bound (Moon *et al.*, 2001; Oh *et al.*, 1997; Toke & Martin, 1996). Another difference is that the growing acyl chain is esterified to CoA instead of to ACP. In yeast and mammals, ELOs incrementally increase the length of an acyl chain using malonyl-CoA as the 2C donor (Leonard *et al.*, 2004). Multiple elongase systems exist (Cinti *et al.*, 1992), each with different chain length

specificity, in order to permit the production of various types of polyunsaturated fatty acids (PUFAs).

In *T. brucei*, a membrane-associated FA synthesis was discovered in a cell-free system (Morita *et al.*, 2000), which was interesting for several reasons. Firstly, a soluble mitochondrial type II FAS had been predicted (and since shown to be active) (Stephens *et al.*, 2007), and as such it was peculiar that FA synthesis was associated with the membrane fraction. Secondly, whereas malonyl-CoA or acetyl-CoA were expected to serve as a primer for FA synthesis, only butyryl-CoA worked as a primer (Morita *et al.*, 2000). Thirdly, knock-down of ACP by RNAi had no effect on bulk FA synthesis in the cell-free system (Lee *et al.*, 2006).

It was recently shown that the membrane-associated FA synthesis in *T. brucei* is due to the action of ELO 1-4. However, unlike yeast and mammalian ELOs, *T. brucei* ELO1 and ELO2 were shown to primarily synthesise myristate (C14) *de novo* from a butyryl-CoA primer in *T. brucei* BSF (Lee *et al.*, 2006). The authors also identified 13 potential ELOs in the *L. major* genome and five potential ELOs in the *T. cruzi* genome. Preliminary data from the cell-free system illustrated that like *T. brucei*, both *L. major* and *T. cruzi* possess membrane-associated FA synthesis activity (Lee *et al.*, 2006). So far, *T. brucei* is the only organism that has been experimentally shown to have adopted the microsomal ELO system for the purpose of *de novo* FA synthesis (Lee *et al.*, 2006; Lee *et al.*, 2007).

1.3.2 Mitochondrial type II FAS

Type II FAS uses ACP to extend an acyl chain, using acetyl-CoA (1C) as a primer and malonyl-CoA as the 2C donor (for review see (White *et al.*, 2005)). Acetyl-CoA can be derived from the glycolytic product pyruvate via the action of mitochondrial pyruvate dehydrogenase. Alternatively, the amino acid threonine can be converted to pyruvate via threonine dehydrogenase, and an intermediate in this pathway can undergo thiolysis with CoA to produce acetyl-CoA and glycine.

The first enzyme involved in type II FAS is acetyl-CoA carboxylase (ACC), which is a biotin-dependent enzyme that catalyses the conversion of acetyl-CoA to malonyl-CoA. The malonyl-moiety is subsequently transferred onto ACP by malonyl-CoA: ACP transferase. The addition of malonyl-ACP to a growing acyl chain is catalysed by the sequential action of four enzymes. Firstly, condensation

of acetyl-CoA/acetyl-ACP and malonyl-CoA to form 3-ketoacyl-ACP is catalysed by 3-ketoacyl-ACP synthase. Next, 3-ketoacyl-ACP is reduced to 3-hydroxyacyl-ACP by 3-ketoacyl: ACP-reductase. The dehydration reaction producing trans-2-enoyl-ACP is catalyzed by 3-hydroxyacyl thioester dehydratase. Lastly, trans-2-enoyl-ACP is reduced by 2-enoyl thioester reductase to produce the acyl-ACP molecule that has been extended by two carbon atoms. The acyl-ACP species can then be used as a substrate by ketoacyl-ACP synthase and the type II FAS cycle continues.

Interestingly, it has been noted in fungi (Wada *et al.*, 1997), plants (Wada *et al.*, 1997), yeast (Brody *et al.*, 1997; Schonauer *et al.*, 2008) and mammals (Feng *et al.*, 2009a; Witkowski *et al.*, 2007) that the principal end product of type II FAS is the 8 carbon (8C) FA octanoyl-ACP (see Section 1.4.3). Type II FAS also produces FA chains up to myristate (14C) in *Neurospora crassa* (Mikolajczyk & Brody, 1990), up to palmitate (16C) in *T. brucei* (Stephens *et al.*, 2007) and up to stearate (18C) in plants (Gueguen *et al.*, 2000), however much less is known with regards to the role of acyl-chains longer than 8C.

The trypanosomatids *T. brucei*, *T. cruzi* and *L. major* all possess genes that potentially encode all components of a type II FAS (Lee *et al.*, 2007; Stephens *et al.*, 2007). Research on *T. brucei* has shown that all components of type II FAS are localised to the mitochondrion (Autio *et al.*, 2008a; Stephens *et al.*, 2007). *T. brucei* mitochondrial type II FAS is essential, since RNAi of ACP causes cell death in PCF and BSF parasites (Guler *et al.*, 2008; Stephens *et al.*, 2007). It was shown that *T. brucei* RNAi resulted in a significant decrease in LA (Stephens *et al.*, 2007). However, the principal cause of a decrease in parasite fitness was found to be due to a significant decrease in the production of certain phospholipids, resulting in altered mitochondrial ultrastructure, mitochondrial membrane potential, and ultimately a decrease in cytochrome-mediated respiration (Guler *et al.*, 2008).

A recent surge of publications have received much attention recently, since they illustrate that type II FAS in fungi, yeast, mammals, plants and *T. brucei* is pivotal to cellular survival. As will be discussed in Section 6.3.3.3, evidence is accumulating that one of the main functions of type II FAS is to produce octanoyl-ACP, which is the precursor of LA (Hiltunen *et al.*, 2009).

1.4 Lipoic acid

LA (6,8-thioctic acid or 1,2-dithiolane-3-pentanoic acid) was first isolated from bovine liver in 1951 (Reed *et al.*, 1951). LA is an eight-carbon fatty acid with two sulphur atoms linked to carbons 6 and 8 that form a disulphide 1,2-dithiolane ring in the oxidised form (see Figure 1.5), which can be reduced to form dihydrolipoic acid (DHLA) (Atmaca, 2004). R-LA is the naturally occurring enantiomer of LA, and from now on will simply be referred to as LA. From bacteria to mammals, LA is an essential cofactor in several multienzyme complexes that are involved in key metabolic processes (Perham, 2000). Additionally, LA is a potent antioxidant, and there have been a plethora of recent studies into the potential therapeutic applications of LA (Estrada *et al.*, 1996) (Beitner, 2003) (Hager *et al.*, 2001). LA can be synthesised from the type II FAS product octanoyl-ACP (see Sections 1.3.2 and 1.4.3), or can be salvaged from the external milieu (see Section 1.4.3.2).

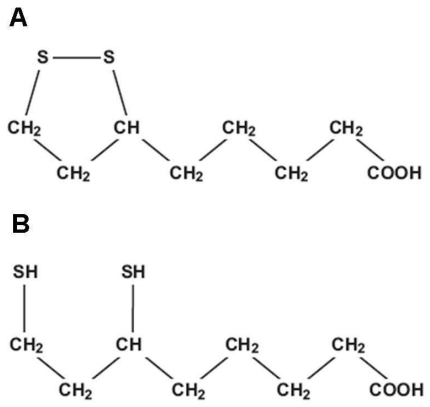


Figure 1.5 Schematic representations of LA and DHLA
Diagrams showing the structure of the oxidised form (*A*) and reduced form, DHLA (*B*) of LA.

1.4.1 Antioxidant properties of the lipoic acid/dihydrolipoic acid redox pair

An antioxidant is defined by several criteria, which include specificity of free radical quenching, metal chelating activity, interactions with other antioxidants, and effects on gene expression. Additional criteria that are important when considering therapeutic applications are bioavailability, concentration and solubility. Vitamin E, for example, is soluble only in the lipid phase and its main role is to quench lipid peroxyl radicals (Burton & Ingold, 1981). On the other hand, the LA/DHLA redox

pair fulfil most of the above criteria, and for this reason have been described as a 'universal antioxidant' (Kagan *et al.*, 1992).

The LA/DHLA redox pair has a redox potential of -0.32 V (Searls & Sanadi, 1960); as a comparison, the redox potential of the glutathione (GSH)/glutathione disulphide (GSSG) pair is -0.24 V (Scott *et al.*, 1963). This low electronegativity means that DHLA is a very strong reductant. As well as being able to reduce GSSG to GSH (Jocelyn, 1967), DHLA is an antioxidant of ascorbate (vitamin C), and due to its solubility in lipid as well as water (Roy & Packer, 1998), DHLA can also reduce lipid-soluble tocopherol (vitamin E) (Podda *et al.*, 1994; Rosenberg & Culik, 1959).

In addition to the recycling of antioxidants as described above, LA is able to scavenge hydroxyl radicals (Scott *et al.*, 1994; Suzuki *et al.*, 1991), hypochlorous acid (Haenen & Bast, 1991) and singlet oxygen (Devasagayam *et al.*, 1991). DHLA is able to scavenge hypochlorous acid (Haenen & Bast, 1991), and unlike LA, DHLA is also an effective buffer of peroxyl radicals (Kagan *et al.*, 1992; Suzuki *et al.*, 1993). DHLA, but not LA, is also able to chelate redox active metal ions such as iron, thus minimising the production of highly reactive hydroxyl radicals through the Fenton reaction (Suh *et al.*, 2004).

Given the important nature of LA/DHLA as an antioxidant, much attention has recently been focused on understanding whether this thiol could be of therapeutic usefulness. It was, for example, demonstrated that LA has beneficial effects in the treatment of diabetes type 1 and type 2, which are both associated with elevated levels of oxidative stress. In both type 1 and type 2 diabetes, the problem is a lack of insulin-mediated glucose uptake into cells. Insulin is ordinarily made within β cells of the pancreas, however type 1 diabetes is an autoimmune disease that results in the destruction of these insulin-generating β cells. The main problem associated with type 2 diabetes is resistance to insulin. *In vitro*, it has been demonstrated that LA stimulated glucose uptake into muscle cells in a manner similar to that induced by insulin (Estrada *et al.*, 1996). In a study whereby 74 patients with type 2 diabetes were administered 600 mg LA or 'placebo' once, twice or three times daily for four weeks, patients that received LA had a significant increase in insulin-stimulated glucose uptake compared to the placebo control (Jacob *et al.*, 1999).

In terms of therapeutic applications, LA/DHLA is also being tested for its protective/regenerative potential in other diseases, such as cataract (Maitra *et al.*, 1995) and Alzheimer's disease (Hager *et al.*, 2001), as well as gaining the attention of cosmetologists and dermatologists (Beitner, 2003).

In addition to its role as an antioxidant, LA has also been shown to affect transcription of certain genes, by regulating redox-sensitive transcription factors present in the cytosol. For example, LA is known to inhibit the translocation of the transcription factor NF-κB from the cytosol to the nucleus (Zhang & Frei, 2001). More recent data indicate the intriguing observation that LA is directly or indirectly involved in RNA processing in yeast (Schonauer *et al.*, 2008) and vertebrate (Autio *et al.*, 2008b) mitochondria, marking an evolutionarily conserved intersection of positive feedback loops between type II FAS and RNA processing.

Overall, LA is a multifaceted molecule, with roles as an antioxidant and as a signalling molecule. The focus of this thesis however, is on the role of LA in energy metabolism, in which it acts as a cofactor for four enzyme complexes, with each playing a key role in either substrate level/oxidative phosphorylation or in the creation of C1 units for folate metabolism.

1.4.2 Role of LA in energy metabolism

1.4.2.1α-Ketoacid dehydrogenase complexes

LA plays a key role in energy metabolism, since it acts as an essential cofactor of α-KADH complexes (α-KADHs) (Perham, 2000). α-KADHs consist of multimers of three proteins that are strongly (but not covalently) associated; a substrate-specific α-ketoacid decarboxylase (E1), an acyltransferase (E2) and a dihydrolipoamide dehydrogenase (E3) (see Figure 1.6). Mammals (De Marcucci & Lindsay, 1985) and yeast (Maeng *et al.*, 1994) also encode a protein that is structurally-similar to E2, called E3 binding protein (E3BP). Three types of KADHs are found in eukaryotes; pyruvate dehydrogenase (PDH), α-ketoglutarate dehydrogenase (α-KGDH) and branched-chain α-ketoacid dehydrogenase (BCKDH). These multienzyme complexes can be up to 10 MDa in size. Recent data indicate that the BCKDH may even exist as a part of an even larger supramolecular complex

with the branched-chain aminotransferase (BCAT), referred to as the branched-chain amino acid metabolon (BCAA metabolon) (Islam *et al.*, 2007).

α-KADHs catalyse the conversion of α-ketoacid, NAD^+ and coenzyme A (CoA) into CO_2, NADH and acyl-CoA (see Figure 1.6). The substrate-specific E1 subunit contains a thiamine pyrophosphate (TPP) cofactor, and catalyses the decarboxylation of an α-ketoacid, generating CO_2 and an acyl-TPP moiety, which remains bound to the E1 subunit. The E2 subunit contains the LA cofactor, which is covalently bound via an amide linkage to the N^6-amino group of a specific lysine residue of the E2 subunit. The lipoamide acts as a swinging arm in the E2-catalysed transfer of the acyl group from acyl-TPP to coenzyme A to form acyl-CoA (see Figure 1.6). The high-energy acyl-CoA thioester is either fed into the TCA cycle or used as a precursor in fatty acid biosynthesis. The last step in the reaction is the re-oxidation of dihydrolipoamide by the FAD-containing E3 subunit, which results in lipoamide and NADH (see Figure 1.6) (Perham, 1991; Perham, 2000; Reed & Hackert, 1990). The reduced NADH derived from the reaction can be used in the electron-transport chain with production of ATP (Voet & Voet, 2005).

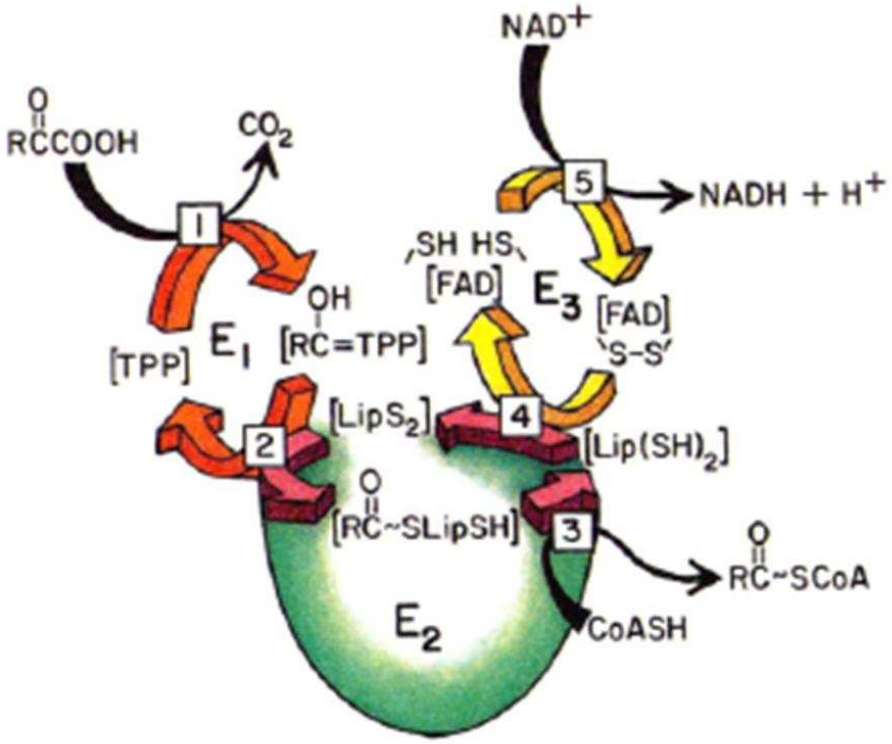


Figure 1.6 Reaction mechanism of α-KADHs
1, The E1 subunit catalyses the decarboxylation of an α-ketoacid to form an acyl group, in a TPP-dependent manner. The E2 subunit catalyses transfer of the acyl group (*2*) onto a CoA receptor (*3*) to form acyl CoA, and in the process the lipoyl-moiety becomes reduced. The E3 subunit catalyses the oxidation of the lipoyl-group (*4*), with the electron acceptor being NAD^+ (*5*). Abbreviations: TPP, thiamin diphosphate; $LipS_2$ and $Lip(SH)_2$, lipoyl moiety and its reduced form. (Image taken from (Reed & Hackert, 1990) with permission from Elsevier.)

In every α-KADH, the E2 components are assembled into either a cubic (24-mer) or a dodecahedral (60-mer) inner core, around which E1 and E3 components bind (Perham, 1991; Perham, 2000; Reed & Hackert, 1990). The E2 protein consists of three structurally- and functionally distinct domains. The *N*-terminal portion is the lipoyl-domain, which contains the signature lysine residue that is post-translationally modified by LA. Depending upon the species and the α-KADH, there may be up to three lipoyl-domains, each spanning approximately 80 amino acid residues (Packman *et al.*, 1984). However, the reason for the existence of more than one lipoyl-domain is not fully understood, since computer modelling (Hackert *et al.*, 1983a; Hackert *et al.*, 1983b) and experimental evidence (Allen *et al.*, 1989; Guest *et al.*, 1985) indicate that only one lipoyl-domain is required for optimal α-KADH activity. The lipoyl-domain is followed by a subunit-binding domain, which confers binding of E1 and E3 subunits to the E2 core. Nevertheless, the PDH E2 subunit in mammals (De Marcucci & Lindsay, 1985) and some yeast (Maeng *et al.*, 1994) (but not bacteria) is not able to bind E3 directly, and instead requires the accessory E3BP; E3 binds to E3BP and is

inserted into the E2p scaffold (Sanderson *et al.*, 1996a; Sanderson *et al.*, 1996b). After the subunit-binding domain is the *C*-terminal catalytic domain, which catalyses the transfer of the acyl-moiety from the E1 subunit onto the CoA acceptor, to produce acyl-CoA (see Figure 1.6).

The three domains that comprise the E2 subunit are separated by linker regions, which are approximately 20-30 amino acids in length, and are typically rich in amino acids alanine and proline, which gives these structures flexibility (Packman *et al.*, 1984; Perham *et al.*, 1981). These flexible linkers permit the lipoyl-domain to act as a 'swinging domain', shuttling reaction intermediates between E1, E2 and E3 subunits (Fries *et al.*, 2007; Green *et al.*, 1992; Perham *et al.*, 1981; Perham, 1991; Reed & Hackert, 1990).

The role of the E1 subunit is to confer substrate-specificity to the α-KADH, and requires TPP and Mg^{2+} cofactors for activity (see Figure 1.6). The E3 protein is a flavoprotein, which is responsible for the re-oxidation of dihydrolipoamide to lipoamide (see Figure 1.6). In most organisms, the E3 protein is shared between different complexes (Bourguignon *et al.*, 1996). The situation is different in organisms that possess α-KADHs localised to different organelles, such as plants and *P. falciparum*; in such cases more than one E3 protein exists, in order to serve α-KADHs in an organellar-specific manner (Lutziger & Oliver, 2000; McMillan *et al.*, 2005).

The PDH catalyses the oxidative decarboxylation of pyruvate to acetyl-CoA (see Figure 1.7). The enzyme complex is generally found in the mitochondrion, where it forms the link between glycolysis and the TCA cycle. Plants possess an additional PDH enzyme complex, which is present in the plastid, where it provides acetyl-CoA and NADH for fatty acid biosynthesis (Mooney *et al.*, 2002).

The BCKDH complex catalyses the oxidative decarboxylation of branched-chain α-ketoacids, namely 2-ketoisocaproate, 2-keto-3-methylvalerate and 2-ketoisovalerate (derived from leucine, isoleucine and valine, respectively) to form an acyl-group (isovaleryl, 2-methylbutyryl or isobutyryl, respectively), in a TPP-dependent process (Massey *et al.*, 1976) (see Figure 1.7). Eventually, the acyl-CoA products can be converted to acetyl-CoA and/or succinyl-CoA for usage in the TCA cycle (Anderson *et al.*, 1998).

The α-KGDH is also found in the mitochondrion, where it catalyses the oxidative decarboxylation of α-ketoglutarate to succinyl-CoA. The production of succinyl-CoA is another branch point into the TCA cycle, other than acetyl-CoA. In *Leishmania* and *T. brucei* insect stage forms, proline (and glutamine) is thought to be a key carbon source. Proline is ultimately catabolised to α-ketoglutarate, which is fed into the TCA cycle through the action of the α-KGDH (see Figure 1.7). As such, α-KGDH is likely to be a key enzyme in *Leishmania* energy metabolism.

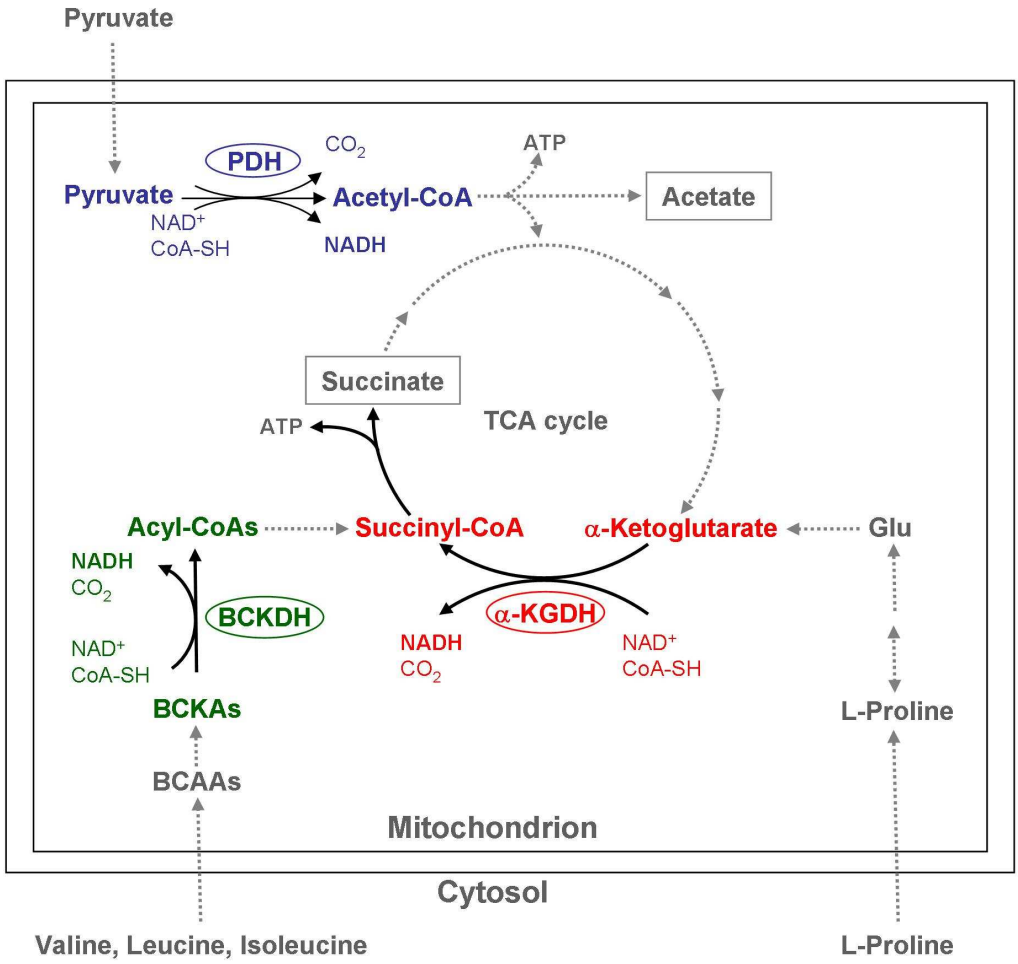


Figure 1.7 Schematic representation of the reactions catalysed by different α-KADHs
This figure is based upon the information presented in Figure 1.4, and highlights the relevance of the different α-KADHs in mitochondrial energy metabolism. Abbreviations: Glu, glutamate.

1.4.2.2Glycine cleavage complex

LA is also an essential cofactor of the glycine cleavage complex (GCC). The GCC consists of multiple copies of four protein subunits; a decarboxylase (P-protein), an aminomethyltransferase (T-protein), a dihydrolipoamide dehydrogenase (L-protein) and a carrier protein (H-protein). The GCC is present in the mitochondria,

plastids and cytosol of plant photosynthetic cells, and in the mitochondria of mammals (Cossins & Chen, 1997; Douce *et al.*, 2001).

The GCC catalyses the oxidative decarboxylation and deamination of glycine, generating CO_2, NH_3, NADH and N_5,N_{10}-methylene tetrahydrofolate (5,10-CH_2-THF) (see Figure 1.8 for reaction mechanism). C1 units provided by 5,10-CH_2-THF are used by folate coenzymes to synthesise essential cellular compounds, including pyrimidines and mitochondrial protein (Appling, 1991).

The P-protein catalyses the decarboxylation of glycine and the reductive methylamination of lipoamide, which is covalently attached to the H-protein. The T-protein is a THF-dependent enzyme, and catalyses the transfer of methylene to THF and the subsequent release of NH_3. The H-protein then reacts with the L-protein, the dihydrolipoamide dehydrogenase described above, to oxidise the dihydrolipoamide, with NAD^+ as the final electron acceptor. The released 5,10-CH_2-THF reacts with another glycine molecule, resulting in the formation of serine, a reaction that is catalysed by serine hydroxymethyltransferase (SHMT) (Douce *et al.*, 2001).

The GCC and α-KADH share similar mechanisms of action, most relevant of which is the parallel between α-KADH E2 subunits and GCC H-protein. Like the E2 subunit, the H-protein is lipoylated, and the bound LA shuttles reaction intermediates between different catalytic subunits, itself being reduced and oxidised. Nevertheless, whereas the E2 subunit bears catalytic activity, the H-protein does not, and is just a carrier protein.

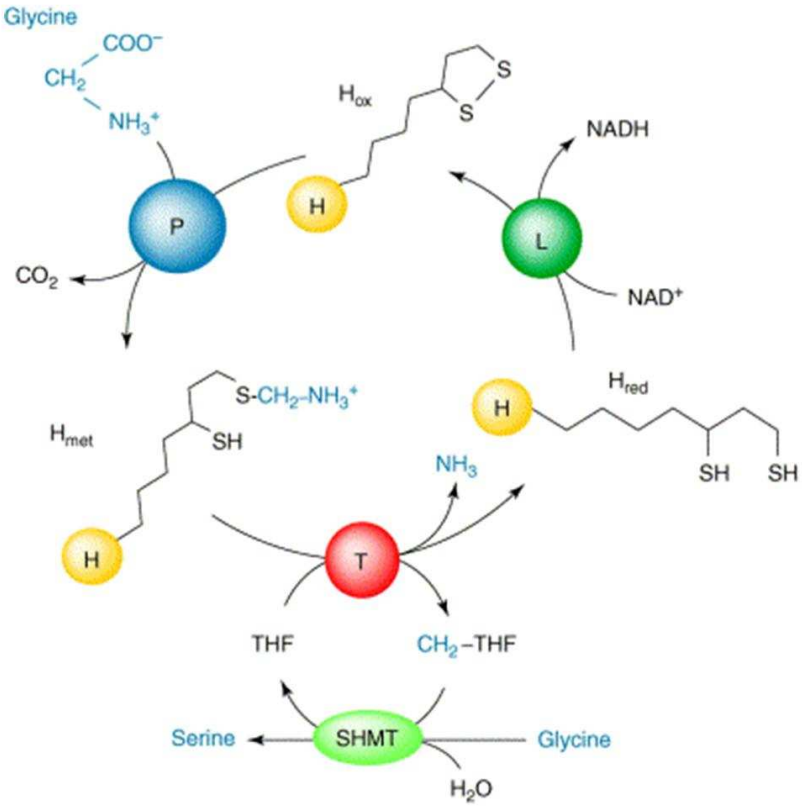


Figure 1.8 Reaction mechanism of the GCC
Outline of the reactions involved in oxidative decarboxylation and deamination of glycine in plant mitochondria. P-, H-, T- and L- are the protein components of the glycine-decarboxylase multienzyme system. The key enzyme in the entire sequence of reactions is the lipoamide-containing H-protein, which undergoes a cycle of reductive methylamination (catalysed by the P-protein), methylamine transfer (catalysed by the T-protein) and electron transfer (catalysed by the L-protein). SHMT: serine hydroxymethyltransferase involved in the conversion of CH_2–THF to THF at the expense of a second molecule of glycine. Note that the methylamine moiety deriving from glycine is passed to the distal sulphur of the dithiolane ring. H_{met}, H_{red} and H_{ox}: methylaminated, reduced and oxidized forms of the H protein, respectively. (Image taken from (Douce *et al.*, 2001) with permission from Elsevier.)

1.4.3 LA metabolism

As discussed in Section 1.4.2, LA is an essential cofactor of α-KADHs and the GCC. LA is ligated to the E2 subunits of α-KADHs and to the H-protein of the GCC in a process referred to as lipoylation. Specifically, LA is attached to the ε-amino group of a strictly conserved lysine residue of apoproteins (Douce *et al.*, 2001; Reche & Perham, 1999).

Two pathways exist to carry out lipoylation of α-KADHs and the GCC (see Figure 1.9). Firstly, LA can be synthesised *de novo*, which necessitates the sequential action of two enzymes; octanoyl-[acyl carrier protein]: protein *N*-octanoyltransferase (LipB) covalently attaches ACP to the apoproteins (Cronan *et al.*, 2005; Zhao *et al.*, 2005). The octanoyl-ACP used as a precursor for LA

synthesis is provided by mitochondrial type II FAS, as discussed in Section 1.3.2 (Gueguen *et al.*, 2000; Jordan & Cronan, 1997b; Miller *et al.*, 2000). Following ligation of OA to the apoprotein, two sulphurs are inserted at carbon positions six and eight, to form the lipoyl-arm. This reaction is catalysed by lipoic acid synthase (LipA), an S-adenosyl-methionine (SAM)-dependent, [Fe-S] cluster-containing enzyme (Cicchillo & Booker, 2005; Miller *et al.*, 2000; Zhao *et al.*, 2003). LA can also be salvaged from the environment, and scavenged LA is ligated directly to the enzyme complexes by either of the two known salvage pathways. Salvage of LA in *E. coli* requires just one enzyme, lipoate protein ligase (LplA) (Morris *et al.*, 1994; Reed *et al.*, 1994), which catalyses the formation of an activated intermediate, lipoyl-AMP, and the subsequent transfer of the lipoyl moiety to the apoprotein. The mammalian system necessitates the sequential action of two enzymes for activation and transfer of LA to the apoprotein; medium-chain acetyl-CoA synthetase 1 (ACSM1) (Fujiwara *et al.*, 2001) and lipoyltransferase (LT) (Fujiwara *et al.*, 1997a; Fujiwara *et al.*, 1999), respectively.

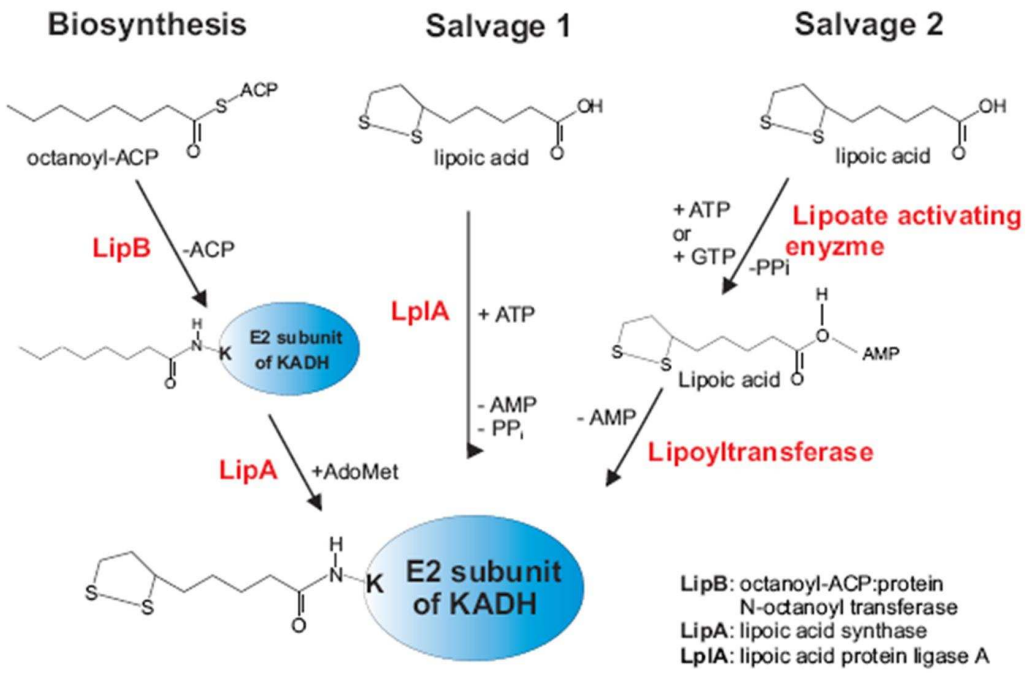


Figure 1.9 Schematic representation of LA metabolism pathways
LA can be acquired through biosynthesis or salvage pathways. In the biosynthesis pathway, octanoyl-ACP from mitochondrial type II FAS is used a substrate for LipB, which transfers the octanoyl-moiety to the E2 subunit of a α-KADH or to the H-protein of the GCC. LipA catalyses the insertion of two sulphur atoms into the octanoyl-moiety to form the lipoyl-moiety. LA is salvaged by bacteria and mammals in an ATP-dependent reaction. In *E. coli*, one enzyme – LplA – catalyses the activation of LA to form lipoyl-AMP, and the subsequent transfer of the lipoyl-moiety to the apo-E2-protein or apo-H-protein. In mammals, salvage of LA necessitates the sequential action of two enzymes; lipoate-activating enzyme (medium-chain acyl-CoA synthetase) catalyses the activation of LA and lipoyltransferase catalyses the transfer of the lipoyl-moiety to the apo-E2-protein or apo-H-protein.

1.4.3.1LIPA

LipA catalyses the production of LA from OA by inserting two sulphur atoms at carbon position six and eight (see Figure 1.9). As described in Section 1.3.2, octanoyl-ACP is provided by mitochondrial type II FAS. In *E. coli*, it was demonstrated that LipA has significantly higher preference for the apoprotein-bound octanoyl-moiety (octanoamide) as a substrate rather than free octanoyl-ACP (Zhao *et al.*, 2003). LipA is a member of the radical SAM superfamily, which use SAM and a specialised $[4Fe\text{-}4S]^{+}$ cluster to cleave non-activated carbon hydrogen (C-H) bonds (Hayden *et al.*, 1992; Miller *et al.*, 2000; Reed & Cronan, 1993; Sofia *et al.*, 2001). All members of this superfamily possess a strictly conserved Cx_3Cx_2C motif; the three cysteine residues within this motif nucleate three of the Fe atoms within the $[4Fe\text{-}4S]^{+}$ cluster, and the fourth Fe atom is bound to SAM (Frey *et al.*, 2008). Radical SAM enzymes catalyse the reductive cleavage of SAM to produce $[4Fe\text{-}4S]^{2+}$-Met and a 5'-deoxyadenosyl radical (5'-dA$^{\bullet}$). 5'-dA$^{\bullet}$ abstracts hydrogen atoms linked to non-activated carbon hydrogen (C-H) bonds. In the case of LipA, the 5'-dA$^{\bullet}$ abstracts hydrogen atoms linked to C6 and C8 of octanoamide, forming C6- and C8 alkyl radicals (Douglas *et al.*, 2006). It was established that the production of one molecule of lipoamide from octanoamide requires two SAM molecules (Cicchillo *et al.*, 2004a). Interestingly, LipA also possesses a second LipA -specific cysteine-rich motif Cx_4Cx_5C (Cicchillo & Booker, 2005; Frey *et al.*, 2008). A second $[4Fe\text{-}4S]^{+}$ cluster is bound to this motif; the C6 and C8 alkyl radicals formed from hydrogen abstraction by the action of 5'-dA$^{\bullet}$ attack μ-sulphido atoms attached to the Cx_4Cx_5C cluster, and the final result is production of a lipoyl-group (Booker *et al.*, 2007; Cicchillo *et al.*, 2004b) (see Figure 1.9).

LipA shares some mechanistic features with biotin synthase (BioB). LA metabolism and biotin metabolism are closely related in terms of reaction mechanisms carried out by the respective enzymes (Reche, 2000). Biotin is an essential cofactor of several enzyme complexes including acetyl-CoA carboxylase and pyruvate carboxylase (Nikolau *et al.*, 2003; Pacheco-Alvarez *et al.*, 2002). Mammals are not able to synthesise biotin, and as such rely upon uptake of this vitamin, which is synthesised in bacteria, plants and some fungi. BioB belongs to the radical SAM superfamily and catalyses the generation of biotin from desthiobiotin (Lotierzo *et al.*, 2005; Sofia *et al.*, 2001). During this reaction, two C-

H bonds are cleaved and one sulphur atom is inserted. Similar to LipA, BioB requires two SAM molecules to form one molecule of biotin (Shaw *et al.*, 1998). The protein also contains a second cysteine-rich motif (Ugulava *et al.*, 2001a), which binds to a $[2Fe\text{-}2S]^{+}$ cluster that contributes a sulphur atom to the formation of biotin (Jarrett, 2005; Ugulava *et al.*, 2001b).

1.4.3.2Octanoic acid and lipoic acid transferases

The enzymes LipB and LplA possess a moderate degree of similarity on the amino acid sequence level to biotin protein ligase (BPL) (Reche, 2000). LipB, LplA and BPL all possess a strictly conserved lysine residue that was thought to be involved in the transfer of LA/biotin to the ε-amino group of a conserved lysine residue on the apoprotein, and this was indeed shown to be the case for LplA and BPL (Bagautdinov *et al.*, 2005; McManus *et al.*, 2005).

However, the reaction catalysed by LipB differs in several respects from those carried out by LplA and BPL. Firstly, *E. coli* LipB (*Ec*LipB) uses ACP-bound OA or LA as substrate and is unable to transfer free OA or LA to apo-E2 or apo-H-protein (Jordan & Cronan, 1997b; Jordan & Cronan, 2003) (see Figure 1.9). Although *Ec*LipB is able to transfer lipoyl-ACP as well as octanoyl-ACP, the fact that *E. coli* LipA (*Ec*LipA) has a significantly higher preference for protein-bound octanoamide (Zhao *et al.*, 2003) dictates that the substrate of LipB under physiological conditions is likely to be octanoyl-ACP. Secondly, unlike LplA and BPL, *Ec*LipB was shown to catalyse octanoylation of apoproteins in a two-step manner, whereby octanoyl-ACP is covalently attached to LipB as an acyl-intermediate before being transferred to the apo-lipoyl-domain (Zhao *et al.*, 2005). Crystal structure data of LipB from *Mycobacterium tuberculosis* has revealed the nature of this two-step reaction, and indicates that *M. tuberculosis* LipB (*Mt*LipB) is a cysteine-lysine acyltransferase (Ma *et al.*, 2006). In the first step of catalysis, octanoyl-ACP forms a thioester linkage with a conserved cysteine residue in *Mt*LipB, and the conserved lysine residue in LipB activates the ε-amino group of the target lysine residue in the apoprotein by deprotonation (Ma *et al.*, 2006). In the second step, *Mt*LipB catalyses the nucleophilic attack of the activated ε-amino group of the target (apoprotein) lysine residue; the result is cleavage of the thioester linkage between the octanoyl-moiety and *Mt*LipB, and the formation of an amide bond between the octanoyl-moiety and the apoprotein (Ma *et al.*, 2006).

The reaction mechanisms of LplA and BPL are similar to one another, yet very different to that of LipB. *E. coli* LplA (*Ec*LplA) is able to use octanoyl-ACP or lipoyl-ACP as substrates, which is based on evidence that an *E. coli* line lacking the *lipB* gene is able to proliferate (albeit poorly) without the provision of exogenous LA (Jordan & Cronan, 2003). Free LA is the optimal substrate for *Ec*LplA (Green *et al.*, 1995; Jordan & Cronan, 2003), and the lipoylation reaction occurs in two steps. The first step involves activation of LA by ATP to form the lipoyl-AMP intermediate, and the release of pyrophosphate (PPi). Subsequently, the activated intermediate is ligated onto the ε-amino group of the target lysine residue in the apoprotein (Fujiwara *et al.*, 2005; Kim do *et al.*, 2005; McManus *et al.*, 2005) (see Figure 1.9).

In mammals, salvage of LA necessitates the sequential action of two enzymes. The first step is activation of LA by ATP or guanosine triphosphate (GTP) to form lipoyl-AMP or lipoyl-GMP (Fujiwara *et al.*, 2007), respectively, which is carried out by a lipoic acid activating enzyme that was shown to be the ACSM (Fujiwara *et al.*, 2001). Secondly, the lipoyl-AMP intermediate is ligated to apoproteins by LT (Fujiwara *et al.*, 1997a; Fujiwara *et al.*, 1999; Fujiwara *et al.*, 2007).

LplA, BPL and LT share a conserved mechanism of ligation of the lipoyl/biotinoyl-moiety to apoproteins from a tightly-bound acyl-adenylate intermediate (Fujiwara *et al.*, 2005; Fujiwara *et al.*, 2007; Kim do *et al.*, 2005; McManus *et al.*, 2005; Reche & Perham, 1999; Reche, 2000; Wilson *et al.*, 1992). Nevertheless, LT is unable to catalyse the first step in the reaction carried out by LplA; LA activation (Fujiwara *et al.*, 2007). Crystal structure data of *Ec*LplA (Fujiwara *et al.*, 2005), the archaebacterium *Thermoplasma acidophilum* LplA (Kim do *et al.*, 2005; McManus *et al.*, 2005), *E. coli* BPL (Wilson *et al.*, 1992) and *Bos taurus* LT (*Bt*LT) (Fujiwara *et al.*, 2007) have helped to better understand how the similarities between these enzymes contribute to common modes of action, and how the differences between these enzymes permit substrate specificity (Kim do *et al.*, 2005; McManus *et al.*, 2005). The *N*-terminal domains of *Ec*LplA, *Ta*LplA and *Bt*LT were shown to possess three highly conserved sequence motifs, which contain key residues involved in the formation of the lipoyl-AMP binding pocket (see Figure 3.4) (Fujiwara *et al.*, 2005; Fujiwara *et al.*, 2007; Kim do *et al.*, 2005; McManus *et al.*, 2005). LA is bound within a hydrophobic cavity, and it is believed that the binding and activation of LA with ATP occurs at the same site, since no large conformational changes were observed after LA binding and activation, and ATP was located in the same position as AMP (Fujiwara *et al.*, 2007; Kim do *et al.*,

2005). A conserved lysine residue is present in *Ec*LplA, *Ta*LplA and *Bt*LT, which directly interacts with the carboxyl-group of LA (Fujiwara *et al.*, 2005; Fujiwara *et al.*, 2007; Kim do *et al.*, 2005; McManus *et al.*, 2005).

Based on size, two types of LplA were uncovered; a shorter form found in *T. acidophilum* (circa 260-270 amino acids) (Kim do *et al.*, 2005; McManus *et al.*, 2005) and a longer form found in *E. coli* (circa 330-340 amino acids) (Fujiwara *et al.*, 2005). Of the two structural reports on *T. acidophilum* LplA (*Ta*LplA) (Kim do *et al.*, 2005; McManus *et al.*, 2005), one of them illustrated that recombinant *Ta*LplA is able to activate LA, yet is unable to ligate it to a lipoate-acceptor protein (McManus *et al.*, 2005). Interestingly, *T. acidophilum* possesses the conserved *N*-terminal domain that has been shown to be involved in activation of LA, yet completely lacks the *C*-terminal domain found in *Ec*LplA (McManus *et al.*, 2005), thus explaining why it is considerably shorter than *Ec*LplA. The authors hypothesised that salvage of LA in *T. acidophilum* necessitates the sequential action of two enzymes, in a similar fashion to LT (McManus *et al.*, 2005).

Unlike *Ta*LplA however, *Bt*LT does not lack a *C*-terminal domain, and is unable to catalyse the activation of LA (Fujiwara *et al.*, 2007). As discussed above, the *N*-terminal domain of LplA has been shown to be important in binding and activation of LA (Fujiwara *et al.*, 2005; McManus *et al.*, 2005), and given that the LT *N*-terminal domain contains conserved motifs that permit the binding of lipoyl-AMP (Fujiwara *et al.*, 2007), one might expect *Bt*LT to have the capacity to activate LA. Interestingly, upon protein purification of recombinant *Bt*LT, LA was bound within the active site, yet LA has never been reported to be bound to purified recombinant *Ec*LplA or *Ta*LplA (Fujiwara *et al.*, 2007). The authors that solved the *Bt*LT crystal structure hypothesise that the reason *Bt*LT is unable to activate LA is due to its *C*-terminal domain (Fujiwara *et al.*, 2007). The *C*-terminal domain of *Bt*LT has the same overall fold as that of *Ec*LplA, but importantly, they are rotated by approximately 180° relative to each other (Fujiwara *et al.*, 2007). Both *Ec*LplA and *Bt*LT proteins possess a loop linking beta strand 12 and alpha helix 6 (β12-α6 loop, or the 'adenylate binding loop'), and this loop is unordered in *Ec*LplA. In *Bt*LT however, the *C*-terminal domain anchors the adenylate binding loop to the active site, which restricts access of ATP to the active site (Fujiwara *et al.*, 2007).

Given that LplA/LT and BPL enzymes catalyse the post-translational ligation of LA or biotin, respectively, to certain apoproteins, the question arises as to how this

event is so specific. An interesting study showed that a single point mutation introduced into the active site of *Ec*BPL resulted in promiscuous biotin addition (biotinylation) to lysine residues on proteins that were not ordinarily biotinylated (Choi-Rhee *et al.*, 2004; Cronan, 2005). The most likely explanation for this "promiscuous biotinylation" was found to be due to the activated biotin intermediate (biotinoyl-AMP) escaping more easily from the active site of mutant *Ec*BPL and biotinylating lysine residues on non-specific substrates by chemical (not enzymatic) acylation, in a concentration-dependent manner (Choi-Rhee *et al.*, 2004; Cronan, 2005).

In order to avoid lipoylation/biotinylation of the wrong proteins, activated lipoyl-AMP and biotinoyl-AMP intermediates are bound tightly to the active site of LplA/LT and BPL, respectively. Also, the exact positioning of the target lysine residue on the apoprotein is pivotal to being recognised by lipoylating/biotinylating enzymes (Reche *et al.*, 1998; Wallis & Perham, 1994). The target lysine residue is located within a lipoyl-domain or biotinoyl-domain, whereby it is found in the D**K**A/V motif or M**K**M motif, respectively. However, the most important factor in LplA/LT and BPL substrate specificity is the subtle structural differences found between the lipoyl/biotinoyl domains (Jones & Perham, 2008).

1.5 Aims of project

a. Determine whether genes coding for subunits of α-KADHs and of the GCC are present within the *L. major* genome, and subsequent *in silico* analyses of the predicted proteins.

b. Assess whether α-KADH E2 subunits and the H-protein of the GCC are lipoylated *in vivo*.

c. Determine whether genes coding for enzymes involved in LA metabolism are present within the *L. major* genome, and subsequent *in silico* analyses of the predicted proteins.

d. Assess whether the proteins predicted to be involved in LA metabolism are functional.

e. Confirmation that the proteins predicted to be involved in LA metabolism localise to the mitochondrion.

f. Attempt gene replacement of one gene from the LA biosynthesis pathway and one from the LA salvage pathway, in order to gain an insight into the level of redundancy between the two metabolic pathways.

This thesis is divided into three results chapters covering points a-d (Chapter 3), points e-f for the LA salvage pathway (Chapter 4) and points e-f for the LA biosynthesis pathway (Chapter 5). The results and their implications relative to published work will be discussed in Chapter 6.

2 Materials and Methods

2.1 Biological and chemical reagents

General chemicals were purchased from either Sigma or Fisher Scientific.

Abgene	2x Reddymix
BD Biosciences	α-His-tag monoclonal antibody
Bio-Rad	Precision plus all blue standards, Bradford protein assay reagent, econo-columns
Biosera	Lipid-containing- and lipid-depleted foetal calf serum (FCS)
Calbiochem	α-Lipoic acid polyclonal antibody
Eurogentec	Custom DNA oligonucleotides, custom polyclonal antibody production
Fisher Scientific	10x phosphate buffered saline (PBS)
GE Healthcare	Alk phos direct Southern blot kit, N^+ membrane, ECL films
IBA	α-Strep-tag monoclonal antibody, strep-tactin sepharose, d-desthiobiotin, strep-tactin regeneration buffer
ICN Biomedicals Inc.	R-lipoic acid
Invitrogen	Accuprime *Pfx* supermix, zero blunt cloning kit, chemically-competent TOP10 *E. coli*, sybr safe, low melting point agarose, HOMEM medium
Melford	Ampicillin, carbenicillin, isopropyl-β-D-thiogalactopyranoside (IPTG), dithiothreitol (DTT)
Millipore	Immobilon western detection kit, Centricon Plus-20 centrifugal filter device
Molecular Probes	Mitotracker CMXRos
New England Biolabs	All restriction endonucleases, 1 kb DNA ladder
Novagen	Bugbuster protein extraction reagent, benzonase nuclease, chemically-competent *E. coli* BLR (DE3) cells
PAA	G418, puromycin, hygromycin
Promega	α-Mouse IgG (HRP-conjugated) antibody, α-rabbit IgG (HRP-conjugated) antibody
Qiagen	Qiaprep spin DNA miniprep kit, Qiaquick gel extraction kit, hi-speed maxi kit, Qiaamp DNA mini kit
Roche	Rapid ligation kit
Sigma	Acrylamide 30 %, Ponceau-S solution, L-polylysine, bovine serum albumin (BSA), lysozyme, phenylmethylsulfonylfluoride (PMSF), pepstatin A, E-64, 1, 10-phenanthroline, 4' -6-diamidino-2-phenylindole dihydrochloride (DAPI), tetracycline hydrochloride, digitonin, peanut agglutinin, propidium iodide
Stratagene	Quickchange lightning mutagenesis kit
Schleicher Schuell	Protran nitrocellulose membrane

Thermo Scientific	Restore western blot stripping buffer, aminolink coupling resin, IgG binding buffer, IgG elution buffer
Upstate	Mouse α-*T. brucei* EF1α antibody

2.1.1 Buffers, solutions and media

General buffers

1x PBS	140 mM NaCl, 3 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.4
1x TAE	40 mM Tris-acetate, 1 mM EDTA, pH 8.0

DNA analyses

DNA loading dye	0.25 % (w/v) bromophenol blue, 0.25 % (w/v) orange-G, 40 % (w/v) sucrose
Depurination solution	0.25 N HCl
Denaturation solution	1.5 M NaCl and 0.5 M NaOH
20x SSC	0.3 M tri-sodium citrate, 3 M NaCl, pH 7.0

Protein analyses

6x loading buffer	62.5 mM Tris/HCl pH 6.8, 2% (w/v) SDS, 10% (v/v) glycerol, 0.001% (w/v) bromophenol blue, 5% (v/v) 2-mercaptoethanol
1x running buffer	25 mM Tris, 192 mM glycine, 0.1 % (w/v) SDS
1x MOPS buffer	50 mM 3-[N-morpholino] propane sulphonic acid, 50 mM Tris, 3.5 mM SDS, 1 mM EDTA
Coomassie stain	40 % (v/v) methanol, 10 % (v/v) acetic acid, 0.1 % (w/v) coomassie brilliant blue R-250
Destain	20 % (v/v) methanol, 10% (v/v) acetic acid
Towbin transfer buffer	25 mM Tris, 192 mM glycine, 20 % (v/v) methanol
Buffer W	100 mM Tris-HCl, 150 mM NaCl, pH 8.0
Buffer E	100 mM Tris-HCl, 150 mM NaCl, 2.5 mM desthibiotin, pH 8.0
Quenching buffer	1 M Tris-HCl, pH 7.4
Wash buffer	1 M NaCl
Neutralisation buffer	1 M Tris-HCl, pH 9.0

Bacteria culture

Luria-Bertani (LB) medium	10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 5 g L^{-1} NaCl (add 15 g L^{-1} agar for LB plates)
M9 minimal medium	10x M9 salt stock: 58 g L^{-1} Na_2HPO_4, 30 g L^{-1} Na_2HPO_4, 5 g L^{-1} NaCl, 10 g L^{-1} NH_4Cl

	Working solution: 1x M9 salts, 0.2% (w/v) glucose, 1 mM $MgSO_4$, 0.001 % (w/v) thiamine (add 15 g L^{-1} agar for M9 plates)
Ampicillin	100 mg ml^{-1} in ddH_20; stored at -20℃
Carbenicillin	100 mg ml^{-1} in ddH_20; stored at -20℃
Kanamycin	50 mg ml^{-1} in ddH_20; stored at -20℃
Tetracycline	5 mg ml^{-1} in ethanol; stored at -20℃
TfbI	100 mM RbCl, 50 mM $MnCl_2$-$4H_2O$, 30 mM KNa, 10 mM $CaCl_2$-$2H_2O$, 15 % glycerol
TfbII	10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2$-$2H_2O$, 15 % glycerol

***L. major* culture**

HOMEM + 10 % FCS:	Modified eagle's medium supplemented with 10 % heat-inactivated FCS (Berens *et al.*, 1976). HOMEM is purchased from GIBCO (Invitrogen), and FCS from Biosera
Electroporation buffer	120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4, 25 mM HEPES, 2 mM EDTA, 2 mM $MgCl_2$, pH 7.6
Freezing solution	30 % (v/v) glycerol in FCS
TELT buffer	50 mM Tris-HCl, 62.5 mM EDTA; 2.5 M LiCl; 4% (v/v) Triton X-100, pH 8.0
Lysis buffer	0.25 M sucrose, 0.25 % Triton x-100, 10 mM EDTA, 10 μM EDTA, 10 μM E-64, 2 μM 1, 10-phenanthroline, 4 μM pepstain A, 1 mM PMSF
TSE buffer	0.025 M Tris-HCl, 0.25 M Sucrose, 1 mM EDTA, 10 μM E-64, 2 μM 1, 10 phenanthroline, 2 μM Pepstatin A, 1 mM PMSF, 0.3 – 5.0 mg ml^{-1} digitonin, pH 7.4

2.1.2 Bacteria strains

TOP10 (Invitrogen)
F$^-$ *mcrA*Δ(*mrr-hsd*RMS-*mcr*BC) ø 80*lac*ZΔM15Δ*lac*X74 *rec*A1 *arc*D139 Δ(*arcleu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nub*G

BLR(DE3) (Novagen)
F$^-$ *ompT hsdS*$_B$(r$_{B-}$ m$_{B-}$) *gal dcm* (DE3) Δ(srl-recA)306::Tn*10* (Tet^R)

XL10-Gold ultracompetent (Stratagene)
Tet^r Δ*(mcrA)183* Δ*(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F' *proAB lacI*^{q}*ZΔM15 Tn10* (Tet^r) Amy Cam^r]^a

KER184 (*lipB*$^-$) (Vanden Boom *et al.*, 1991)
F$^-$ *rpsL lipB182*::Tn*1000*dKn

KER176 (*lipA*$^-$) (Vanden Boom *et al.*, 1991)
F$^-$ *rpsL lipA150*::Tn*1000*dKn

2.1.3 *L. major* strain

MHOM/IL/80/Friedlin

2.1.4 Oligonucleotide primers

2.1.4.1For transfection in *L. major*

Lm1	GCGC**ATCGAT**ATGCTGCGCTGCTGCTCTGC (*Cla*I)
Lm2	GCGC**ACCGGT**TCACGCAATCGCAGTACCTGCG (*Age*I)
Lm3	GCGC**AAGCTT**GTTGTGTGTTCGCTGGCAGCG (*Hind*III)
Lm4	GCGC**GTCGAC**GCTGACGTTCACGAG (*Sal*I)
Lm5	GCGC**CCCGGG**AGAGGAGGTGCATGTGCAATGAGC (*Sma*I)
Lm6	GCGC**AGATCT**CAAGGGTGCGCGCCTATCTC (*Bgl*II)
Lm13	GCGC**ATCGAT**ATGTGGCAGACTGTCGTGCG (*Cla*I)
Lm14	GCGC**ACCGGT**TTAGGTCGCGATGTCGAAC (*Age*I)
Lm15	GCGC**AAGCTT**CGACTTCCTCGGCTCCCAGC (*Hind*III)
Lm16	GCGC**GTCGAC**GGTTGCACACGATCCAGTTGC (*Sal*I)
Lm17	GCGC**CCCGGG**TCCGTCGAGCCTTCGCGTCG (*Sma*I)
Lm18	GCGC**AGATCT**CGTAGGGGGGAGGAGTTGCC (*Bgl*II)
Lm19	GCGC**CATATG**CTGCGCTGCTGCTCTGCTCTG (*Nde*I)
Lm20	GCGC**GGTACC**CGCAATCGCAGTACCTGCGTTG (*Kpn*I)
Lm21	GCGC**GGTACC**GGGCGTTGAAATGCCTGCAG (*Kpn*I)
Lm22	GCGC**CATATG**AAGGCATTTTTCATCGGCAAAC (*Nde*I)
Lm23	GCGC**CATATG**TGGCAGACTGTCGTGCGGCG (*Nde*I)
Lm24	GCGC**AGATCT**GGTCGCGATGTCGAACACATT (*Bgl*II)
Lm64	GTGCCGTTTTCGCAGACGAGGG
Lm65	CCCTCGTCTGCGAAAACGGCAC

2.1.4.2For recombinant protein expression

Lm25	ATGGTA**GGTCTC**A**AATG**CTGCGCTGCTGCTCTGCTCTG
Lm26	ATGGTA**GGTCTC**A**AATG**TCTGCTCTGATGTGCCCGACG
Lm27	ATGGTA**GGTCTC**A**AATG**CCGGTTGCTGCTGCTGCCGC
Lm28	ATGGTA**GGTCTC**A**AATG**CAGTCGGACAAGACGGGCATG
Lm29	ATGGTA**GGTCTC**A**GCGCT**CGCAATCGCAGTACCTGCGTTG
Lm30	ATGGTA**GGTCTC**A**AATG**AAGGCATTTTTCATCGGCAAACGC
Lm31	ATGGTA**GGTCTC**A**AATG**AGTGTGCGCCGCGGCGCG
Lm32	ATGGTA**GGTCTC**A**AATG**GGCGCGTCGAAGCTGCCGC
Lm33	ATGGTA**GGTCTC**A**GCGCT**GGGCGTTGAAATGCCTGCAGTG
Lm34	ATGGTA**GGTCTC**A**AATGT**GGCAGACTGTCGTGCGGCG
Lm35	ATGGTA**GGTCTC**A**AATG**CCCACGTCCCTCGCAGCCTT
Lm36	ATGGTA**GGTCTC**A**AATG**GACTCGAACCTTGTCGTCGCTG
Lm37	ATGGTA**GGTCTC**A**GCGCT**GGTCGCGATGTCGAACACATTC

All primers used for PCR of DNA fragments to be used in recombinant protein expression contain directional *Bsa*I restriction endonuclease sites.

2.1.5 Antibodies

1^{ary} antibody	Animal	Clonality	Dilution	Source	Reference
α-Lipoic acid	Rabbit	Polyclonal	1/6,000	Calbiochem	Gunther *et al.* (2007)
α-Strep-tag	Mouse	Monoclonal	1/7,500	IBA	N/A
α-His-tag(6x)	Mouse	Monoclonal	1/10,000	BD Biosciences	N/A
α-GFP	Mouse	Monoclonal	1/5,000	Roche	N/A
α-*Lm* LIPA	Rabbit	Polyclonal	1/20	Ryan Bissett (Eurogentec)	N/A
α-*Lm* LIPB	Rabbit	Polyclonal	1/1000	Ryan Bissett (Eurogentec)	N/A
α-*Lm* CS	Rabbit	Polyclonal	1/7,500	Dr Rod Williams	N/A
α-*Tb* EF1α	Mouse	Polyclonal	1/10,000	Upstate	Besteiro *et al.* (2008)
α-*Lm* HASPB	Rabbit	Polyclonal	1/4,000	Prof. Debbie Smith	Flinn *et al.* (1994)

2^{ary} antibody	Animal	Clonality	Dilution	Source	Reference
α-Mouse (H + L), HRP	Goat	Polyclonal	1/10,000	Promega	N/A
α-Rabbit (H + L), HRP	Goat	Polyclonal	1/10,000	Promega	N/A

Table 2.1 Primary and secondary antibodies and their dilutions
Table of all primary and secondary antibodies used in this thesis. The 'Animal' field indicates the origin of the antibodies. '*Lm*' and '*Tb*' indicate that the antigens used for antibody production are of *L. major* or *T. brucei* origin. 'HRP' is an abbreviation for 'horseradish peroxidise', to which secondary antibodies are conjugated. 'N/A' is an acronym for 'not applicable.'

2.2 *L. major* cell culture

2.2.1 *L. major* promastigote culture

L. major cultures were initiated by purification of amastigotes from a BALB/c mouse by Mrs Denise Candlish (University of Glasgow) and differentiated into promastigotes by inoculation into HOMEM medium supplemented with 10 % (v/v) heat-inactivated FCS, which will be referred to as HOMEM + 10 % FCS. In order to maintain *L. major* promastigotes, cultures were incubated at 25 ℃ and sub-passaged to a density of 10^5 parasites ml^{-1} on a weekly basis.

Parasite numbers were determined by transferring the parasites into 1 % formaldehyde in PBS, and counting them using a Neubauer haemocytometer with an objective magnification of 400 x on an inverted light microscope (Zeiss Wetzlar, Germany). Growth rates through promastigote development were examined by determining parasite numbers daily for up to 120 h, starting from a parasite density of 5 x 10^5 cells ml^{-1}.

When necessary, antibiotics were added to transgenic promastigote cell lines as follows: G418 at 10- or 50 μg ml^{-1}; hygromycin B at 50 μg ml^{-1}; neurseothricin at 75 μg ml^{-1}; puromycin at 100 μg ml^{-1}.

Stabilates of parasites were generated by transferring 500 μl of exponentially growing cells into 500 μl of 30 % glycerol (v/v) in FCS and transferring the mixture into – 80 ºC for 24 h before the stabilate was transferred into liquid nitrogen for long-term storage.

2.2.2 Isolation of concentrated *L. major* metacyclic promastigotes

Isolation of metacyclic promastigotes was carried out based upon a negative selection (of procyclic promastigotes) protocol described by Sacks (Sacks *et al.*, 1985). Briefly, *L. major* stationary culture promastigotes were incubated with sterile peanut agglutinin dissolved in PBS, pH 7.2 at a concentration of 100 μg ml^{-1} and incubated for 30 min at room temperature. The parasites were then centrifuged for 5 min at 100 *g* at room temperature in order to fractionate non-metacyclic promastigotes (pellet fraction) from metacyclic promastigotes (supernatant fraction).

2.2.3 Harvest, lysis and fractionation of parasites

2.2.3.1Isolation of genomic DNA from *L. major*

Genomic DNA (gDNA) was isolated from parasites by one of two methods. For PCR analyses, gDNA was isolated by phenol-chloroform extraction (as described in (Sambrook *et al.*, 1989)) after lysis with TELT buffer (see Section 2.1.1). For Southern blotting, gDNA was isolated using the QIAamp DNA mini kit according to instructions provided by the manufacturer (Qiagen).

2.2.3.2Isolation of protein from *L. major*

Protein was extracted from parasites by cell lysis in lysis buffer (see Section 2.1.1). Soluble- and pellet fractions were separated by centrifugation for 30 min at 13,000 rpm (Fisher Scientific accuSpin MicoR with 24-plate motor) at 4 ℃. The protein concentration in the soluble fraction was determined using the Bradford

method with BSA as a standard (Bradford, 1976). Then, 6x sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) loading dye was added and aliquots were denatured at 100 ºC for 5 min before the samples were stored at – 20 °C.

2.2.3.3Cellular pre-fractionation

Parasites were permeabilised with increasing quantities of digitonin in order to pre-fractionate (or concentrate) different organelles. Previous digitonin-titration experiments in *L. mexicana* showed that cytosolic-, glycosomal- or mitochondrial protein markers could be detected by western blotting (see Section 2.5.3) when cells were treated with 0.3 mg ml^{-1}, 1.5 mg ml^{-1} or 3.0 mg ml^{-1} to 5.0 mg ml^{-1} digitonin, respectively (Leroux *et al.*, 2006). The same digitonin concentrations were used here. Briefly, 2 x 10^9 stationary phase parasites were harvested by centrifugation for 15 min at 1000 *g* at room temperature. Cells were washed twice at room temperature in TSE buffer (TSEB) (see Section 2.1.1) (Leroux *et al.*, 2006). Cells were then resuspended in 200 µl TSEB supplemented with the lowest concentration of digitonin (dissolved in sterile water) and incubated for 5 min at room temperature. 40 µl of 0.3 M sucrose was added to the cells, which were then centrifuged for 10 min at 18,000 *g* at 4 ℃. The pellet fraction was subsequently treated in the same way as described above except the second lowest concentration of digitonin was added. Soluble fractions from each digitonin-titration step were used to extract soluble protein for SDS-PAGE and western blotting.

2.2.4 Targeted gene replacement in *L. major*

Gene replacement by double homologous recombination is the method of choice for gene knockout studies in *L. major* (Cruz & Beverley, 1990; Cruz *et al.*, 1991). The plasmids used consist of an antibiotic resistance gene (for selection of transgenic parasites), flanked by regions of DNA specific to the gene under study. The knockout plasmid is linearised either side of the flanking regions of the target gene to give rise to the knockout cassette, which is then introduced into *Leishmania* by electroporation (see Section 2.2.8). The fact that *Leishmania* is a diploid organism means that two rounds of homologous recombination with different selectable markers are necessary in order to ablate both endogenous alleles of the gene (Cruz & Beverley, 1990; Cruz *et al.*, 1991).

2.2.4.1Gene knockout

In order to create an *L. major LIPA* (*LmjLIPA*) gene knockout construct, the 534 bp 5' flanking region (using primer pairs Lm3-Lm4) and the 550 bp 3' flanking region (using primer pairs Lm5-Lm6) were cloned into plasmid into pGL158 (see Figure 2.1 and Section 2.1.4.1) to give rise to *LIPA-SAT*. The *LPLA-SAT* gene knockout construct was created by cloning the 516 bp 5' flanking region (using primer pairs Lm15-Lm16) and the 545 bp 3' flanking region (using primer pairs Lm17-Lm18) into the pGL158 plasmid (see Figure 2.1 and Section 2.1.4.1). In both cases, 5' and 3' flanking regions were cloned into pGL158 by directional cloning with *Hind*III/*Sal*I and *Sma*I/*Bgl*II, respectively (see Section 2.4.2).

The pGL158 plasmid contains the neurseothricin-resistance gene; streptothricin-acetyl-transferase (*SAT*) (Joshi *et al.*, 1995). *LIPA-HYG* and *LPLA-HYG* were generated by replacing the *Spe*I/*Bam*HI cassette containing the neurseothricin-resistance gene with a *Spe*I/*Bam*HI cassette containing the hygromycin-resistance gene; hygromycin B phosphotransferase (*HYG*) (Cruz *et al.*, 1991), which was derived from plasmid pGL345 (see Figure 2.2).

Knockout cassettes were created by *Hind*III/*Bgl*II linearisation of knockout constructs; the cassettes were gel purified, precipitated into a smaller volume and then transfected into parasites (see Section 2.2.8).

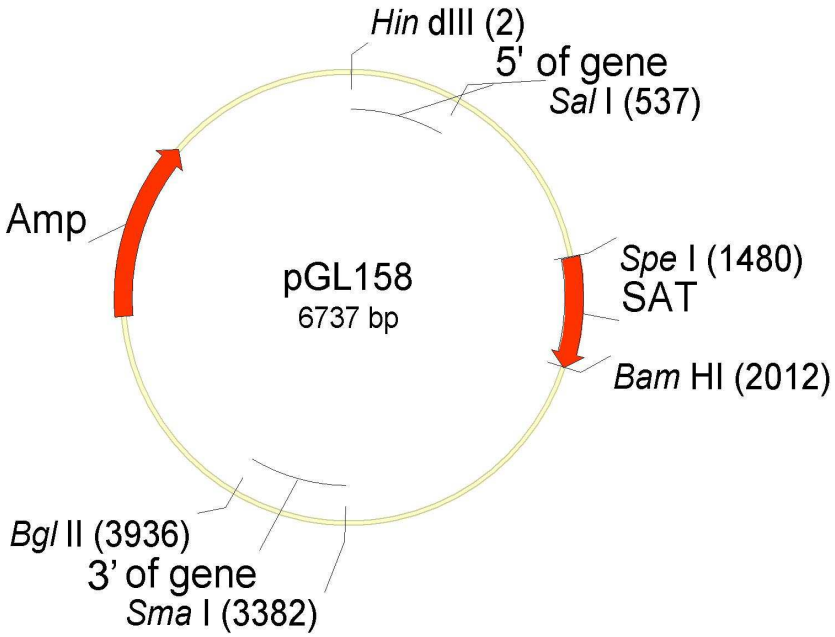


Figure 2.1 pGL158 plasmid
The pGL158 plasmid contains the following features: 5' of gene, the 5' FR of the gene of interest; 3' of gene, the 3' FR of the gene of interest; SAT, streptothricin acetyltransferase gene conferring resistance to neurseothricin; Amp, ampicillin-resistance gene. Restriction endonucleases are illustrated, and numbers in brackets mark the cut sites of the restriction endonucleases.

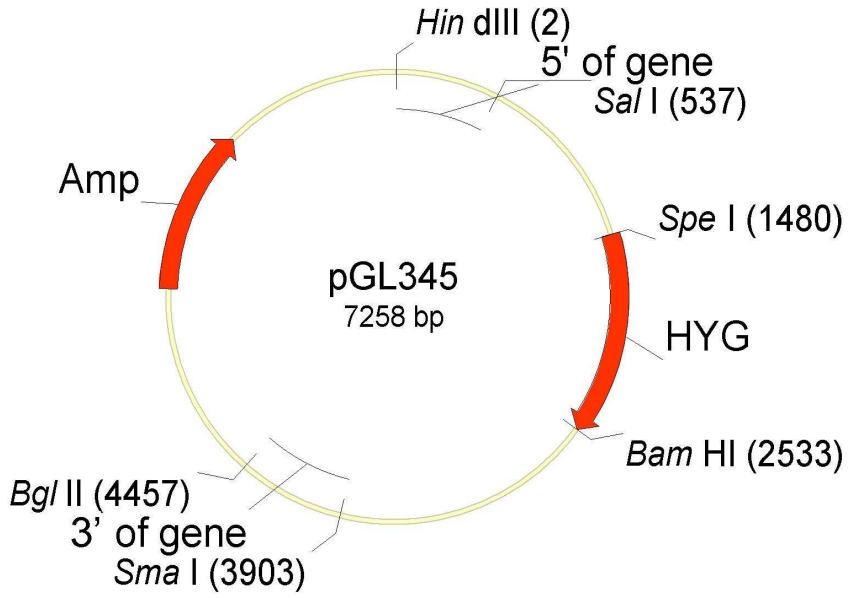


Figure 2.2 pGL345 plasmid
The pGL345 plasmid contains the following features: 5' of gene, the 5' FR of the gene of interest; 3' of gene, the 3' FR of the gene of interest; SAT, streptothricin acetyltransferase gene conferring resistance to neurseothricin; Amp, ampicillin-resistance gene. Restriction endonucleases are illustrated, and numbers in brackets mark the cut sites of the restriction endonucleases.

2.2.5 Re-expression of the target gene

In order to assess the functionality of genes in *Leishmania*, it is necessary to create a parasite line re-expressing the gene of interest, before or after ablating the endogenous gene copies. The pGL631 (or pRB vector) (see Figure 2.3) was chosen for this purpose, which is designed to replace (knock-in) one of the twelve copies of the *18S SSU ribosomal rRNA* (*rRNA*) gene (Misslitz *et al.*, 2000). Additionally, a re-expression construct that integrates is more likely to mimic expression levels of the endogenous gene, compared to the (typically) very high expression/over-expression observed from episomal expression constructs (see Section 5.4.1).

The pRB vector (see Figure 2.3) contains DNA sequences corresponding to 5' and 3' regions within the *18S SSU rRNA* gene, in order to permit replacement of one of the 12 *18S SSU rRNA* gene copies by homologous recombination (Misslitz *et al.*, 2000). The puromycin resistance gene, puromycin *N*-acetyltransferase (PAC) (Freedman & Beverley, 1993), is upstream of the intergenic region of the *L. mexicana* cysteine peptidase B 2.8 (*CPB 2.8*) gene, in order to permit high expression of the resistance marker if the transfected line is to be used in amastigotes (Misslitz *et al.*, 2000). In the pRB vector, open reading frames (ORFs) to be expressed in parasites are located downstream of a splice acceptor site (SAS), which is necessary in order to permit *trans*-splicing of the pre-messenger RNA (mRNA) into monocistronic units (Agabian, 1990).

LIPA-pRB and *LPLA*-pRB re-expressor constructs were created by directional cloning of the respective ORFs into the pRB vector, using *Cla*I/*Age*I. The *LIPA* and *LPLA* genes were amplified from *L. major* gDNA using primer pairs Lm1-Lm2 and Lm13-Lm14, respectively (see Section 2.1.4.1), and then cloned into plasmid pRB (also referred to as pGL631) (see Figure 2.3).

Knockout cassettes were created by *Pac*I/*Pme*I linearisation of knockout constructs; the cassettes were gel purified, precipitated into a smaller volume and then transfected into parasites (see Section 2.2.8).

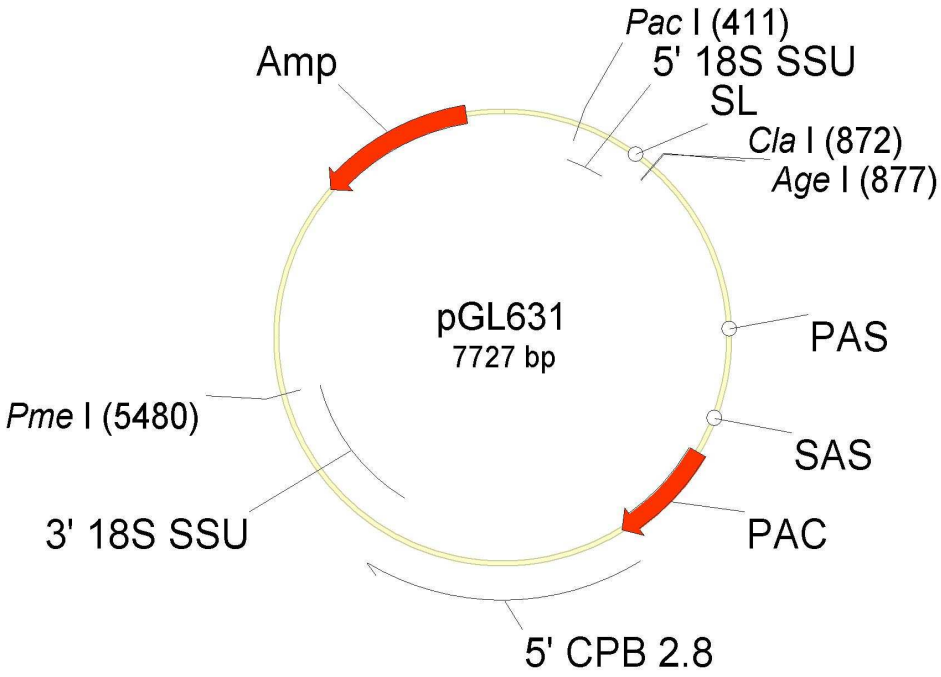


Figure 2.3 pGL631 (pRB) plasmid
The pGL631/pRB plasmid contains the following features: 5' 18S SSU, the 5' FR of the *18S SSU* locus; SL, splice leader sequence; PAS, polyadenylation site; SAS, splice acceptor site; 5' CPB 2.8, intergenic region of the *L. mexicana* cysteine peptidase B 2.8 (*CPB 2.8*) gene; PAC, puromycin *N*-acetyltransferase gene conferring resistance to puromycin; 3' 18S SSU, the 3' FR of the *18S SSU* locus; Amp, ampicillin-resistance gene. Restriction endonucleases are illustrated, and numbers in brackets mark the cut sites of the restriction endonucleases. The gene of interest to be expressed in *L. major* is directionally cloned into pGL631 using *Cla*I and *Age*I restriction sites.

2.2.6 Episomal gene expression in *L. major*

Stable extrachromosomal expression of genes in *Leishmania* is possible using a series of expression vectors, named pNUS vectors, which have been designed for the trypanosomatids *Crithidia* and *Leishmania* (Tetaud *et al.*, 2002). These vectors contain a rRNA promoter, intergenic sequences from the *Crithidia fasciculata* (*C. fasciculata*) phosphoglycerate kinase (*PGK*) locus at the 5' and 3 ends of the target gene, and the 3' untranslated sequence of the *C. fasciculata* glutathionyl-spermidine synthetase (*GSPS*) locus 3' to the antibiotic resistance gene. Two variations of the vectors were used in this thesis: pGL1132 adds a *C*-terminal green-fluorescent protein (GFP)-tag to the insert gene (see Figure 2.4), and pGL1137 adds a *C*-terminal His_{6x}-tag (see Figure 2.5). The antibiotic resistance gene used in all studies was neomycin phosphotransferase (NEO), which confers resistance to neomycin (Curotto de Lafaille *et al.*, 1992).

In order to create *LIPA-GFP* and *LPLA-GFP* constructs, full-length *LIPA* and *LPLA* genes lacking their stop codons, were amplified from *L. major* gDNA using primer

pairs Lm19-Lm20 and Lm23-Lm24 (see Section 2.1.4.1), and cloned into plasmid pGL1132 (see Figure 2.4). The *LIPB-His* construct was created by amplifying the full-length *LIPB* gene lacking its stop codons from *L. major* gDNA using primer pair Lm21-Lm22 (see Section 2.1.4.1), and cloning into plasmid pGL1137 (see Figure 2.5). The *LPLA-His* construct was created by excision (*Nde*I/*Bgl*II digest) of the *LPLA* gene from pGL1132, and sub-cloning into *Nde*I/*Bgl*II-digested pGL1137. The *LPLA*^{H118A}*-His* construct was created by site-directed mutagenesis PCR (see Section 2.4.1.3), using the *LPLA-His* construct as a DNA template.

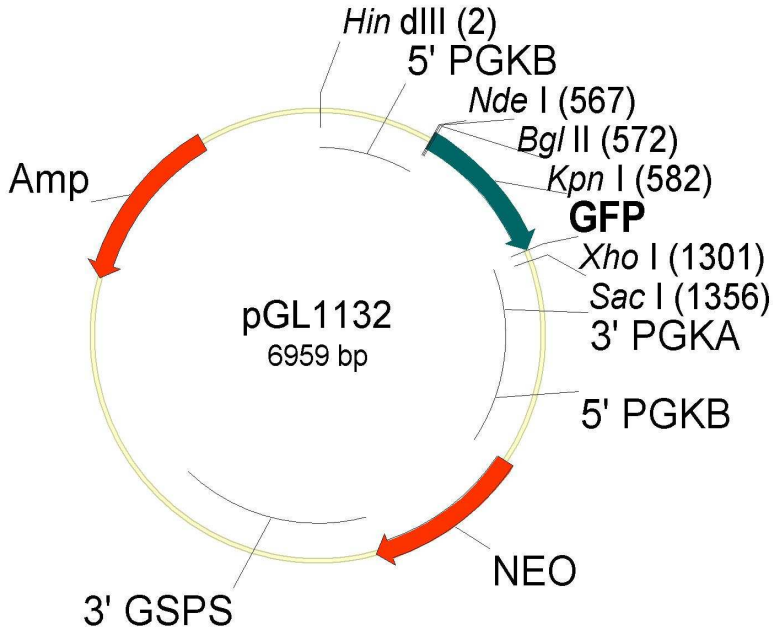


Figure 2.4 pGL1132 plasmid
The pGL1132 plasmid contains the following features: 5' PGKB, the 5'FR of the *C. fasciculata* phosphoglycerate kinase B (*PGKB*) gene; GFP, GFP gene; 3' PGKB, the 3'FR of the *C. fasciculata* phosphoglycerate kinase A (*PGKA*) gene; 3' GSPS, the 3'FR of the *C. fasciculata* glutathionyl-spermidine synthetase (*GSPS*) gene; NEO, neomycin phosphotransferase gene conferring resistance to G418; Amp, ampicillin-resistance gene. Restriction endonucleases are illustrated, and numbers in brackets mark the cut sites of the restriction endonucleases. The gene of interest to be expressed in *L. major* is directionally cloned into pGL631 using *Nde*I and *Kpn*I (or *Bgl*II) restriction sites.

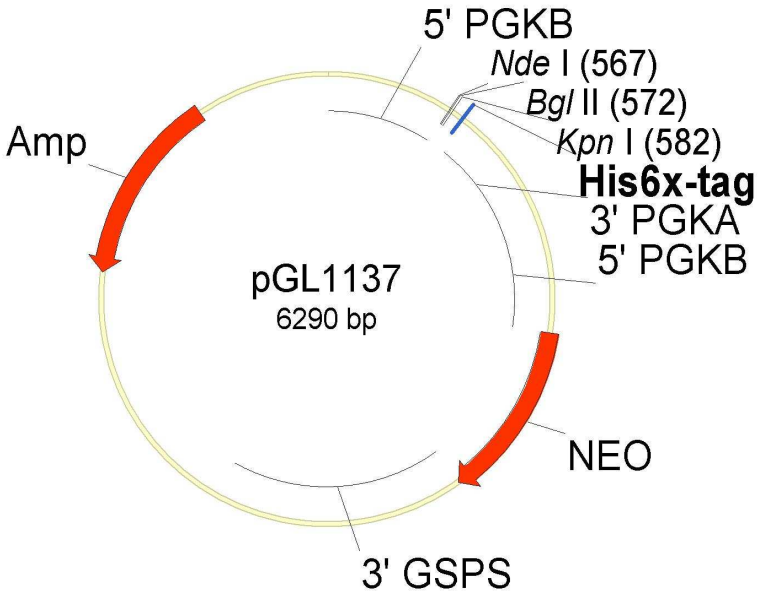


Figure 2.5 pGL1137 plasmid
The pGL1137 plasmid contains the following features: 5' PGKB, the 5'FR of the *C. fasciculata* phosphoglycerate kinase B (*PGKB*) gene; GFP, GFP gene; 3' PGKB, the 3'FR of the *C. fasciculata* phosphoglycerate kinase A (*PGKA*) gene; 3' GSPS, the 3'FR of the *C. fasciculata* glutathionyl-spermidine synthetase (*GSPS*) gene; NEO, neomycin phosphotransferase gene conferring resistance to G418; Amp, ampicillin-resistance gene. Restriction endonucleases are illustrated, and numbers in brackets mark the cut sites of the restriction endonucleases. The gene of interest to be expressed in *L. major* is directionally cloned into pGL631 using *Nde*I and *Kpn*I (or *Bgl*II) restriction sites.

2.2.7 Resistance markers

In order to select for transgenic *Leishmania* parasites, six resistance markers have been developed, four of which were used in this thesis (see Table 2.2).

Gene	Protein	Antibiotic	Reference
NEO	Neomycin phosphotransferase	G418	Cruz *et al.* (1990)
SAT	Streptothricin acetyltransferase	Neurseothricin	Joshi *et al.* (1995)
PAC	Puromycin *N*-acetyltransferase	Puromycin	Freedman & Beverley (1993)
HYG	Hygromycin B phosphotransferase	Hygromycin B	Cruz *et al.* (1991)

Table 2.2 Antibiotic resistance genes used for genetic manipulation of *Leishmania*

2.2.8 Transfection of *L. major*

For constructs designed to integrate into the *L. major* genome, 100 µg of plasmid was digested with the appropriate restriction enzymes (*Hind*III/*Bgl*II for knockout plasmids and *Pac*I/*Pme*I for pRB plasmids). The knock-out cassettes were separated by gel electrophoresis, gel-extracted and –purified, and then precipitated with ethanol according to standard procedures (Sambrook *et al.*,

1989) before resuspension in 15 μl sterile water, with a final DNA concentration of about 1 – 2 μg μl^{-1}. For ectopic constructs, 20 μg of circular plasmid was used for transfections, in a final volume of 15 μl.

For transfections, mid-log phase (at a density of 8 x 10^6 – 1 x 10^7 parasites ml^{-1}) promastigote cultures were used. Each transfection necessitated 5 x 10^7 cells and 10 – 20 μg of purified plasmid/linear DNA in 15 μl sterile water. Two different transfection methods were used. The first is using the protocol developed by Robinson and Beverley (Robinson & Beverley, 2003). Briefly, cells were counted and the appropriate number centrifuged for 10 min at 1,000 *g* at 4 ℃ and washed with electroporation buffer (EPB: 120 mM KCl, 015 mM $CaCl_2$, 10 mM K_2HPO_4, 25 mM HEPES, 2 mM EDTA, 2 mM $MgCl_2$ at pH 7.6). Cells were resuspended in 500 μl EPB to a concentration of 1 x 10^8 parasites ml^{-1}, and added to 15 μl of appropriate plasmid/linear DNA in an ice-cold 4 mm gap cuvette (Gene Pulse, Bio-Rad, Hemel Hempstead, UK) and incubated on ice for 10 min. The cuvette was exposed to two electroporation pulses at 25 μF, 1.5 kV (3.75 kV cm^{-1}), pausing 10 sec between pulses, using a Bio-Rad Gene Pulser II machine.

The second method was only used for transfection of integration cassettes. Briefly, cells were counted and the appropriate number centrifuged for 10 min at 1,000 *g* at 4 ℃ and resuspended in Nucleofactor Solution fr om the Amaxa Human T Cell Nucleofactor Kit (Amaxa AG, Cologne, Germany) to give 5 x 10^7 cells 100 μl^{-1}. The resuspended cells were mixed with 15 μl linear DNA in the cuvette from the Nuclefeofactor kit. Program U-033 was used to electroporate the cells using the Amaxa Nucleofactor Device.

In both methods, electroporated cells were incubated on ice for 10 min before transferring into a flask containing 10 ml HOMEM + 10 % FCS. Drug selection and cloning (of integrative constructs) was commenced the following day. In order to generate a population of cells for ectopic expression of a plasmid, the appropriate antibiotic(s) was added directly to the flask of cells.

2.2.9 Cloning of *L. major* by limiting dilution

In order to generate clones, three dilutions of the transfected cells were made (1/6, 1/72 and 1/864) in HOMEM + 10 % FCS including appropriate antibiotic(s), and

plated in 96-well plates. After approximately 2–3 weeks of incubation, plates containing ≤ 10 wells with parasite growth (equating to ≤ 0.1 parasite well^{-1}) were used for further analysis; the clones were sub-passaged into 3 ml HOMEM + 10 % FCS in 12-well plates and incubated until parasite density was sufficient to extract gDNA. 2 ml of the clonal cultures were used for gDNA extraction and PCR to screen for knock-out/knock-in cassette integration and/or for the presence of the endogenous gene. The remaining 1 ml of clonal culture with the desired genotype was sub-passaged into 10 ml HOMEM + 10 % FCS with appropriate drugs, and once parasite density was high enough, stabilates were made and gDNA was extracted for Southern blot analysis.

2.2.10 Fluorescence microscopy

Sub-cellular localisation of LIPA-GFP and LPLA-GFP fusion proteins expressed in *L. major* was determined by fluorescence microscopy using an AxioScop-2 mot plus microscope (Zeiss) equipped with a Hamamatsu C4742-95 CCD camera. Preparation of cells to be analysed involved adding 10 μg ml^{-1} DAPI to 10 ml of mid-log phase culture and immediately centrifuging for 1 minute at 1000 *g* at room temperature. Cells were washed twice with 20 ml PBS, and to the second wash 1 nM Mitotracker CMXRos was added. Finally, pelleted cells were resuspended in 20 μl PBS and loaded on a microscope slide with cover slip sealed with nail varnish to avoid drying-out of specimen.

Images were captured at 100x magnification using differential interference contrast (DIC) microscopy. The GFP-fusion proteins were observed using the fluorescent filter FITC ($\lambda_{excitation}$ = 494 nm and $\lambda_{emission}$ = 518 nm). The mitochondrion was selectively stained using Mitotracker CMXRos (Invitrogen), and was visualised using the fluorescent filter for rhodamine ($\lambda_{excitation}$ = 570 nm and $\lambda_{emission}$ = 590 nm). The nucleus and kinetoplast DNA of parasites were visualised using DAPI (Sigma) and the fluorescent filter for DAPI ($\lambda_{excitation}$ = 345 nm and $\lambda_{emission}$ = 458 nm).

2.2.11 Cell viability assay

The alamar blue assay was employed in order to determine the leishmanicidal/leishmanistatic effects of a compound on *Leishmania* cells.

Metabolising parasites reduce resazurin (alamar blue) and the fluorescence can be detected; the higher the level of fluorescence, the less effect the compound has at that specific concentration.

Promastigotes were cultured in the appropriate medium to a density of 5 x 10^6 ml^{-1} and then diluted to 2 x 10^6 ml^{-1} in medium. In a 96-well plate, 100 µl of neat compounds (lipoid acid, octanoic acid or 8' bromooctanoic acid) dissolved in 100 % EtOH were added to wells in the first columns; each treatment was carried out in duplicate and a negative solvent (EtOH) control included in order to normalise the data; the concentration of the EtOH solvent control was the same as the concentration of EtOH in the well containing the most concentrated of each compound, which was 0.5 % EtOH. 100 µl medium was added to each well, and then serial dilutions of drug were carried out along rows. 100 µl of cells at a density of 2 x 10^6 ml^{-1} were added to each well, thus giving a final density of 1 x 10^6 parasites ml^{-1}. Plates were incubated at 25 °C for 120 h. After the incubation period, 20 µl resazurin was added to each well (to a final concentration of 12.5 µg ml^{-1}), followed by further incubation for 48 h at 25 °C. Plates were then read using an LS 55 luminescence spectrometer with $\lambda_{emission}$ = 535 nm and $\lambda_{excitation}$ = 620 nm. All data were normalised against respective column negative controls and expressed as bar charts using the software Prism.

2.3 Bioinformatics

2.3.1 Identifying genes in the *L. major* genome

Genes were identified in the *L. major* GeneDB database (http://www.genedb.org/genedb/leish/) by performing TBLASTN searches using known protein sequences from *Escherichia coli* or *H. sapiens*. Predicted translations of the sequences identified were queried back against the SWISSPROT protein database at NCBI using BLASTP (http://www.ncbi.nlm.nih.gov/blast).

2.3.2 Multiple sequence alignments

Protein sequence analyses were performed using VECTOR NTI Suite (Invitrogen), and sequence alignments carried out using the program ClustalW (Thompson *et al.*, 1994).

2.3.3 Subcellular localisation predictions

Predictions of *N*-terminal targeting sequences were performed using TARGETP (http://www.cbs.dtu.dk/services/TargetP) (Emanuelsson *et al.*, 2000) and MITOPROT (http://mips.biochem.mpg.de/cgi-bin/proj/medgen/mitofilter) (Claros & Vincens, 1996).

Based on predicted mitochondrial targeting peptides, different *N*-terminal truncated constructs of the three genes (*LIPA, LIPB* and *LPLA*) were created for expression in *E. coli* (see Section 2.5.6).

2.4 Methods in molecular biology

2.4.1 Polymerase chain reaction

2.4.1.1AccuPrime *Pfx* SuperMix

L. major genes/flanking regions were amplified from *L. major* gDNA using AccuPrime *Pfx* SuperMix (Invitrogen), which contains 1.1 mM $MgSO_4$, 330 μM deoxyribonucleotide triphosphate (dNTPs) and 22 U ml^{-1} *Pfx* DNA polymerase. Polymerase chain reaction (PCR) was set up using 100 ng gDNA and 10 μM of each primer. The final volume per PCR reaction was 25 μl and PCR was carried out under the following conditions:

Initial denaturation	95 ℃, 5 min

30 cycles of:

Denaturation	95 ℃, 15 sec
Annealing	primer-specific temperature, 30 sec
Elongation	68 ℃, 2 min kb^{-1} amplified

Final elongation	68 ℃, 10 min

The PCR products were analysed on a 1 % agarose gel containg Sybr Safe at a 1/10,000 dilution, and fragments of the expected size were cloned into pCR-BluntII-TOPO using the Zero Blunt TOPO PCR cloning kit (Invitrogen).

2.4.1.2ReddyMix PCR Master Mix

In order to screen for bacterial clones possessing the desired insert in plasmid, and for *L. major* clones having correct knockout cassette integration (and for the presence/absence of the endogenous gene), PCR was carried out using ReddyMix PCR Master Mix (Thermo Scientific). Individual PCR reactions contain 0.625 U of ThermoPrime *Taq* DNA polymerase, 1.5 mM $MgCl_2$ and 0.2 mM of each of the four dNTPs. The PCR was set up using bacterial colony/up to 100 ng gDNA and 10 μM of each primer. The final volume per PCR reaction was 10 μl and PCR was carried out under the following conditions:

Initial denaturation	95 ℃, 5 min

30 cycles of:

Denaturation	95 ℃, 15 sec
Annealing	primer-specific temperature, 30 sec
Elongation	72 ℃, 2 min kb $^{-1}$ amplified

Final elongation	72 ℃, 10 min

The PCR products were analysed on a 1 % agarose gel.

2.4.1.3Site-directed mutagenesis PCR

Site-directed mutagenesis was used to introduce three point mutations into *LPLA* in order to produce the mutant *LPLA*^{H118A}. The *LPLA-His* construct (see Section 2.2.6) was used as a DNA template for mutagenesis PCR using primer pair Lm64-Lm65 (see Section 2.1.4.1). PCR was carried out using the QuickChange Lightning Site-Directed Mutagenesis Kit (Stratagene). The reaction contained 50 ng of plasmid DNA, 125 ng of each primer, 1x reaction buffer, 1.5 μl of QuickSolution reagent, 1 μl of dNTP mix and 1 μl of 'QuickChange Lightning' DNA

polymerase (provided with kit). The PCR was carried out under the following conditions:

Initial denaturation	95 ℃, 2 min
18 cycles of:	
Denaturation	95 ℃, 20 sec
Annealing	69 ℃, 10 sec
Elongation	68 ℃, 4 min (for 7.8 kb plasmid)
Final elongation	68 ℃, 5 min

2 µl of the provided *Dpn*I enzyme (Stratagene) was added to the finished PCR reaction, and incubated at 37 ℃ for 5 min in order to degrade methylated (plasmid) DNA. The PCR products were then transformed into XL-10 Ultracompetent bacteria (Stratagene). Bacterial colonies containing the plasmid with insert were verified by restriction digestion and then mutagenesis of the *LPLA* insert was confirmed by sequencing (Dundee Sequencing Centre). The *LPLA*^{H118A}-*His* construct (see Section 2.2.6) was then transfected into WT parasites to create the WT[*LPLA*^{H118A}-*His*] line.

2.4.2 Cloning techniques

2.4.2.1TOPO cloning of PCR products

In order to facilitate the cloning of PCR products into their destination vectors, inserts were firstly ligated into an intermediate vector. Given that PCR products amplified with *Pfx* polymerase have blunt ends, the Zero Blunt TOPO PCR cloning kit containing the vector pCR-BluntII-TOPO was chosen for intermediate vector cloning (see Figure 2.6). Ligations were carried out according to instructions provided by the manufacturer (Invitrogen) and 1 µl of the ligation transformed into TOP10 cells (see Section 2.1.2).

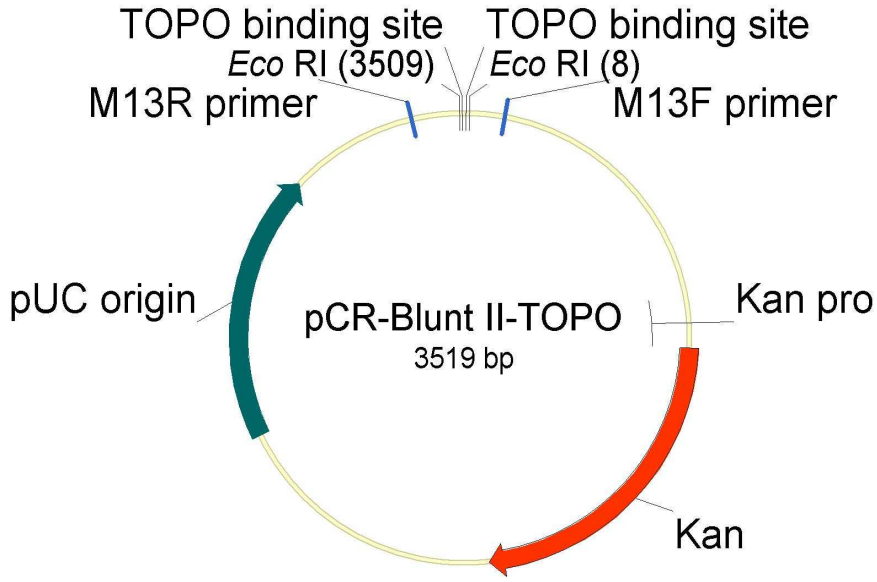


Figure 2.6 Vector map of pCR-BluntII-TOPO
This figure displays the important features of the pCR-Blunt II-TOPO plasmid, which was used to clone PCR products amplified with polymerases with proofreading abilities. The plasmid contains the kanamycin resistance cassette (KanR; 795 bp) for selection in *E. coli*. The topoisomerase I is covalently bound to the TOPO binding sites. After cloning, the generated construct was analysed by diagnostic digests using the *Eco*RI restriction sites shown, and the sequence of the cloned PCR product was verified by sequencing using the primers M13F and M13R.

2.4.2.2Sub-cloning into destination plasmids

Bacterial clones containing TOPO plasmid with the correct insert (as determined by colony PCR (see Section 2.4.1.2) or by restriction endonuclease digest (see Section 2.4.6) were used in subsequent cloning steps, after verification of the correct DNA sequence (see Section 2.4.9). The insert was isolated from the TOPO plasmid using the appropriate restriction endonucleases (see Section 2.4.6). The destination plasmid was linearised using the same restriction enzymes and the DNA fragments were separated by agarose gel electrophoresis (see Section 2.4.10). The digested insert and destination plasmid were excised from the gel and purified using the Qiagen Gel Extraction kit as per manufacturer's instructions. The insert was then ligated into the destination plasmid using the Rapid DNA Ligation kit, according to instructions provided by the manufacturer (Roche). A molar vector: insert ratio of 1: 3 was used in the ligation reactions. 10 µl of the ligation reactions were transformed into competent TOP10 *E. coli* and plated on LB-plates containing the appropriate antibiotics. Individual colonies were analysed by DNA plasmid miniprep (see Section 2.4.5) and restriction endonuclease digestion (see Section 2.4.6). Plasmids containing the correct insert size were analysed by DNA sequencing (see Section 2.4.9).

2.4.3 Preparation of chemically-competent bacteria

Competent bacteria used during this work were either bought from Invitrogen (TOP10), Novagen (BLR (DE3)) or Stratagene (XL10-Gold Ultracompetent) (see Section 2.1.2), or were made chemically competent. Briefly, a 5 ml overnight culture of the appropriate *E. coli* strain was set up and then used to inoculate 200 ml LB-medium with the appropriate (if necessary at all) antibiotics. The culture was grown at 37 ℃ at 220 rpm until OD_{600} reached 0.6. The culture was transferred into a sterile centrifuge tube and incubated for 15 min at 4 ℃. The cells were then pelleted for 10 min at 6,000 *g* at 4 ℃. The pellet was resuspended in 30 ml of ice-cold TfbI buffer (100 mM RbCl, 50 mM $MnCl_2$-$4H_2O$, 30 mM KNa, 10 mM $CaCl_2$-$2H_2O$ and 15 % glycerol) and incubated at 4 ℃ for 30 min. The cells were then centrifuged for 5 min at 6,000 *g* at 4 ℃ and resuspended in 6 ml of ice-cold TfbII buffer (10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2$-$2H_2O$ and 15 % glycerol). Aliquots of competent cells were snap frozen in dry ice and immediately transferred to – 80 ℃ for storage.

2.4.4 Transformation of competent bacteria

Chemically-competent *E. coli* were used to transform plasmid DNA and ligation reactions. The bacteria were removed from storage at – 80 ℃ and thawed on ice for 10 min. The DNA (1 µl of 1/10 diluted miniprep plasmid DNA or 10 µl of ligation) was added to the bacteria and then incubated on ice for 30 min. Heat shock was carried out for 45 sec at 42 ℃ and then the bacteria were kept on ice for 10 min. 250 µl LB was added to the bacteria, which were then incubated for 30 min to 1 h shaking at 37 ℃ at 250 rpm. Finally, the cells were pelleted and resuspended in 50 µl LB and plated onto agar plates supplemented with the appropriate antibiotic (see Section 2.1.1).

2.4.5 Plasmid DNA isolation from bacteria

Plasmid DNA was isolated from *E. coli* by one of two methods, depending upon the amount of DNA required. In order to purify small amounts of DNA to screen for correct insert ligation into a plasmid, 5 ml LB with the appropriate antibiotics was inoculated with a single bacterial colony overnight at 37 ℃ at 250 rpm. The bacteria were pelleted and DNA purification carried out using the Qiaprep Spin

Miniprep kit according to manufacturer's instructions (Qiagen). Large quantities of DNA to be used in transfection of *Leishmania* were isolated using a similar silica column-based method, but using a higher volume of culture (250 ml) and the Hi-Speed Plasmid Maxi kit (Qiagen). DNA was precipitated into a volume of 500 µl using the QiaPrecipitator module (Qiagen), and the concentration determined by spectrophotometric means (see Section 2.4.8).

DNA obtained from miniprep and maxiprep protocols was subjected to restriction endonuclease digestion (see Section 2.4.6) in order to determine/verify that the plasmid contained the correct insert in the correct orientation.

2.4.6 Restriction endonuclease digestion

Restriction endonuclease digestion was routinely carried out on miniprep DNA in order to verify the correct cloning of an insert into an intermediate/destination vector. Unless stated otherwise, 3 µl of miniprep DNA was digested with 1 U of restriction enzyme in the appropriate buffer (NEB). Reactions were typically incubated for 2 h at 37 ℃ (or 50 ℃ for *Bsa*I). Restriction patterns were visualised by mixing with 6xDNA loading dye and subjection to agarose (1 % w/v) electrophoresis (see Section 2.4.10)

2.4.7 Ethanol precipitation of gDNA

DNA was precipitated using 0.1 volumes of 3 M Na-acetate pH 5.2 and three volumes of ice-cold ethanol. Additionally, 1 µl of glycogen (20 mg ml^{-1} stock) was added in order to visualise the DNA pellet throughout centrifugation steps. The mixture was incubated for 20 min at – 80 ℃ and then the precipitated DNA pelleted at 13,000 rpm (Fisher Scientific accuSpin MicoR with 24-plate motor) at 4 ℃ for 30 min. The DNA was washed with ice-cold 70 % (v/v) ethanol, and then allowed to air-dry (under a sterile hood if the DNA was to be used for transfections in *Leishmania*). Finally, the DNA was re-suspended in water or an appropriate buffer.

2.4.8 Determining DNA concentration

DNA concentration was measured using a UV-spectrophotometer (Shimadzu) at $OD_{260\ nm}$. $OD_{260\ nm}$ = 1 corresponded to 50 µg ml^{-1} for double-stranded DNA. DNA

purity was determined by calculating the ratio of DNA ($OD_{260\ nm}$) to protein ($OD_{280\ nm}$).

2.4.9 DNA sequencing

DNA sequencing was carried out externally by the Dundee Sequencing Centre (www.dnaseq.co.uk) using Applied Biosystems Big-Dye Ver 3.1 chemistry on an Applied Biosystems model 3730 automated capillary DNA sequencer. It was necessary to provide 200 – 300 ng of DNA template, along with 3.2 pmoles of sequencing primer, per sequencing reaction.

2.4.10 Agarose gel electrophoresis

DNA quantity and quality was routinely examined by agarose gel electrophoresis using the Sub-Cell GT system (Bio-Rad). 1 % (w/v) agarose was dissolved in 1x TAE buffer by boiling the mixture in a microwave. SYBR safe (Invitrogen) DNA stain was added at 1/10,000 and the agarose poured into gel trays with the appropriate size comb. The DNA samples were mixed with 6x DNA loading dye and electrophoresis was performed at 100 V in 1x TAE buffer until DNA fragments were sufficiently separated. 1 kb DNA ladder (NEB) was run alongside DNA samples to permit size determination of the bands visualised by UV illumination at 302 nm or 365 nm using the Gel Doc XR system (Bio-Rad).

2.4.11 Southern blot analysis

In order to analyse the genotypes of genetically-manipulated parasite lines, firstly their gDNA was isolated (see Section 2.2.3.1), and then subjected to diagnostic digests with *Nru*I overnight at 37 ℃. Fully-digested gDNA was separated on a 0.8 % agarose gel overnight at 20 V. The DNA was transferred onto positively-charged nylon membrane Hybond-N+ (GE Healthcare). For the transfer the membrane was placed on the VacuGene XL apparatus (GE Healthcare) and was covered with 0.25 N HCl depurination- (0.25 N HCl), denaturation- (1.5 M NaCl and 0.5 M NaOH) and transfer (20x SSC) solutions were incubated on the gel for 30 min, 30 min and 1 h, respectively. DNA was subsequently cross-linked to the membrane. Blocking, probe-making, hybridisation, washing and detection steps were all carried out using the AlkPhos Direct Kit according to instructions provided by the manufacturer (GE Healthcare). Pre-hybridisation-, hybridisation- and

primary wash steps were carried out at 61 ℃ and the probes used recognised *L. major LIPA* and *LPLA* genes and *HYG* and *SAT* genes. *LIPA*- and *LPLA* gene probes were made using full-length *LIPA* and *LPLA* genes amplified from *L. major* gDNA using primer pairs Lm1-Lm2 and Lm13-Lm14, respectively (see Section 2.1.4.1). *HYG* and *SAT* gene probes were derived from digesting (*Bam*HI/*Spe*I) *HYG* and *SAT* genes from plasmids pGL435 and pGL158, respectively (see Figure 2.1 and Figure 2.2). The gene fragments were gel extracted and purified (see Section 2.4.2.2) and then used to make probes.

2.5 Biochemical methods

2.5.1 SDS-PAGE

Proteins were separated by SDS-PAGE (Laemmli, 1970). The NuPAGE electrophoresis system (Invitrogen) was used with either 4 – 12 % Novex Bis-Tris pre-cast gels or with non-gradient polyacrylamide gels. Gradient gels were utilised for the analyses of parasite extract whereas non-gradient polyacrylamide gels were used for examining purity of recombinant protein. In non-gradient gels, the running gel consisted of 6 ml running gel buffer (375 mM Tris pH 8.9, 0.1 % (w/v) SDS, 10 – 15 % (v/v) acrylamide) which was polymerised by addition of 5 ml N,N,N',N'-Tetramethylethylenediamine (TEMED) and 25 µl ammonium persulphate (APS) (10 mg ml^{-1} stock). The gel was poured into empty plastic cassettes (Invitrogen) and allowed to set. The stacking gel, consisting of 2 ml stacking gel buffer (122 mM Tris pH 6.7, 0.1 % (w/v) SDS and 5 % (v/v) acrylamide) was poured on top of the running gel and an appropriate comb was placed in the stacking gel. Protein samples were prepared while the stacking gel was in the process of setting. 10 µg parasite extract or 1 µg recombinant protein was mixed with 6x loading buffer and denatured at 100 ℃ for 5 min. The samples were loaded and the non-gradient polyacrylamide gels were run in 1x running buffer (25 mM Tris, 192 mM glycine, 0.1 % (w/v) SDS) and pre-cast gels in 1x MOPS buffer (Invitrogen) at 40 mA (200 V maximum). Following electrophoresis, gels were either stained with Coomassie blue (see Section 2.5.2) or were analysed by western blotting (see Section 2.5.3).

2.5.2 Coomassie blue staining

Coomassie blue staining was carried out in order to determine the purity of elution fractions from protein purification (see Section 2.5.6). After SDS-PAGE (see Section 2.5.1), the gel was removed from the cassette and covered with Coomassie blue stain (40 % (v/v) methanol, 10 % (v/v) acetic acid, 0.1 % (w/v) and coomassie brilliant blue R-250) for at least 1 h at room temperature. For destaining, the Coomassie blue solution was removed and the gel was incubated in destain solution (20 % (v/v) methanol and 10% (v/v) acetic acid) for at least 1 h, changing the destain solution two times in the process.

2.5.3 Western blotting

Proteins were transferred from gel to membrane using the Trans-Blot SD Semi-Dry Electrophoretic Transfer Cell (Bio-Rad) for 1 h at 20 V at room temperature. Ponceau-S (reversible) staining was used to visualise equal loading in all lanes, before continuing the experiment. Blots were blocked for 1 h in 5 % (w/v) milk in PBS (5 % MPBS) at room temperature, followed by incubation with primary, and secondary antibodies, in 2 % (w/v) MPBS with 0.1 % (v/v) Tween-20 (2 % MPBST). In between primary- and secondary antibody incubations, and after the secondary antibody step, blots were washed 3 x 10 min in PBS containing 0.1 % (v/v) Tween-20 (PBST). Detection of secondary antibody (linked to horseradish peroxidase)-bound protein was carried out using Immobilon Western Blot Detection Kit (Millipore) for 5 min at room temperature. The resulting chemiluminescence was detected on autoradiography film (Kodak).

Quantification of western blot band density was carried out by using a ChemiDoc XRS machine (Bio-Rad) to detect chemiluminescence signal generated by western blotting. Subsequently, Quantity One software (Bio-Rad) was used to determine band density. A standard curve was produced with band intensity as a function of the quantity of recombinant protein, in order to determine the quantity of a specific protein in promastigote protein lysate.

2.5.4 Determining protein concentration

The protein concentration of parasite extract and recombinant protein was determined by the Bradford assay (Bradford, 1976), using the Bio-Rad protein

assay reagent. The absorbance of a protein solution mixed with Bradford reagent (Bio-Rad) was measured at 595 nm and the protein concentration was determined relative to a standard curve of known BSA concentrations.

2.5.5 Estimation of protein molecular mass

After SDS-PAGE and Coomassie staining or western blotting, a standard curve was made by calculating the relative mobility (Rf) of standard proteins and then logging the values. The values were used to construct a standard curve, which was used to calculate the molecular mass of bands of an unknown size.

2.5.6 Expression and purification of proteins with a Strep-tag

2.5.6.1Cloning of expression constructs

In order to express *L. major* LIPA, LIPB and LPLA in *E. coli*, different truncated versions of the genes were cloned into the vector pASK-IBA3 (see Figure 2.7 and Section 2.3.3), which adds a *C*-terminal Strep-tag to the translated transgene. The Strep-tag is a stretch of 8 amino acids that specifically binds to a modified version of streptavidin called Strep-Tactin, and as such a recombinant protein bearing a Strep-tag can be purified by affinity chromatography using Strep-Tactin resin (IBA).

The genes of interest were cloned into pASK-IBA3 by directional *Bsa*I cloning (see Section 2.1.4.2). Table 2.3 highlights the primers used in order to amplify the different versions of *LIPA, LIPB* and *LPLA* by PCR, and the size of the *N*-terminal truncation. *N*-terminal truncations were based upon targeting predictions made by MitoProt and/or TargetP software as well as by alignments with *E. coli* homologues (see Sections 2.3.3 and 3.4). PCR fragments were cloned into the intermediate vector pCR-BluntII-TOPO (see Section 2.4.2.1), and correct sequences were verified by DNA sequencing (see Section 2.4.9). DNA inserts were then sub-cloned into pASK-IBA3 by directional *Bsa*I cloning; correct ligation of insert into vector was confirmed via *Xba*I/*Hind*III double digest and restriction digest analysis by agarose electrophoresis (see Sections 2.4.6 and 2.4.10). The nomenclature of the final constructs will be referred to as the name of the DNA fragment (see Table 2.3) followed by IBA3 (representing vector pASK-IBA3).

Name of DNA fragment	Primer pairs used in PCR	Size of PCR fragment (kbp)
$LIPA_{FL}$	Lm25-Lm29	1.250
$LIPA_{T1}$	Lm26-Lm29	1.235
$LIPA_{T2}$	Lm27-Lm29	1.193
$LIPA_{T3}$	Lm28-Lm29	1.067
$LIPB_{FL}$	Lm30-Lm33	0.791
$LIPB_{T1}$	Lm31-Lm33	0.704
$LIPB_{T2}$	Lm32-Lm33	0.689
$LPLA_{FL}$	Lm34-Lm37	1.535
$LPLA_{T1}$	Lm35-Lm37	1.499
$LPLA_{T2}$	Lm36-Lm37	1.430

Table 2.3 Cloning genes into pASK-IBA3 for protein expression
Table providing information about the cloning of full-length (FL) and N-terminally truncated versions (T1-T3) of *LIPA, LIPB* and *LPLA* to be used in protein expression trials. All primers in this table have *Bsa*I restriction sites (see Section 2.1.4.2).

2.5.6.2Trial protein expression

In order to determine the most yielding conditions for expression of recombinant proteins, trial expressions were carried out. Variables included: full-length or *N*-terminally truncated expression construct (see Section 2.5.6.1); temperature of incubation after induction of protein expression (15 ℃, 30 ℃ or 37 ℃); length of incubation after induction of protein expression (1 h, 2 h, 3 h or overnight). Briefly, 5 ml of LB supplemented with appropriate antibiotic(s) was inoculated with a single colony from a transformation with a protein expression construct and incubated overnight at 37 ℃ at 250 rpm. This starter culture was diluted 1/100 into 50 ml LB with appropriate antibiotics and incubated at 37 ℃ at 250 rpm until OD_{600} nm = 0.6. Protein expression from the expression construct was then induced (with anhydrotetracycline (ATc) for pASK-IBA3 constructs (see Section 2.5.6.3). Post-induction, cultures incubated at different temperatures and for different lengths of time were centrifuged at for 1 min at 13,000 rpm (Eppendorf microcentrifuge with F45-24-11 rotor) at room temperature, the supernatant was discarded and the pellet was resuspended in 100 µl BugBuster Protein Extraction Reagent (Novagen). 0.5 µl benzonase (Novagen) was also included in order to degrade DNA. The mixture was incubated at 4 ℃ with constant shaking. After 15 min, the mixture was centrifuged for 15 min at 13,000 rpm (Fisher Scientific accuSpin MicroR with 24-place rotor) at 4 ℃. The supernatant was transferred into a fresh

tube, and the pellet was resuspended in 100 µl PBS. For analyses, 5 µl of the pellet fractions and 10 µl of the supernatant fractions were separated by SDS-PAGE (see Section 2.5.1) and evaluated by western blotting using specific antibodies against the protein-tag (see Table 2.1).

2.5.6.3Protein purification

Trial expressions indicated that of all the constructs tested, *LIPA$_{T3}$-IBA3*, *LIPB$_{FL}$-IBA3* and *LPLA$_{T2}$-IBA3* resulted in the production of the highest level of the respective proteins, LIPA$_{T3}$-Strep, LIPB$_{FL}$-Strep and LPLA$_{T2}$-Strep. The expression of these constructs in BLR (DE3) cells was carried out as described in Section 2.5.6.2. The plasmid contains the ampicillin-resistance gene (AMP) to allow selection of the transfected plasmid with ampicillin. At optical density $(OD)_{600\ nm}$ = 0.6, the expression of recombinant protein was induced by addition of 200 ng ml^{-1} ATc. The protein was expressed at optimal conditions, which were determined previously using the BugBuster protein extraction kit (see Section 2.5.6.2), and which are outlined for LPLA, LIPB and LIPA in Section 3.5. After the optimal time of expression, the culture was centrifuged for 15 min at 6,000 *g* at 4 ℃. The bacterial pellet was resuspended in a small volume of chilled buffer W (100 mM Tris-HCl (pH 8.0), 150 mM NaCl). The resuspended pellet was stored at – 20 ℃ with protease inhibitors at the same concentration as used in *L. major* lysis buffer (see Section 2.1.1).

For purification, the pellet was thawed and 50 mg ml^{-1} lysozyme was added. The pellet was incubated for 30 min on ice. Additional protease inhibitors were added (at the same concentrations as above). Bacterial cells were lysed using the One Shot Cell Disrupter (Constant Systems) at 15 kPsi. Disrupted bacteria were centrifuged for 30 min at 20,000 rpm (Beckman J2-H5 centrifuge with a JA- 20 rotor) at 4 ℃. The cleared lysate was subsequently filtered through a 0.4 µm disposable filter unit (Satorius) in order to remove any remaining particulate matter. Affinity chromatography with Strep-Tactin Sepharose (IBA) was used to purify Strep-tagged recombinant protein, according to instructions provided by the manufacturer. Briefly, 2 ml Strep-Tactin slurry was transferred to an Econo-Pac column (Bio-Rad) to give a column volume (CV) of 1 ml Strep-Tactin Sepahorse. The sepharose was equilibrated with two CV of buffer W. The cleared lysate was then added to the column and allowed to pass through by gravity flow. The column

was then washed with five CV of buffer W. Cleared lysate, flow-through and wash fractions were retained for analysis by SDS-PAGE and Coomassie blue staining. Elution of recombinant protein was carried out with buffer E (buffer W with 2.5 mM desthiobiotin), and six fractions were collected. The protein concentration of each elution fraction was determined by Bradford assay (see Section 2.5.4). Finally, SDS-PAGE (see Section 2.5.1) of all samples taken during the purification and elution fractions was carried out and analysed by Coomassie blue staining (see Section 2.5.2) or western blotting with α-Strep-tag antibody (see Section 2.5.3).

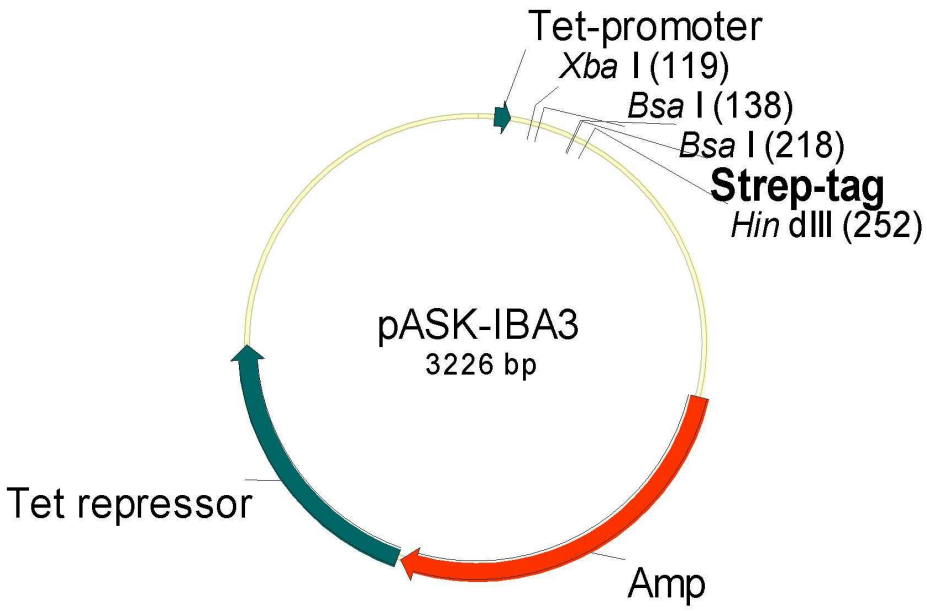


Figure 2.7 pASK-IBA3 plasmid
The pASk-IBA3 plasmid contains the following features: Tet promoter, anhydrotetracycline-inducible promoter; Amp, ampicillin-resistance gene; Tet repressor, repressor of Tet promoter. Restriction endonucleases are illustrated, and numbers in brackets mark the cut sites of the restriction endonucleases. The gene of interest to be expressed in *E. coli* is directionally cloned into pASK-IBA3 using *Bsa*I restriction sites.

2.5.7 Antibody production and purification

Rabbit polyclonal antibodies were raised against recombinant $LIPA_{T3}$-Strep, $LIPB_{FL}$-Strep and $LPLA_{T2}$-Strep proteins, by Eurogentec (Belgium), following their standard immunization protocols. 400 μg of each antigen was subjected to SDS-PAGE and Coomassie blue staining, the band excised and injected into rabbit; two rabbits were inoculated per antigen.

Antibodies specific to the recombinant proteins were purified from antisera by affinity chromatography purification using AminoLink Coupling resin, according to instructions provided by the manufacturer (Thermo Scientific). Briefly, 1 mg of

recombinant protein used to inoculate rabbits for antibody production was dialysed into 1 ml PBS using a Centricon Plus-20 centrifugal filter device (Millipore), in order to dilute out any amines from the elution fraction. The protein was then covalently linked to 2 ml AminoLink Coupling slurry (1 ml CV of resin) by treatment with cyanoborohydride. Any remaining binding sites were blocked by treatment with a quenching buffer (1 M Tris-HCl, pH 7.4). The covalently bound protein was then washed with wash buffer (1 M NaCl) and equilibrated with 5 CV of Protein A IgG binding buffer. 5 ml of antiserum to be purified was added to 2.5 ml Protein A IgG binding buffer before transferring the mixture to the prepared column. The column was washed with 20 ml Protein A IgG binding buffer. Ten 500 µl elutions were carried out with Protein A IgG elution buffer into tubes containing 25 µl of neutralisation buffer (1 M Tris-HCl, pH 9.0).

2.6 DNA content analysis

Fluorescence activated cell sorting (FACS) analysis was used in order to determine the DNA content of transgenic parasites. Mid-log promastigotes were centrifuged for 5 min at 1,000 *g* at 4 ℃, washed with PBS and then resuspended in 1 ml of 70 % (v/v) methanol in PBS. Fixation in methanol was carried out at 4 ℃ for at least 1 h to overnight. Prior to analysis, the cells were washed once in PBS and then resuspended in 1 ml of PBS containing 10 µg ml^{-1} propidium iodide and 10 µg ml^{-1} RNAse A, and DNA labeling was carried out at 37 ℃ for 1 h in a light-protected box.

FACS analysis was carried out with a Becton Dickinson FACSCalibur, using the FL2-A (fluorescence intensity at 585/642 nm, note $\lambda_{emission}$ propidium iodide = 620 nm under an $\lambda_{excitation}$ = 488 nm (blue, Uniphase Argon Ion Laser)), the Forward Scatter (FSC, relative cell size) and the Side Scatter detectors (SSC, cell granulometry or internal complexity). For each sample, at least 10,000 events (cells) were counted. Data was interpreted with CellQuestPro software (BD Bioscience).

2.7 Statistical analyses

All graphs and statistical analyses were carried out using the program Prism 3.0 (GraphPad), unless otherwise stated. All experiments with statistical data attached

were carried out in duplicate, and one-way ANOVA with Tukey post-test used to determine the significance of results; *p*-values < 0.05 were considered to be significant.

2.8 Functionality assays

As discussed in Section 2.5.6, full-length- and different truncated versions of *LIPA, LIPB* and *LPLA* were cloned into pASK-IBA3, in order to test for the optimal expression construct for recombinant protein purification. In order to determine whether the different constructs encoded functional proteins, *LIPA-IBA3* constructs were transformed into the *lipA* deficient *E. coli* strain KER176 (Vanden Boom *et al.*, 1991), and *LIPB-IBA3* and *LPLA-IBA3* constructs were transformed into the *lipB* deficient *E. coli* strain KER184 (Vanden Boom *et al.*, 1991). Functionality was assessed by whether the constructs would complement the growth defect of the bacterial lines when grown on minimal medium (see Section 2.1.1). Both cell lines were grown on M9 minimal plates containing 100 mg ml^{-1} ampicillin (to select for the construct), 50 mg ml^{-1} kanamycin (to select for the transgenic cell line) and 200 ng ml^{-1} ATc (to induce recombinant protein expression) and without- or with supplementation of 10 μM LA.

3 *In silico* and functionality studies

3.1 Introduction

In all organisms, α-KADHs and the GCC are intricately linked to energy metabolism (Perham, 2000) and (Douce *et al.*, 2001). LA is an essential component of these complexes, covalently linked to the acyl transferase subunits of α-KADHs and to the H-protein of the GCC. The post-translational event that transfers LA to apoproteins is referred to as lipoylation. Two pathways exist to carry out lipoylation; LA biosynthesis and LA salvage. Biosynthesis enzymes LIPA and LIPB have been studied to varying degrees in *E. coli* (Jordan & Cronan, 1997a; Reed & Cronan, 1993), *M. tuberculosis* (Sassetti *et al.*, 2003), *P. falciparum* (Wrenger & Muller, 2004), *T. gondii* (Thomsen-Zieger *et al.*, 2003), *A. thaliana* (Yasuno & Wada, 2002) and *H. sapiens* (Morikawa *et al.*, 2001). Salvage of LA in *E. coli* requires just one enzyme, LPLA (Morris *et al.*, 1994; Reed *et al.*, 1994), which catalyses the formation of an activated intermediate, lipoyl-AMP, and the subsequent transfer of the lipoyl moiety to the apoprotein. The mammalian system necessitates the sequential action of two enzymes for activation and transfer of LA to the apoprotein; ACSM1 (Fujiwara *et al.*, 2001) and LT (Fujiwara *et al.*, 1997a; Fujiwara *et al.*, 1999), respectively.

In this chapter, an *in silico* approach was employed in order to identify and analyse potential components comprising α-KADHs, the GCC and the enzymes involved in LA acquisition and ligation in *L. major*. The next aim was to determine whether lipoylation of the acyl transferase and H-protein subunits occurs in *L. major* promastigotes. Subsequently, the genes encoding *L. major* LIPA, LIPB and LPLA were heterologously expressed in *E. coli* deficient in either *lipA* or *lipB*, in order to ascertain whether *L. major* encodes active lipoylating machinery. Finally, the effects of LA analogues on promastigote growth and lipoylation were studied.

3.2 Sequence analyses of lipoylated protein complexes

In order to identify potential homologues of α-KADH subunits and subunits of the GCC, the *L. major* genome database was searched by TBLASTN using *H. sapiens* and *E. coli* amino acid sequences as queries (as described in Section 2.3). Potential homologues of α-KADH subunits were identified (see Table 3.1):

1. Four genes encoding the α-ketoglutarate dehydrogenase (α-KGDH) complex; two α-ketoglutarate dehydrogenase isoenzymes (designated E1k-A and E1k-B) and one succinyl transferase (E2k).

2. Four genes encoding the pyruvate dehydrogenase (PDH) complex; two pyruvate dehydrogenase heteromers (E1p-α and E1p-β) and one acetyl transferase (E2p).

3. Four genes encoding the branched-chain α-ketoacid dehydrogenase (BCKDH) complex; two branched-chain dehydrogenase heteromers (E1b-α and E1b-β) and one branched-chain transacylase (E2b).

4. One gene encoding the lipoamide dehydrogenase subunit (LipDH), which is presumably common to all α-KADHs.

In addition, potential homologues of components of the GCC were identified (see Table 3.1):

1. One lipoyl-domain protein (H-protein).

2. One glycine dehydrogenase (P-protein).

3. Two aminomethyl transferase isoenzymes (T-protein-A and T-protein-B).

4. One gene encoding the LipDH subunit, which is presumably the same as that used by α-KADHs.

	Predicted Sizes (kDa)			Systematic Name	Targeting Predictions (%)		Identity of *L. major* to other species (%)	
Subunit	***E. coli***	***H. sapiens***	***L. major***	***L. major***	**Mitoprot**	**Target P**	***H. sapiens***	***E. coli***
E1p	99.7 (AAC73225)	N/A	N/A	N/A	N/A	N/A	N/A	N/A
E1p-α	N/A	43.3 (NP_000275)	42.9	LmjF18.1380	79 MT	47 MT	43	N/A
E1p-β	N/A	39.2 (NP_000916)	37.9	LmjF25.1710	99 MT	87 MT	57	N/A
E1k	105.1 (ACB01934)	115.9 (NP_002532)	N/A	N/A	N/A	N/A	N/A	N/A
E1k-like	N/A	114.5 (NP_060715)	N/A	N/A	N/A	N/A	N/A	N/A
E1k-A	N/A	N/A	114.4	LmjF36.3470	97 MT	92 MT	35	33
E1k-B	N/A	N/A	112.7	LmjF27.0880	95 MT	78 MT	34 (36 to LmE1k -A)	32
E1b-α	N/A	50.5 (NP_000700)	53.3	LmjF21.1430	80 MT	85 MT	36	N/A
E1b-β	N/A	43.1 (NP_000047)	40	LmjF35.0050	84 MT	73 MT	49	N/A
P-protein	104.4 (AAC75941)	112.7 (NP_000161)	108.5	LmjF26.0030	95 MT	86 MT	47	46
E2p	66.1 (AAC73226)	68.9 (NP_001922)	48.7	LmjF36.2660	86 MT	52 MT	24	18
						62 SP		
E2p-like	N/A	54.1 (NP_001128496)	46.1	LmjF21.0550	61 MT	5 MT	15 to *Hs* pE2	N/A
						85 SP	16 to *Hs* E3BP**	
E2k	44 (AAA23898)	48.6 (NP_001924)	41.7	LmjF28.2420	96 MT	78 MT	38	39
E2b	N/A	53.5 (NP_001909)	50.2	LmjF05.0180	100 MT	92 MT	41	N/A
H-protein	13.8 (AAC75942)	18.9 (NP_004474)	15.2	LmjF35.4720	76 MT	50 MT	36	37
LipDH	50.7 (AAC73227)	54.2 (NP_000099)	50.6	LmjF32.3310	9 MT (ATG start)	21 MT (ATG start)	47	38
						16 SP, 57 OT (ATG start)		
					23 MT (GTG start)	83 MT (GTG start)*		
T-protein	40.1 (AAC75943)	43.9 (NP_000472)	N/A	N/A	N/A	N/A	N/A	N/A
T-protein A	N/A	N/A	43.2	LmjF36.3800	17 MT	12 MT	29	38
						25 SP, 38 OT		
T-protein B	N/A	N/A	41.3	LmjF36.3810	17 MT	12 MT	29	39
						25 SP, 38 OT		

Table 3.1 Sequence analyses of subunits of *L. major* α-KADHs and of the GCC
L. major homologues of subunits of α-KADHs and of the GCC were identified as described in Section 2.3. Red subunits are those which catalyse the oxidative decarboxylation of substrates. Blue subunits are those that carry out transferase reactions. The green subunit is common to both α-KADHs and the GCC and catalyses the re-oxidation of dihydrolipoic acid (DHLA) to oxidised LA. In the 'Protein Size' column, accession numbers of *E. coli* and *H. sapiens* proteins are shown in brackets. Mitochondrial targeting predictions were carried out by analysing *L. major* protein sequences with MitoProt or TargetP. The results obtained indicate the percent likelihood of mitochondrial targeting (MT). Some proteins were predicted to possess signal peptides (SP), or an alternative processing type (OT). The mitochondrial targeting for the LipDH subunit was unexpectedly low. However, by starting the ORF at a GTG 75 bp upstream of the predicted ATG start, the likelihood of mitochondrial targeting was significantly increased (*). Amino acid sequence identity of *L. major* α-KADH- and GCC subunits to *E. coli* and *H. sapiens* homologues was determined using Vector NTI (see Section 2.3). In the case of the *L. major* E2p-like protein, in addition to aligning with *H. sapiens* E2p, an alignment was also performed with the human E3 binding protein (E3BP) (**).

3.2.1 α-Ketoglutarate dehydrogenase (α-KGDH) complex

3.2.1.1 α-Ketoglutarate dehydrogenase isoenzymes (E1k-A and E1k-B)

The E1k subunit of the α-KGDH complex catalyses the oxidative decarboxylation of α-ketoglutarate to form CO_2 and a succinyl group, in a TPP-dependent manner (Perham, 1991; Perham, 2000). Whereas the PDH E1 (E1p) and BCKDH E1 (E1b) subunits in most organisms studied tend to form heterotetramers, E1k proteins are active as homodimers. Recent *in silico* analyses indicate that mammalian species such as *H. sapiens* and *M. musculus* possess an E1k isoenzyme, E1k-like (E1kL) (Bunik & Degtyarev, 2008). *H. sapiens* E1k (*Hs*E1k) and E1kL (*Hs*E1kL) share 75 % sequence identity. Experimental evidence has shown *Hs*E1kL to have different kinetic properties from *Hs*E1k, and different tissue specificity (Bunik *et al.*, 2008).

L. major encodes two potential E1k proteins, *Lmj*E1k-A and *Lmj*E1k-B (see Table 3.1). The genes are located on chromosomes 36 and 27, respectively (see Table 3.1). *Lmj*E1k-A and *Lmj*E1k-B are predicted to be 114.4 kDa and 112.7 kDa, respectively, and as such are both similar in size to *Hs*E1k (115.9 kDa) and *Hs*E1kL (114.5 kDa). Unlike the high sequence identity shared between *Hs*E1k and *Hs*E1kL (75 %), *Lmj*E1k-A shares only 36 % identity to *Lmj*E1k-B, and both have similar identities to *Hs*E1k (35 % and 34 %, respectively) and *E. coli* E1k (*Ec*E1k) proteins (33 % and 32 %, respectively) (see Table 3.1 and Appendix Figure 7.1 and Table 7.1). The two potential *Lmj*E1k isoenzymes are predicted to be mitochondrial by prediction programmes MitoProt (95 % and 75 % confidence

levels, respectively) and TargetP (78 % and 92 % confidence levels, respectively) (see Table 3.1).

The fact that *H. sapiens* possesses the functional E1k isoenzyme E1kL has been discussed two paragraphs above. In human, *in silico* analyses resulted in the identification of another E1k-like protein named dehydrogenase E1 and transketolase domain-containing 1 (*Hs*DHTDK1) protein (Bunik & Degtyarev, 2008). *Hs*DHTDK1 lacks three conserved motifs that are required for Ca^{2+} activation, and in this sense is more similar to bacterial E1k proteins (Bunik & Degtyarev, 2008). ClustalW alignment of *Lmj*E1k-A and *Lmj*E1k-B with different vertebrate E1k, E1kL and DHTDK1 proteins revealed some potentially interesting differences (see Appendix Figure 7.1). Firstly, *Lmj*E1k-A and *Lmj*E1k-B possess most of the required motifs for α-ketoacid dehydrogenase activity, such as lipoyl-domain-binding motifs and TPP-binding motifs (see Appendix Figure 7.1). However, *Lmj*E1k-A and *Lmj*E1k-B lack all three conserved motifs required for binding of Ca^{2+}. In addition, as for DHTDK1, *Lmj*E1k-A and *Lmj*E1k-B have altered α-ketoglutarate substrate binding motifs (see Appendix Figure 7.1). Bunik *et al.* (2008a) discuss this phenomenon in the context of DHTDK1, and hypothesise that DHTDK1 could use substrates other than α-ketoglutarate, such as glyoxylate.

3.2.1.2 Succinyl transferase subunit (E2k)

The E2k subunit of the α-KGDH complex catalyses the transfer of the succinyl group from E1k onto CoA, in a LA-dependent manner (Perham, 1991; Perham, 2000). In all α-KADHs, the E2 components are assembled into either a cubic (24-mer) or a dodecahedral (60-mer) inner core, around which E1 and E3 components bind.

Bioinformatics analyses indicate that *L. major* possesses a single E2k gene encoding the succinyl transferase, *Lmj*E2k. *Lmj*E2k has a relatively high sequence identity to *H. sapiens* E2k (*Hs*E2k) (38 %) and *E. coli* E2k (*Ec*E2k) (39 %) (see Table 3.1). *Lmj*E2k is predicted to be mitochondrial by prediction programmes MitoProt and TargetP (96 % and 78 % confidence levels, respectively) (see Table 3.1).

ClustalW alignment with *H. sapiens* and *E. coli* homologues illustrates that *Lmj*E2k has a single lipoyl-domain at the *N*-terminus. However, the strictly conserved motif

^{42}TDK^{44} in *Ec*E2k is ^{64}SDK^{66} in *Lmj*E2k (see Appendix Figure 7.7), and this is the same for all other trypanosomatid species with sequenced genomes available (*L. infantum, L. braziliensis, T. brucei* and *T. cruzi*). This seemingly kinetoplastid-specific anomaly could be of significance given that the Thr^{42} residue of the ^{42}TDK^{44} motif has been shown to be very important in the reductive succinylation of the E2k-bound lipoyl moiety by E1k (Jones & Perham, 2008). On the other hand, threonine and serine are similar residues in that both are polar and uncharged, and as such may be interchangeable.

*Lmj*E2k has the conserved ^{358}DHRxxDG^{364} motif at the *C*-terminus (see Appendix Figure 7.7), the His^{359} of which is required for succinyl transferase catalytic activity (Reed & Hackert, 1990).

However, unlike bacterial and mammalian E2k and E2b proteins, *Lmj*E2k lacks a strictly conserved Arg^{140} residue (Ciszak *et al.*, 2006) that permits binding to both E1 and E3 proteins, and instead has Lys^{141}, which given its similarity in size and charge to arginine, may fulfil a similar role (see Appendix Figure 7.9).

3.2.2 Pyruvate dehydrogenase (PDH) complex

3.2.2.1 Pyruvate dehydrogenase subunits (E1p-α and E1p-β)

The E1p subunit of the PDH complex catalyses the oxidative decarboxylation of pyruvate to form CO_2 and an acetyl group in a TPP-dependent manner (Perham, 1991; Perham, 2000). Eukaryotic E1p subunits (Wexler *et al.*, 1991) and those from gram-positive bacteria such as *Bacillus* species (Lessard & Perham, 1994) are comprised of heterotetramers, whereas E1p subunits from gram-negative bacteria such as *E. coli* are homodimers (de Kok *et al.*, 1998).

Searching the *L. major* genome for homologues resulted in the identification of two genes, the products of which have been designated *Lmj*E1p-α and *Lmj*E1p-β (see Table 3.1). *Lmj*E1p-α and *Lmj*E1p-β are predicted to be 42.9 kDa and 37.9 kDa, respectively, and are therefore more similar in size to *H. sapiens* E1p-α (*Hs*E1p-α) and *H. sapiens* E1p-β (*Hs*E1p-β) (43.2 kDa and 39.2 kDa, respectively) than to the single *E. coli* E1p (*Ec*E1p) (99.7 kDa) (see Table 3.1). This information (combined with that provided below) is in accordance with *Lmj*E1p-α and *Lmj*E1p-β forming

heteromers. Both *L. major* E1p proteins have high identities to the respective *Hs*E1p-α and *Hs*E1p-β subunits (43 % and 57 %, respectively) (see Table 3.1 and Appendix Figure 7.2 and Figure 7.3). Both proteins are predicted to be mitochondrial by MitoProt and TargetP programmes, with confidence values of 79 % and 47 % respectively, for *Lmj*E1p-α, and 99 % and 87 % respectively, for *Lmj*E1p-β (see Table 3.1).

ClustalW alignment of *Lmj*E1p-α and *Lmj*E1p-β with mammalian homologues revealed that some of the amino acids within conserved motifs are different in *L. major* E1p proteins, however the overall conservation is high. For example, the conserved motif in E1p-α proteins involved in interacting with the thiamine diphosphate (TPP) cofactor, catalysing the decarboxylation of pyruvate, and interacting with the lipoyl-domain of the E2p subunit (^{285}TYRY(H/g)GHSMSDPG^{298} in *Hs*E1p-α) is instead ^{276}**C**YRY**M**GHSMSDP**D**^{289} in *Lmj*E1p-α (see Appendix Figure 7.2). Another example is a motif found in E1p-β proteins that is involved in binding the lipoyl-domain of the E2p subunit (^{109}EFM(T/s)FNFSMQAID^{121} in *Hs*E1p-α), which is instead ^{101}EFMTFNF**A**MQAID^{113} in *Lmj*E1p-β (see Appendix Figure 7.3). The significance of such differences remains unknown, yet they could reflect subtle species-specific adaptations to different substrates for example.

3.2.2.2 Acetyl transferase (E2p) and E3-binding protein (E3BP) subunits

The E2p subunit of the α-PDH complex catalyses the transfer of the acetyl group from E1p onto the acceptor protein coenzyme A (CoA), in a LA-dependent manner (Perham, 1991; Perham, 2000). In all α-KADHs, the E2 components are assembled into either a cubic (24-mer) or a dodecahedral (60-mer) inner core, around which E1 and E3 components bind. Mammals (De Marcucci & Lindsay, 1985) and some yeast (Maeng *et al.*, 1994) (but not bacteria) also encode a structurally similar protein, E3-bindng protein (E3BP). In these organisms, the E2p subunit cannot bind the E3 subunit directly. As such, E3 binds to E3BP and is inserted into the E2p scaffold (Sanderson *et al.*, 1996a; Sanderson *et al.*, 1996b).

Querying the *L. major* genome with *H. sapiens* E2p (*Hs*E2p) and *E. coli* E2p (*Ec*E2p) protein sequences resulted in the identification of a single *Lmj*E2p-encoding gene, and a gene encoding an E2p-like protein (*Lmj*E2pL). The

predicted size of *Lmj*E2p is 48.7 kDa, which is 20 kDa smaller than the *Hs*E2p and *Ec*E2p proteins (see Table 3.1). The size discrepancy can be explained by 20 kDa *N*-terminal extensions in *Hs*E2p and *Ec*E2p proteins compared to *Lmj*E2p (see Appendix Figure 7.8). The large *N*-terminal extensions in *Ec*E2p and *Hs*E2p consist of two additional lipoyl-domains (L1 and L2 lipoyl-domains) (Guest *et al.*, 1985). *Lmj*E2p has 24 % and 18 % sequence identity to *Hs*E2p and *Ec*E2p proteins, respectively (see Table 3.1). Subcellular targeting prediction programmes give different results with regards to *Lmj*E2p subcellular localisation; MitoProt predicts mitochondrial targeting (86 % confidence level), yet TargetP predicts with only moderate confidence that *Lmj*E2p is mitochondrial (52 % confidence level) and with higher confidence (62 %) that *Lmj*E2p possesses a secretory signal peptide (see Table 3.1). Given that E2p is integral to the formation and function of the PDH complex, one would expect *Lmj*E2p to be mitochondrial, and indeed evidence is provided that this is probably the case in *L. major* (see Section 5.4.1). *Lmj*E2p possesses a single lipoyl-domain with the motif ^{60}TDKA^{63}, as well as the *C*-terminal motif ^{619}DHRxxDG^{625}; the former is involved in binding LA and the latter is integral to the transfer of the acetyl group from the E1p subunit to a CoA acceptor protein (see Appendix Figure 7.8) (Reed & Hackert, 1990).

As mentioned, a second gene that shares some homology to *Lmj*E2p, *Lmj*E2pL, is predicted in the *L. major* genome (see Table 3.1). The most obvious candidate for the *Lmj*E2pL protein would be the E3BP. Human and yeast E3BP have reasonable sequence identities to their respective E2p proteins (30 % and 20 %, respectively). Similarly, *Lmj*E2p and *Lmj*E2pL share 18 % amino acid sequence identity. The sequence identities of *Lmj*E2pL to *Hs*E2p and *Hs*E3BP are low (15 % and 16 %, respectively) (see Table 3.1). As for *Lmj*E2p, the mitochondrial targeting prediction by MitoProt is reasonably high (61 % confidence level), whereas TargetP predicts with high confidence (85 %) that *Lmj*E2pL has a secretory signal sequence (see Table 3.1).

In yeast and humans, E3BP has the conserved ^{80}TDKA^{83} motif (*Hs*E3BP numbering), the Lys^{82} of which can be lipoylated by lipoyl transferases (Sanderson *et al.*, 1996b). *Lmj*E2pL also possesses the conserved ^{51}TDKA^{54} motif, and possibly a second functional motif ^{177}TDKA^{180} (see Appendix Figure 7.8). Yeast and human E3BP do not possess acetyl transferase activity, since the strictly conserved His^{620} residue within the motif ^{619}D**H**RxxDG^{625} (*Hs*E2p numbering), is instead Ser^{460} (*Hs*E3BP numbering) (Neagle *et al.*, 1989). *Lmj*E2pL also lacks this

histidine residue and the motif is instead ^{366}**YKS**xxD**T**^{372} (with bold residues representing those residues that are not conserved in *Lmj*E2pL) (see Appendix Figure 7.8). The *H. sapiens* PDH (*Hs*PDH) complex requires E3BP to insert E3 into the E2p scaffold, since *Hs*E2p only has the capacity to interact with E1 and E3BP (Ciszak *et al.*, 2006). Interestingly, *Lmj*E2p has Val^{175} and Arg^{177} residues (see Appendix Figure 7.9), which could permit binding of both E1 and E3 subunits according to predictions made by Cizak *et al.* (2006). However, *Lmj*E2pL does not possess the strictly conserved ^{170}SPAxRNxxE^{178} motif, or the Pro^{192} and Ile^{195} residues (see Appendix Figure 7.9) that have been shown in *Hs*E3BP to make important contacts with the E3 subunit (Ciszak *et al.*, 2006). As such, although *in silico* analyses generally infer that *Lmj*E2pL may be E3BP, experimental evidence would be required to confirm the localisation of *Lmj*E2pL, as well as the nature of its interactions with other subunits of the PDH complex.

3.2.3 Branched chain α-ketoacid dehydrogenase (BCKDH)

3.2.3.1 Branched chain ketoacid dehydrogenase subunits (E1b-α and E1b-β)

The E1b subunit of the BCKDH complex catalyses the oxidative decarboxylation of branched-chain α-ketoacids 2-ketoisocaproate, 2-keto-3-methylvalerate and 2-ketoisovalerate (derived from leucine, isoleucine and valine, respectively) to form CO_2 and an acyl group (isovaleryl, 2-methylbutyryl or isobutyryl, respectively), in a TPP-dependent process (Massey *et al.*, 1976). E1b subunits in human (Aevarsson *et al.*, 2000) and gram-negative bacterium *Pseudomonas putida* (Aevarsson *et al.*, 1999) are heterotetrameric.

Searching the *L. major* genome for homologues resulted in the identification of two genes, *Lmj*E1b-α and *Lmj*E1b-β (see Table 3.1). *Lmj*E1b-α and *Lmj*E1b-β are predicted to be 53.3 kDa and 40.0 kDa, respectively, and are therefore similar in size to *H. sapiens* E1p-α (*Hs*E1p-α) and E1p-β (*Hs*E1p-β) (50.5 kDa and 43.1 kDa, respectively) (see Table 3.1). This information (combined with that provided below) is in accordance with *Lmj*E1b-α and *Lmj*E1b-β forming heteromers. Both *L. major* E1b proteins have high identities to the respective *Hs*E1b-α and *Hs*E1b-β subunits (36 % and 49 %, respectively) (see Table 3.1 and Appendix Figure 7.4 and Figure 7.5). Both proteins are predicted to be mitochondrial by MitoProt and

TargetP programmes, with confidence values of 80 % and 84 %, respectively, for *Lmj*E1b-α, and 85 % and 73 %, respectively, for *Lmj*E1b-β (see Table 3.1).

ClustalW alignment of *Lmj*E1b-α and *Lmj*E1b-β with mammalian homologues revealed that some of the amino acids within conserved motifs of *Lmj*E1b-β are different to those found in mammalian E1b-β proteins; however, the overall conservation is high. For example, the conserved motif in E1b-β proteins involved in interacting with the TPP cofactor (^{146}E[I/m]QF^{149} in *Hs*E1b-β) is instead ^{123}E**V**QF^{126} in *Lmj*E1b-β (see Appendix Figure 7.5). The significance of such differences remains unknown.

3.2.3.2 Branched-chain transacylase subunit (E2b)

The E2b subunit of the BCKDH complex catalyses the transfer of the acyl group (either isovaleryl, 2-methylbutyryl or isobutyryl) from E1b onto the acceptor protein CoA, in a LA-dependent manner (Perham, 1991; Perham, 2000). In all α-KADHs, the E2 components are assembled into either a cubic (24-mer) or a dodecahedral (60-mer) inner core, around which E1 and E3 components bind.

Bioinformatics analyses indicate that *L. major* possesses a single gene encoding the branched-chain transacylase, *Lmj*E2b. *Lmj*E2b has a relatively high sequence identity to *H. sapiens* E2b (*Hs*E2b) (40 %) (see Table 3.1). *Lmj*E2b is predicted to be mitochondrial by prediction programmes MitoProt and TargetP (100 % and 92 % confidence levels, respectively) (see Table 3.1).

ClustalW alignment with mammalian homologues illustrates that *Lmj*E2b has a single lipoyl-domain at the *N*-terminus, with the diagnostic motif ^{86}SDKA^{89}, the Lys^{88} of which is lipoylated by lipoyl transferases (see Appendix Figure 7.10). *Lmj*E2b is likely to be catalytically active, given that the ^{446}DHRxxDG^{452} motif is present, the His^{447} of which is required for acyl transferase catalytic activity (Reed & Hackert, 1990).

However, unlike bacterial and mammalian E2k and E2b proteins, *Lmj*E2k lacks a conserved Arg^{199} residue (Ciszak *et al.*, 2006) that permits binding to both E1 and E3 proteins, and instead has Lys^{297}, which given its similarity in size and charge to arginine, may fulfil a similar role (see Appendix Figure 7.9).

3.2.4 Glycine cleavage complex (GCC)

3.2.4.1 Lipoyl-domain subunit (H-protein)

In the GCC, H-protein is monomeric (Nakai *et al.*, 2003a; Pares *et al.*, 1994), unlike the E2 subunits of α-KADHs, which form 24-mers or 60-mers (Perham, 1991; Reed & Hackert, 1990). However, 27 H-protein monomers are found within the GCC of *Pisum satisvum* (*Ps*GCC) (Oliver *et al.*, 1990). Similarly to E2 subunits, H-protein possesses a lipoyl-domain which binds LA and acts as a "swinging arm" domain that permits access of each catalytic subunit in turn, to the essential lipoyl moiety. Nevertheless, unlike E2 subunits, H-protein bears no catalytic activity (Macherel *et al.*, 1992).

Bioinformatics analyses indicate that *L. major* encodes a single lipoyl-domain protein, *Lmj*H-protein. *Lmj*H-protein has a reasonable sequence identity to *H. sapiens* H-protein (*Hs*H-protein) (36 %) and *E. coli* H-protein (*Ec*H-protein) (37 %) (see Table 3.1). *Lmj*H-protein is predicted to be mitochondrial by prediction programmes MitoProt and TargetP (76 % and 50 % confidence levels, respectively) (see Table 3.1).

ClustalW alignments with mammalian homologues reveals that *Lmj*H-protein possesses the motif ^{76}SV**K**A^{79}, the strictly conserved Lys^{78} of which becomes lipoylated by lipoyl transferases (see Appendix Figure 7.11). Additionally, strongly conserved negatively-charged residues Glu^{48}, Asp^{52} and Glu^{75} that make contacts with P-protein and T-protein (Nakai *et al.*, 2003a) are present, however one glutamate residue is instead Asn^{81} (see Appendix Figure 7.11).

3.2.4.2 Glycine dehydrogenase subunit (P-protein)

The P-protein subunit of the GCC complex is a pyridoxal 5'-phosphate (PLP)-dependent enzyme that catalyses the oxidative decarboxylation of glycine to form CO_2 and aminomethylated H-protein (Douce *et al.*, 2001). Eukaryotic P-proteins (for example, human) (Kume *et al.*, 1991) and some prokaryotic P-proteins (for example, *E. coli*) (Okamura-Ikeda *et al.*, 1993) occur as homodimers, whereas most prokaryotic P-proteins typically form heterotetramer P-protein complexes (for example, *Thermus thermophilus*) (Nakai *et al.*, 2003b). In the *Ps*GCC two

homodimers of P-protein are necessary and sufficient for glycine decarboxylase activity (Oliver *et al.*, 1990).

The *L. major* genome possesses one gene encoding *Lmj*P-protein. The *L. major* homologue has relatively high sequence identity to *E. coli* P-protein (*Ec*P-protein) (46 %) and *H. sapiens* protein (*Hs*P-protein) (47 %) (see Table 3.1). *Lmj*P-protein is predicted to be mitochondrial by MitoProt and TargetP, with confidence levels of 95 % and 86 %, respectively (see Table 3.1). The mitochondrial localisation of *Lmj*P-protein has been verified experimentally in *L. major* promastigotes by over-expression of an *Lmj*P-protein-GFP reporter construct (Scott *et al.*, 2008).

*Lmj*P-protein contains all of the residues known to be important in lining the active site (such as the ^{541}PLGSCTMKLN^{550} motif), and for interaction with lipoamide (Gln^{354} and His^{638}) (see Appendix Figure 7.6). The only potentially important differences are two changes in residues involved in binding H-protein, whereby *Lmj*P-protein has Lys^{357} instead of Arg^{326}, and at another position has Arg^{360} instead of Lys^{329} (see Appendix Figure 7.6). However, the most important factor involved in binding H-protein is thought to be the positive charge within this motif (Nakai *et al.*, 2005), and the changes just outlined would keep the charge the same as that found in *Ec*P-protein and *Hs*P-protein. Indirect evidence has proven that *Lmj*P-protein (and as such the GCC in *L. major*) is active (Scott *et al.*, 2008); inactivation of the *Lmj*P-protein gene in promastigotes resulted in a significant loss of [2-^{14}C] glycine incorporation into DNA, which is ordinarily derived from 5,10-methylene-tetrahydofolate (5,10-CH_2-THF) produced via the GCC (Scott *et al.*, 2008).

3.2.4.3 Aminomethyl transferase subunit isoenzymes (T-protein-A and T-protein-B)

The T-protein subunit of the GCC complex catalyses the transfer of the methylene group (CH_2) from aminomethylated H-protein to a tetrahydrofolate (THF) acceptor, to produce 5,10-CH2-THF and ammonia (NH_3) (Douce *et al.*, 2001). T-protein is a monomeric protein, and forms a 1:1 ratio with H-protein (Okamura-Ikeda *et al.*, 1993). A total of 9 T-protein monomers have been determined for the *Ps*GCC (Oliver *et al.*, 1990).

Searching the *L. major* genome revealed two potential T-protein genes, which are arranged in tandem on chromosome 36. The GCC studied in all other organisms involves one T-protein. As such, it is possible that there is an annotation/sequencing mistake in the *L. major* genome project. However, *L. infantum* has two predicted T-protein genes, yet *L. braziliensis* has only one, which potentially indicates that a gene duplication event occurred in *L. major* and *L. infantum* but not in *L. braziliensis*. The two *L. major* T-proteins, *Lmj*T-protein-A and *Lmj*T-protein-B, are identical, except that *Lmj*T-protein-A has a *C*-terminal extension of 17 amino acids, which is not present in either the human or *E. coli* homologues. The two *L. major* homologues share moderate sequence identity to *E. coli* T-protein (*Ec*T-protein) (29 %) and *H. sapiens* T-protein (*Hs*T-protein) (38 %) (see Table 3.1). Given that both *Lmj*T-protein-A and *Lmj*T-protein-B are identical apart from the terminal 17 amino acids, they possess exactly the same *N*-terminal sequences involved in subcellular targeting. Mitochondrial targeting predictions by MitoProt and TargetP are poor (17 % and 12 %, respectively) (see Table 3.1). However, given the experimentally-proven mitochondrial localisation and function of *Lmj*P-protein (Scott *et al.*, 2008), one would certainly expect *Lmj*T-protein (isoform A or B) to also be mitochondrial.

Both *Lmj*T-protein-A and *Lmj*T-protein-B possess the conserved motifs (Lee *et al.*, 2004) involved in binding H-protein, such as the ^{195}GYTGExGxE^{203} motif (see Appendix Figure 7.12). Additionally, the ^{225}GLGARDx$_2$RxEAx$_3$LYG^{244} motif that surrounds the aliphatic portion of LA, is present in *Lmj*T-protein-A and *Lmj*T-protein-B (see Appendix Figure 7.12). Lastly, *Lmj*T-protein-A and *Lmj*T-protein-B possess the conserved motif ^{318}Gx$_2$TSGx$_2$SPxL^{329} that makes key contacts with the T-protein cofactor, THF (see Appendix Figure 7.12).

3.2.5 Dihydrolipoamide dehydrogenase subunit (LipDH)

Most eukaryotic organisms and bacteria possess a single gene encoding lipDH. The LipDH protein is a subunit of α-KADHs (E3 subunit) and the GCC (L-protein) (Perham, 2000). In both complexes, the role of LipDH is to re-oxidise DHLA to LA with the concomitant generation of reduced NADH, in a process that requires the cofactor flavin adenine dinucleotide (FAD). Eukaryotic (for example, human) (Ciszak *et al.*, 2006) and prokaryotic (for example, *Azotobacter vinelandii*) (Mattevi *et al.*, 1991) LipDH proteins form tightly-bound homodimers. One L-protein dimer is associated with the GCC (Oliver *et al.*, 1990), whereas α-KADHs require 12

copies of E3. Unlike in the α-KGDH and BCKDH complexes (and the PDH complex in bacteria), where E3 homodimers directly bind to the E2 scaffold, interaction of E3 with E2p in the mammalian PDH complex requires E3BP. There is a 1:1 stoichiometry of E3 homodimer binding to E3BP monomer, and 12 E3BP monomers bind to E2p (Ciszak *et al.*, 2006).

Querying the *L. major* genome with *H. sapiens* LipDH (*Hs*LipDH) and *E. coli* LipDH (*Ec*LipDH) proteins resulted in four significant hits. Analysis of all four *L. major* predicted proteins revealed that only one – *LmjF*32.3310 – possessed all motifs diagnostic of LipDH enzymes (see Appendix Figure 7.13). *Lmj*LipDH has relatively high sequence identity to human and *E. coli* LipDH homologues (47 % and 38 %, respectively). However, the confidence of mitochondrial targeting is very low if one assumes the open reading frame of the gene to begin at a classical ATG codon (9 % for MitoProt and 21 % for TargetP) (see Table 3.1). If the ORF is presumed to start at a GTG codon that is 75 bp upstream of the ATG start codon, the resulting protein is predicted to be mitochondrial with higher confidence values (23 % for MitoProt and 83 % for TargetP) (see Table 3.1). Nevertheless, I was not able to identify any studies indicating that GTG is a valid start codon in *Leishmania*. Given the high degree of confidence in mitochondrial localisation predicted by MitoProt and TargetP for the majority of components of the α-KADHs and GCC in *L. major*, combined with the experimental evidence showing *Lmj*P-protein to be mitochondrial (Scott *et al.*, 2008), one would expect *Lmj*LipDH to be mitochondrial also.

ClustalW alignment of *Lmj*LipDH with *Hs*LipDH and *Ec*LipDH proteins revealed that *Lmj*LipDH possesses all motifs diagnostic of LipDH enzymes. Firstly, *Lmj*LipDH possesses the $^{43}GxGx_2G^{48}$ motif, which is involved in binding of the pyrophosphate moiety of FAD (see Appendix Figure 7.13) (Mattevi *et al.*, 1991). Secondly, the redox active $^{75}CLNVGC^{80}$ motif is present in *Lmj*LipDH (see Appendix Figure 7.13) (Thorpe & Williams, 1976). Thirdly, *Lmj*LipDH has the conserved glycine-rich $^{215}GxGxIGxEx_3Vx_4G^{231}$ motif, which is important in binding NADH (see Appendix Figure 7.13) (Mattevi *et al.*, 1992). Lastly, *Lmj*LipDH possesses the acid/base catalyst motif $^{480}HPTx_2E^{485}$, which is an integral component of the active site (see Appendix Figure 7.13).

3.2.6 α-KADH kinases and phosphatases

In mammals, but not lower eukaryotes such as yeast or prokaryotes, α-KADH complex activities are regulated by specific mitochondrial kinases and phosphatases.

3.2.6.1 α-KADH kinases

In human, four highly homologous isoforms of PDH kinase (PDK-1–4), and one BCKDH kinase (BCKDK) have been identified, although the α-KGDH complex has no known associated kinases or phosphatases. PDK forms homodimers, whereas BCKDK is more stable in a tetrameric formation (Machius *et al.*, 2001; Wynn *et al.*, 2000). PDK and BCKDK are mitochondrial serine kinases belonging to the ATPase/kinase superfamily (Bowker-Kinley & Popov, 1999; Dutta & Inouye, 2000), and are distantly related to the protein histidine kinase (PHK) family (Koretke *et al.*, 2000).

PDK and BCKDK down-regulate the activity of their respective α-KADHs by phosphorylating specific serine residues on E1 subunits (Roche *et al.*, 2001; Roche & Hiromasa, 2007). It has been shown that α-KADH kinases bind to the lipoyl-domain of E2p (specifically to the second lipoyl-domain, L2) and E2b subunits (Bao *et al.*, 2004; Roche *et al.*, 2003; Wynn *et al.*, 2000), and that the level of phosphorylation of E1 subunits is dependent upon the lipoylation state of the lipoyl-domain (Roche & Hiromasa, 2007).

In order to identify potential homologues of α-KADH kinases, the *L. major* genome database was queried with *H. sapiens* PDK-1-4 and BCKDK protein sequences. In total, two potential genes encoding PDK/BCKDK were identified; PDH kinase (lipoamide) (*LmjF*24.0010) and a phosphoprotein-like protein (*LmjF*20.0280) (see Table 3.2). For simplicity, the genes and gene products will be referred to as *LmjF*24.0010 and *LmjF*20.0280. *LmjF*24.0010 (50.9 kDa) and *LmjF*20.0280 (55.2 kDa) predicted proteins are larger than *H. sapiens* PDK-1-4 (49.2 kDa, 46.2 kDa, 46.9 kDa and 46.5 kDa, respectively) and BCKDK (46.4 kDa) (see Table 3.2). *LmjF*20.0280 protein has higher sequence identity (20 – 22 %) to the five human α-KADH kinases than does *LmjF*24.0010 protein (15 – 17 %) (see Table 3.2). Mitochondrial targeting predictions for *LmjF*20.0280 and *LmjF*24.0010 proteins by

MitoProt and Target P are high (92 % and 82 % compared to 83 % and 48 %, respectively) (see Table 3.2).

ClustalW alignment of the *LmjF*20.0280 protein with the five human α-KADH kinases revealed that *LmjF*20.0280 possesses the strictly conserved N-, G1- and G2 boxes (see Appendix Figure 7.14), which constitute the ATP domain that is characteristic of the ATPase/kinase superfamily (Bilwes *et al.*, 1999). However, the F-box, which is present in all human α-KADH kinases (Bowker-Kinley & Popov, 1999; Wynn *et al.*, 2000), is not found in the *LmjF*20.0280 protein. *LmjF*24.0010 lacks strictly conserved residues comprising the N-, G1- and G2 boxes; Asp^{364} instead of Asn (N box), Ala^{404} instead of glycine (G1 box) and $^{438}SxPxR^{442}$ instead of GxGxG (G2 box) (see Appendix Figure 7.15). As such, contrary to the current GeneDB annotation of *LmjF*24.0010 as the PDH kinase, it is more likely, based on the conservation of ATP-binding motifs, that *LmjF*20.0280 is a PDH/BCKDH kinase.

However, compared to human PDK-1-4, both *LmjF*20.0280 and *LmjF*24.0010 proteins completely lack the *C*-terminal portion containing the motif $^{416}EAxDWx_2PSxEP^{427}$ (see Appendix Figure 7.14 and Figure 7.15), which is involved in binding the L2 domain of the PDH E2 (Kato *et al.*, 2005). Interestingly, human BCKDK also lacks the *C*-terminal tail that is found in all four PDK isoforms (Kato *et al.*, 2005). Unfortunately, structural information is not available with regards to how human BCKDK binds to the lipoyl-domain of E2b. As such, the key residues involved are not known, and therefore it is not possible to predict whether *LmjF*20.0280 may be a BCKDH kinase.

	Predicted sizes (kDa)		Systematic Name (GeneDB)	Targeting Predictions		Identity (%) to *H. sapiens*				
Subunit	***H. sapiens***	***L. major***	***L. major***	**Mitoprot**	**Target P**	**PDK-1**	**PDK-2**	**PDK-3**	**PDK-4**	**BCKDK**
PDK-1	49.2 (NP_002601)	N/A	N/A	N/A	N/A	N/A	64	62	61	27
PDK-2	46.2 (NP_002602)	N/A	N/A	N/A	N/A	64	N/A	65	63	25
PDK-3	46.9 (NP_005382)	N/A	N/A	N/A	N/A	62	65	N/A	61	23
PDK-4	46.5 (NP_002603)	N/A	N/A	N/A	N/A	61	63	61	N/A	24
BCKDK	46.4 (NP_005872)	N/A	N/A	N/A	N/A	27	25	23	24	N/A
Phosphoprotein-like	N/A	50.9	LmjF20.0280	92 % MT	82 % MT	22	21	21	20	21
PDH (lipoamide) kinase	N/A	55.2	LmjF24.0010	83 % MT	48 % MT, 79 % SP	16	15	15	15	17

Table 3.2 Sequence analyses of potential *L. major* PDH/BCKDH kinases
Predicted *L. major* homologues of subunits of PDK/BCKDK were identified as described in Section2.3. Red proteins are those which have been identified in human. Blue proteins are predicted *L. major* homologues of the human kinases. In the 'Protein Size' column, accession numbers of *H. sapiens* proteins are shown in brackets. Mitochondrial targeting predictions were carried out by analysing *L. major* protein sequences with MitoProt or TargetP. The results obtained indicate the percent likelihood of mitochondrial targeting (MT). Some proteins were predicted to possess signal peptides (SP). Sequence identities shared between human PDK/BCKDK proteins and *L. major* predicted kinases were determined using Vector NTI (See Section2.3). 'N/A' is an acronym for 'non applicable'.

3.2.6.2 α-KADH phosphatases

In human, two PDH phosphatases (PDP) exist; each shares the same regulatory subunit (PDPr), but differs in the catalytic subunit (PDP-1c or PDP-2c) (Huang *et al.*, 1998; Lawson *et al.*, 1993; Lawson *et al.*, 1997; Teague *et al.*, 1982). BCKDH phosphatase (BCKDP or PTMP) activity was first documented in 1984 (Damuni *et al.*, 1984; Damuni & Reed, 1987), yet only recently has the human BCKDP been cloned and partially characterised; although only one isoform of the catalytic subunit has been identified, and no regulatory subunit is known (Joshi *et al.*, 2007). The role of PDP and BCKDP is to catalyse the dephosphorylation of their respective enzyme complexes, and as such to repress the down-regulation imposed upon PDH and BCKDH by PDK and BCKDK, respectively. Based upon some strictly conserved motifs and the crystal structure of PDP-1c (Vassylyev & Symersky, 2007), PDP-1c, PDP-2c and BCKDP are related to the PPM family of serine/threonine phosphatases, whose defining member is PP2C (Barford *et al.*, 1998; Bork *et al.*, 1996). Although the sequence identities shared between human PP2C-1/PP2C-2 and PDP-1c/PDP-2c/BCKDP are low (21 – 24 %), 10/11 of the sequence motifs characteristic of PP2C enzymes are conserved. Information is available with regards to the binding of PDP-1c to the L2 domain of E2p (Vassylyev & Symersky, 2007); however the residues involved are not conserved in BCKDP. As such, no structural data exists describing the mode of BCKDP substrate binding and catalysis.

In order to identify potential homologues of α-KADH phosphatases, the *L. major* genome database was queried with *H. sapiens* PDP-1c, PDP-2c, PDPr and BCKDP protein sequences. Interestingly, significant hits were not obtained from BLAST analysis with PDP subunit sequences. However, querying with the human BCKDP sequence retrieved seven significant hits from *L. major* GeneDB (*LmjF*34.2510, *LmjF*30.0380, *LmjF*36.0530, *LmjF*25.0750, *LmjF*32.1690, *LmjF*15.0170 and *LmjF*27.2320) (see Table 3.3). The seven predicted proteins are all PP2C-like or PP2C-putative, which infers that some of the residues involved in PP2C activity are conserved in these proteins. Although overall sequence identities to PDP-1c, PDP-2c and BCKDP are low (7 – 20 %), the predicted proteins have a high degree of conservation within the 10/11 domains known to be important in the functioning of PP2C and PP2C-like enzymes (data not shown) (Barford *et al.*, 1998; Bork *et al.*, 1996), and catalytically-important residues are

conserved. Mitochondrial targeting predictions vary, although confidence values are high for *LmjF*30.0380 and *LmjF*32.1690 (see Table 3.3).

	Predicted sizes (kDa)		Systematic Name (GeneDB)	Targeting Predictions		Identity (%) to *H. sapiens*		
Subunit	***H. sapiens***	***L. major***	***L. major***	**Mitoprot**	**Target P**	**PDP-1c**	**PDP-2c**	**BCKDP**
PDP-1c	61.0 kDa (NP_060914)	N/A	N/A	N/A	N/A	N/A	47	15
PDP-2c	60.0 kDa (NP_065837)	N/A	N/A	N/A	N/A	47	N/A	17
PDPr	99.4 kDa (NP_060460)	N/A	N/A	N/A	N/A	N/A	N/A	N/A
BCKDP	41.0 kDa (NP_689755)	N/A	N/A	N/A	N/A	15	17	N/A
PP2C-like	N/A	29.8*	LmjF34.2510	0 % MT	4 % MT, 92 % OT	12	10	16
	N/A	47.7**		0 % MT	9 % MT, 88 % OT	13	11	15
PP2C-putative	N/A	41.7*	LmjF30.0380	86 % MT	35 % MT, 56 % OT	10	9	16
	N/A	64.6**		99 % MT	88 % MT	12	9	15
PP2C-like	N/A	32.6	LmjF36.0530	3 % MT	11 % MT, 88 % OT	10	10	20
PP2C-putative	N/A	45	LmjF25.0750	41 % MT	21 %, 79 % OT	12	7	11
PP2C-putative	N/A	68.3	LmjF32.1690	97 % MT	24 % MT, 71 % SP	15	11	12
PP2C-putative	N/A	44.3	LmjF15.0170	40 % MT	24 % MT, 84 % OT	14	12	19
PP2C-like	N/A	34.4*	LmjF27.2320	45 % MT	20 % MT, 81 % OT	13	11	17
	N/A	45.7**		0 % MT	6 % MT, 94 % SP	17	13	17

Table 3.3 Sequence analyses of potential *L. major* PDH/BCKDH phosphatases
Predicted *L. major* homologues of subunits of PDP/BCKDP were identified as described in Section2.3. Red proteins are those which have been identified in human. Blue proteins are predicted *L. major* homologues of the human phosphatases. In the 'Protein Size' column, accession numbers of *H. sapiens* proteins are shown in brackets. Three of the *L. major* genes predicted by GeneDB do not correlate with ORFs predicted by Vector NTI. In such cases, translated protein from GeneDB predictions (*) as well as translated protein from Vector NTI predictions (**) were analysed. A possible explanation for these size discrepancies is that the predicted ATG start codon provided by GeneDB is wrong, and is in fact located further upstream. Mitochondrial targeting predictions were carried out by analysing *L. major* protein sequences with MitoProt or TargetP. The results obtained indicate the percent likelihood of mitochondrial targeting (MT). Some proteins were predicted to possess signal peptides (SP), or an alternative processing type (OT). Sequence identities shared between human PDP/BCKDP proteins and *L. major* predicted phosphatases were determined using Vector NTI (See Section2.3). 'N/A' is an acronym for 'non applicable'.

3.3 Lipoylation patterns in *L. major*

The *in silico* analyses discussed in Section 3.2 indicate that *L. major* encodes machinery to form functional α-KADHs and the GCC. In addition, it was determined that the E2 subunits of α-KADHs and the H-protein of the GCC all have one lipoyl-domain at their *N*-termini. Lipoylation of the E2 subunits and H-protein is essential for activity of these complexes. Therefore, in order to address the question as to whether α-KADHs and the GCC were potentially active in *L. major*, protein lysates of promastigotes and amastigotes were probed by western blot with a polyclonal antibody that specifically detects protein-bound LA. This assay has previously been used successfully in other organisms, for example, *T. gondii* and *P. falciparum* (Allary *et al.*, 2007; Crawford *et al.*, 2006; Gunther *et al.*, 2007).

In order to correlate promastigote growth phase with lipoylation pattern, wild-type parasites were cultured in standard HOMEM + 10 % FCS medium and growth readings were taken every 24 h (see Figure 3.1A), and soluble protein was harvested every 24 h during 48–120 h growth (see Figure 3.1B). In order to analyse the lipoylation pattern in amastigotes, lesions from two mice infected with *L. major* were isolated and amastigotes purified (isolations carried out by Mrs Denise Candlish, University of Glasgow) and soluble protein was extracted, as described in Section 2.2.3.2. Western blotting with α-LA antibody, which detects protein-bound LA, resulted in just four prominent bands in promastigotes. The observed band sizes (calculations are described in Section 2.5.5) are 14.4 kDa, 40.6 kDa and 47.3 kDa (see Figure 3.1B), which combined with predicted sizes after cleavage of mitochondrial targeting peptides (see Table 3.4), correlates well with the predicted sizes of *Lmj*H-protein (15.2 kDa), *Lmj*E2k (41.7 kDa) and *Lmj*E2p (48.7 kDa). The 55 kDa band does not correlate with the predicted size of *Lmj*E2b (50.2 kDa), although this band will be referred to as *Lmj*E2b henceforth as it is the only remaining protein that has the capacity to be lipoylated (see Section 6.2.3).

The banding patterns observed in each of the two amastigote protein lysates are similar to one another, and share some similarities with those observed in promastigotes (see Figure 3.1C). The 47.8 kDa band most likely corresponds to the *Lmj*E2p (which is 47.3 kDa in the promastigote blot), although a band larger

than 50 kDa (assumedly corresponding to *Lmj*E2b) is not present in amastigote protein lysates (see Figure 3.1C). One of two bands (39.4 kDa and 41.4 kDa) could be *Lmj*E2k (which is 40.6 kDa in the promastigote blot). Although the nature of the second band is not clear, it could correspond to the *Lmj*E2pL protein (similar to the human E3BP), since sequence analyses identified *Lmj*E2pL as possessing at least one lipoyl-domain (see Section 3.2.2.2). Two bands of 15.5 kDa and 13.4 kDa are similar in size to the 14.4 kDa band observed in promastigote lysates, which most likely corresponds to *Lmj*H-protein (see Table 3.4). It should also be noted that the protocol for isolating *L. major* amastigotes from an infected mouse is such that the final amastigote preparation contains some contamination with mouse proteins. On the other hand, none of the expected band sizes of mouse α-KADH E2 subunits (see Table 3.4) are observed in Figure 3.1C; suggesting that all lipoylated proteins observed are derived from *L. major*.

Figure 3.1 shows that during exponential growth (48 h) promastigotes primarily lipoylate E2k. Also, lipoylation of H-protein is very low. During early (72 – 96 h) and late-stationary phase (120 h), lipoylation of all three α-KADH E2 subunits and H-protein is apparent. Metacyclic promastigotes purified from mixed promastigote cultures have a similar lipoylation pattern to exponential growth phase parasites; only E2k being lipoylated to an appreciable extent. This indicates that metacyclic promastigotes are metabolically different from procyclic promastigotes. The consistent lipoylation of E2k throughout the 120 h time course indicates that α-ketoglutarate (derived from proline or glutamine) is an important carbon source for energy production, and most likely represents an important entry point into the tricarboxylic acid (TCA) cycle. Pyruvate and branched-chain α-ketoacids could be important during stationary phase, whereas metacyclic promastigotes apparently do not require BCKDH or PDH activity. Lipoylation of H-protein follows a similar pattern to that of E2b and E2p, potentially indicating that production of one-carbon CH_2 units via the GCC is important only at specific times during promastigote development. In fact, it has been experimentally determined that the GCC is not essential in *L. major* promastigotes grown in a similar medium to HOMEM + 10 % FCS (importantly, with the same glycine and serine concentrations) (Scott *et al.*, 2008). In terms of amastigotes, the lipoylation profile is not as clear, although if the bands predicted to be E2p and E2k are correct, the lipoylation intensities of these two proteins are similar, which is in contrast to the situation observed in promastigotes.

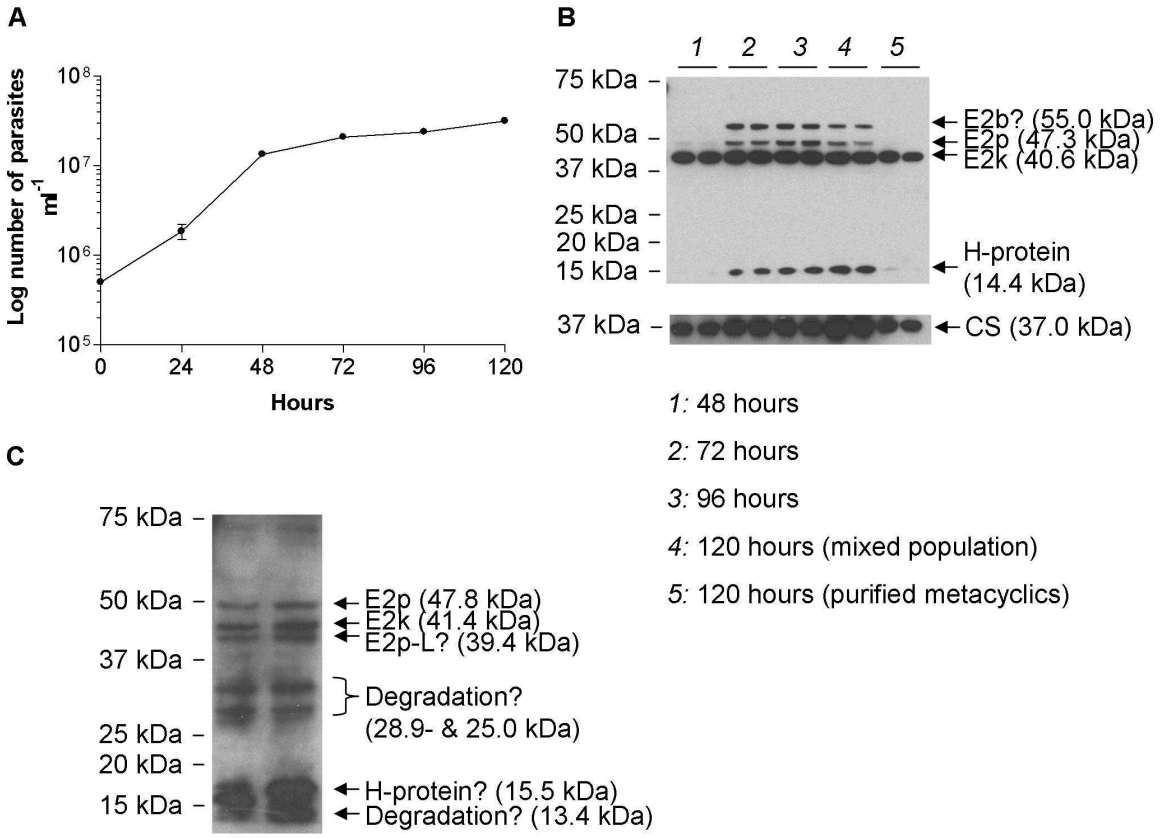


Figure 3.1 Lipoylation of α-KADH E2 subunits and H-protein of GCC in *L. major* promastigotes and amastigotes

A, In order to observe lipoylation patterns in promastigotes, cultures were initiated with starting densities of 5 x 10^5 parasites ml^{-1}. Growth readings were taken every twenty-four hours to produce a growth curve. *B*, Approximately 1 x 10^8 cells were harvested from the cultures at time points from 48 – 120 h. Cells were lysed, and the protein concentration of soluble fractions determined by Bradford assay. At 120 h, in addition to preparing soluble protein from mixed cultures, metacyclic promastigotes were purified from the cultures using peanut agglutinin negative selection, and soluble protein isolated. 10 μg protein from each time point was loaded on a 4 – 12 % SDS-PAGE gradient gel. Western blotting was carried out with α -LA antibody at 1/6,000 and α-cysteine synthase (α-CS) antibody to assess loading, at 1/5,000. Each time point was carried out in duplicate. *C,* Western blot of 30 μg of two different amastigote protein lysates using α-LA antibody at 1/6,000. Observed band sizes (shown in brackets) were calculated as described in Section 2.5.5. Predicted (full-length) sizes of E2k, E2p and E2b are 41.7 kDa, 48.7 kDa and 50.2 kDa, respectively. The predicted size of full-length H-protein is 15.2 kDa (see Table 3.1).

			Predicted protein sizes (kDa)				
Lipoylated protein	**Species**	**Accession number**	**FL**	**MitoProt**	**TargetP**	**Proven size (kDa)**	**Reference**
	Lm	LmjF35.4720	15.2	13.1	13.1	13.4 - 15.5	This work (western blots with α-LA antibody)
H-protein	*Mm*	NP_080848	18.6	12.3	14.7	N/A	N/A
	Hs	NP_004474	18.9	12.3	16.8	N/A	N/A
	Lm	LmjF28.2420	41.7	39.9	33.4	39.4 - 41.4	This work (western blots with α-LA antibody)
E2k	*Mm*	NP_084501	49.0	43.8	41.6	N/A	N/A
	Hs	NP_001924	48.8	41.9	34.8	48	Migliaccio *et al.* (1998)
	Lm	LmjF36.2660	48.7	45.9	45.7	47.3 - 47.8	This work (western blots with α-LA antibody)
E2p	*Mm*	NP_663589	67.9	66.6	58.8	67	Nakagome *et al.* (2007)
	Hs	NP_001922	69.0	N/A	59.6	70	Yeaman *et al.* (1988), Migliaccio et al. (1998)
	Lm	LmjF05.0180	50.2	44.3	39.8	55	This work (western blots with α-LA antibody)
E2b	*Mm*	NP_034152	53.2	46.8	49.7	N/A	N/A
	Hs	NP_001909	53.5	46.3	50.9	52	Migliaccio *et al.* (1998)
	Lm	LmjF21.0550	46.1	N/A	37.9	39.4 - 41.4?	This work (western blots with α-LA antibody)
E3BP/E2p-L	*Mm*	NP_780303	54.0	50.0	49.0	N/A	N/A
	Hs	NP_001128496	51.5	49.8	N/A	50	Jilka *et al*. (1996)

Table 3.4 Predicted- and proven molecular masses of α-KADH E2 subunits and the H-protein of the GCC in *L. major*, *H. sapiens* and *M. musculus*
Predicted (full-length, FL) protein sequences of human (*Hs*), *M. musculus* (*Mm*) and *L. major* (*Lmj*) α-KADH E2 proteins and the H-protein of the GCC were obtained as described in Section 2.3. Predicted cleavage sites for mitochondrial transit were made by MitoProt and TargetP, and the subsequent processed protein sizes were recorded. Where possible, the proven molecular weights of the different apoproteins are given, along with the associated authors who published the data. The proven sizes for *L. major* are based upon α-LA antibody-probed western blots, and the range of potential sizes given are based upon results obtained from western blots of both promastigote and amastigote protein lysates with α-LA antibody (see Figure 3.1B and C). The presence of a lipoylated E2pL protein (which could be E3BP) in amastigotes is questionable, and as such is marked with '?' after the calculated protein size.

3.4 Sequence analyses of lipoylating proteins

Given that lipoylation of α-KADHs and the GCC occurs in promastigotes, the next question to be addressed was whether lipoylating proteins are found within the *L. major* genome. Potential homologues for biosynthesis enzymes LIPB and LIPA and the salvage enzyme LPLA were identified (as described in Section 2.3) (see Table 3.5).

3.4.1 LIPB

LIPB catalyses the first step involved in the biosynthesis of LA from octanoyl-acyl carrier protein; the transfer of an octanoyl-group from a phosphopantetheine-bound cofactor of ACP onto E2- and H-protein apoproteins to form E2/H-protein-octanoamide.

Searching the *L. major* genome revealed the presence of three potential genes encoding LIPB (*Lmj*LIPB); *LmjF*36.3080, *LmjF*07.1060 and *LmjF*31.1070. *LmjF*07.1060 is most likely to be LPLA (*Lmj*LPLA) (see Section 3.4.3), *LmjF*31.1070 is most similar to biotin protein ligase (bpl), and *LmjF*36.3080 is the most likely candidate for *Lmj*LIPB. *Lmj*LIPB (29.5 kDa) is similar in size to *H. sapiens* LIPB (*Hs*LIPB) (25.2 kDa) and *E. coli* LIPB (*Ec*LIPB) (23.9 kDa) (see Table 3.5). Compared to *Hs*LIPB and *Ec*LIPB, *Lmj*LIPB (and *T. brucei* LIPB (*Tb*LIPB)) has a circa 20 amino acid *C*-terminal extension (see Figure 3.2). Both *Hs*LIPB and *Lmj*LIPB (and *Tb*LIPB) have approximately 5-10 amino acid *N*-terminal extensions, which contain targeting signals for mitochondrial transit (see Figure 3.2). *Lmj*LIPB has reasonable sequence identity to *Hs*LIPB (21 %) and *Ec*LIPB (21 %). *Lmj*LIPB is predicted to be mitochondrial by MitoProt and TargetP programmes, with confidence values of 60 % and 82 %, respectively (see Table 3.5).

The crystal structure of *M. tuberculosis* LIPB (*Mt*LIPB) complexed with decanoic acid (DA) highlighted residues that are important in the binding of DA/octanoic acid (OA) to form the LIPB-DA/OA intermediate (Ma *et al.*, 2006). ClustalW alignment of *Lmj*LIPB with *Mt*LIPB, *Hs*LIPB and *Ec*LIPB illustrated the presence of strictly conserved residues in *Lmj*LIPB (see Figure 3.2). Firstly, residues that form a hydrophobic tunnel for binding of the aliphatic portion of DA/OA in *Mt*LIPB (Ma *et*

al., 2006) are conserved in *Lmj*LIPB. For example, the motif ^{80}RGGx$_2$TxHx$_7$Y^{95} is present, although the importance of Tyr^{95} is questionable given that in *Hs*LIPB the residue is instead His^{100} (see Figure 3.2). Most importantly, *Lmj*LIPB possesses the strictly conserved Lys^{157} and Cys^{191} residues (see Figure 3.2). In *Mt*LIPB, Cys^{176} (Cys^{191} in *Lmj*LIPB) has been shown to form an essential covalent thioether linkage with the C3 atom of DA/OA, producing the LIPB-DA/OA intermediate (Ma *et al.*, 2006). Importantly, LPLA enzymes lack this Cys residue, and this is indeed also the case for the predicted *Lmj*LPLA (see Section 3.4.3). The Lys^{142} residue in *Mt*LIPB (Lys^{157} in *Lmj*LIPB) is conserved in all biotin protein ligases (BPL), LPLA and LIPB proteins (Reche, 2000). However, the role of the invariable lysine residues is thought to differ in LIPB and LPLA enzymes (Ma *et al.*, 2006).

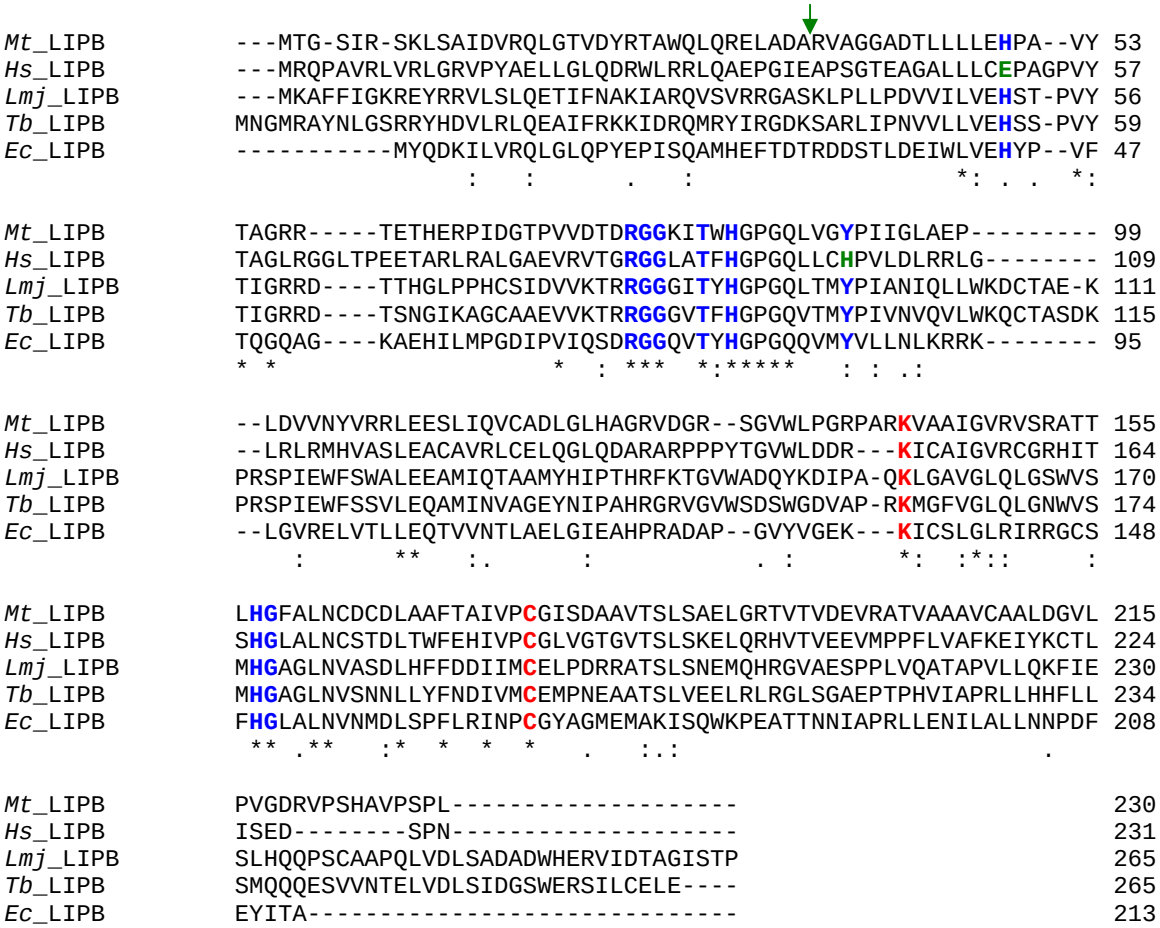


Figure 3.2 LIPB protein alignment
ClustalW alignment of *L. major* LIPB (*Lmj*LIPB) (accession number *LmjF*36.3080) with homologues in *T. brucei* (*Tb*LIPB) (accession number *Tb*11.01.1160), *H. sapiens* (*Hs*LIPB) (accession number A6NK58), *E. coli* (EcLIPB) (accession number AAC73731) and *M. tuberculosis* (*Mt*LIPB) (accession number CAA94273) was carried out. The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues highlight the strictly conserved lysine and Cys residues that comprise the catalytic dyad within the active site of LIPB enzymes (Ma *et al.*, 2006). Blue residues are those that are involved in binding decanoic acid in *Mt*LIPB (Ma *et al.*, 2006). Green residues highlight residues in *Hs*LIPB that are not conserved in other species. MitoProt programme predicts the targeting peptide for mitochondrial transit as being 34 amino acids (green arrow, LIPB$_{T1}$) (see Section 3.5.2).

3.4.2 LIPA

LIPA is an [4Fe – 4S] $^{+}$ cluster-containing protein that catalyses the second step involved in the biosynthesis of LA; via a radical S-adenosylmethionine (SAM)-dependent mechanism LIPA catalyses the replacement of two hydrogen atoms with sulphur atoms, at positions C6 and C8 of octanoamide-E2/H-protein, to form lipoamide-E2/H-protein (Booker *et al.*, 2007).

Searching the *L. major* genome revealed the presence of one potential gene encoding LIPA (*Lmj*LIPA). *Lmj*LIPA (45.9 kDa) is similar in size to *H. sapiens* LIPA (*Hs*LIPA) (41.9 kDa), yet 10 kDa larger than the *E. coli* homologue (*Ec*LIPA) (36.1 kDa) (see Table 3.5). Compared to *Hs*LIPA and *Ec*LIPA, *Lmj*LIPA has a circa 17 amino acid *C*-terminal extension that is rich in hydrophobic amino acids alanine and glycine (see Figure 3.3). Both *Hs*LIPA and *Lmj*LIPA (and *Tb*LIPA) have approximately 40 amino acid *N*-terminal extensions, which contain targeting signals for mitochondrial transit (see Figure 3.3 and Section 5.2). *Lmj*LIPA has reasonable sequence identity to *Hs*LIPA (40 %) and *Ec*LIPA (33 %). *Lmj*LIPA is predicted to be mitochondrial by MitoProt and TargetP programmes, with confidence values of 75 % and 54 %, respectively (see Table 3.5).

ClustalW alignment of *Lmj*LIPA with human and bacterial homologues indicates that *Lmj*LIPA possesses two motifs that are integral to its function: $^{133}Cx_4Cx_5C^{144}$ and $^{165}Cx_3Cx_2C^{172}$ (see Figure 3.3 and Section 1.4.3.1). The Cx_3Cx_2C motif is present in all radical S-adenosylmethionine (SAM) proteins. The cysteine residues nucleate a [4Fe – 4S] $^{+}$ cluster, which binds SAM. LIPA catalyses the reductive cleavage of SAM to produce [4Fe – 4S] $^{2+}$-Met and the 5'-deoxyadenosyl radical (5'-dA$^{\bullet}$). 5'-dA$^{\bullet}$ abstracts hydrogen atoms linked to C6 and C8, of the octanoyl substrate (Douglas *et al.*, 2006). The motif Cx_4Cx_5C is specific to LIPA enzymes. After hydrogen abstraction by 5'-dA$^{\bullet}$, C6- and C8 alkyl radicals are formed, which attack μ-sulphido atoms attached to the Cx_4Cx_5C cluster; the final result is production of a lipoyl-group (Cicchillo *et al.*, 2004b).

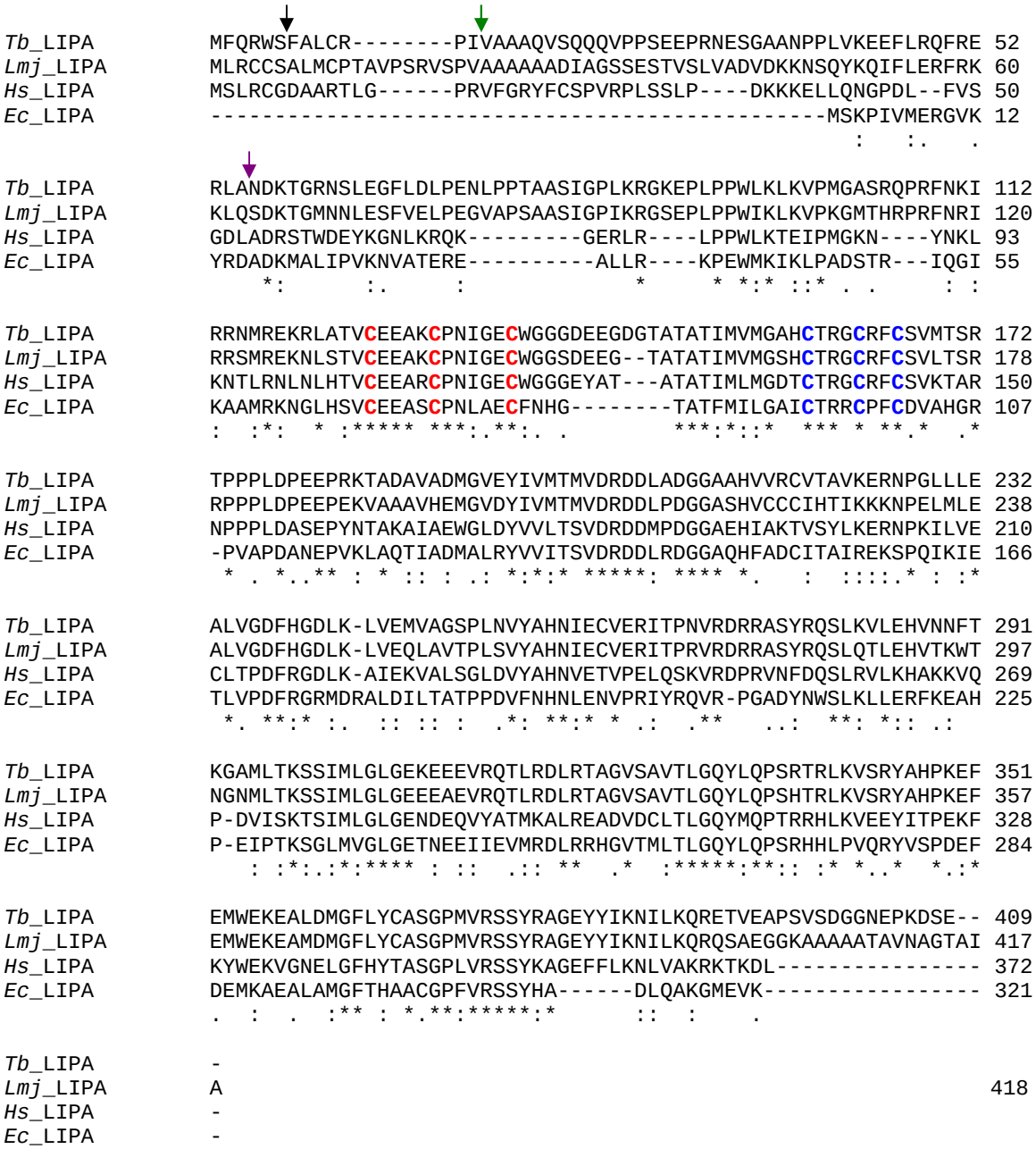


Figure 3.3 LIPA protein alignment
ClustalW alignment of *L. major* LIPA *(Lmj*LIPA) (accession number *LmjF*19.0350) with homologues in *T. brucei* (*Tb*LIPA) (accession number *Tb*10.61.1530), *H. sapiens* (*Hs*LIPA) (accession number NP_006850) and *E. coli* (*Ec*LIPA) (accession number AAC73729) was carried out. The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues highlight strictly conserved Cys residues that belong to the **C**x_4**C**x_5**C** motif that is specific to LIPA enzymes (Cicchillo *et al.*, 2004b). Blue residues represent strictly conserved Cys residues belonging to the **C**x_3**C**x_2**C** motif, which is specific to radical SAM proteins (Cicchillo *et al.*, 2004b). Bold letters represent strictly conserved residues and 'x' represents any amino acid. TargetP and MitoProt programmes predict the targeting peptide for mitochondrial transit as being five amino acids (black arrow, $LIPA_{T1}$) and 15 amino acids (green arrow, $LIPA_{T2}$), respectively. However, in order to produce soluble recombinant protein, an *N*-terminal 62 amino acid truncated version of LIPA was constructed (purple arrow, $LIPA_{T3}$) (see Section 3.5.1).

	Predicted Sizes (kDa)			Systematic Name (GeneDB)	Targeting Predictions		Identity (%) to other species	
Enzyme	***E. coli***	***H. sapiens***	***L. major***	***L. major***	**Mitoprot**	**Target P**	***H. sapiens***	***E. coli***
LIPA	36.1 (AAC73729)	41.9 (NP_006850)	45.9	LmjF19.0350	75 % MT	54 % MT	40	33
LIPB	23.9 (AAC73731)	25.2 (A6NK58)	29.5	LmjF36.3080	60 % MT	82 % MT	21	21
LPLA	37.9 (AAC77339)	42.5 (LT) (NP_057013)	55.2	LmjF07.1060	63 % MT	76 % MT	26 (LT)*	29*

Table 3.5 Sequence analyses of lipoic acid synthase (LIPA) and lipoyl/octanoyl transferase (LIPB and LPLA) proteins
L. major homologues of LIPA, LIPB and LPLA were identified as described in Section 2.3. In the 'Protein Size' column, accession numbers of *E. coli* and *H. sapiens* proteins are shown in brackets. Mitochondrial targeting predictions were carried out by analysing *L. major* protein sequences with MitoProt or TargetP. The results obtained indicate the percent likelihood of mitochondrial targeting (MT). Amino acid sequence identity of *L. major* LIPA, LIPB and LPLA proteins to *E. coli* and *H. sapiens* homologues was determined using Vector NTI. In the case of LPLA sequence identities, only the *N*-termini of *L. major* LPLA (1-222 amino acids), *E. coli* LPLA (1-182) and *H. sapiens* LT (1-250) were aligned(*).

3.4.3 LPLA

As described in Section 1.4.3.2, salvage of LA in most bacteria, plants and protozoans requires one enzyme, LPLA, whereas mammals necessitate the sequential action of two enzymes, ACSM1 and LT (see Table 3.5). Crystal structure data indicate that *E. coli* LPLA (*Ec*LPLA) (Fujiwara *et al.*, 2005), *T. acidophilum* LPLA (*Ta*LPLA) (Kim do *et al.*, 2005; McManus *et al.*, 2005) and *B. taurus* LT (*Bt*LT) (Fujiwara *et al.*, 2007) have relatively high sequence identities in their *N*-terminal domains, especially within the hydrophobic lipoyl-AMP binding site. However, unlike *Ec*LPLA, *Ta*LPLA and *Bt*LT cannot activate LA with ATP to form lipoyl-AMP. *Ta*LPLA completely lacks a *C*-terminal domain, and it is proposed that another enzyme may be required to catalyse the activation of LA to lipoyl-AMP (McManus *et al.*, 2005). In contrast, *Bt*LT possesses a *C*-terminal domain with similar folds to that of *Ec*LPLA, however the domain is shifted 180° relative to the *C*-terminal domain of *Ec*LPLA (Fujiwara *et al.*, 2007). The authors hypothesise that the more ordered conformation of the *Bt*LT *C*-terminal domain results in stabilisation of the adenylate-binding loop, which consequently increases the binding affinity for LA, yet excludes the bulkier ATP (Fujiwara *et al.*, 2007). The steric exclusion of ATP is the current hypothesis as to why LT cannot activate LA, and yet the less bulky lipoyl-AMP can enter the active site and be ligated to the apoprotein.

Searching the *L. major* genome resulted in the discovery of three potential lipoate/biotin protein ligases, *LmjF*36.3080, *LmjF*07.1060 and *LmjF*31.1070. As mentioned in Section 3.4.1, *LmjF*36.3080 most likely corresponds to the *LmjLIPB* gene, *LmjF*31.1070 corresponds to the *LmjBPL* gene, and *LmjF*07.1060 could encode *Lmj*LPLA or *Lmj*LT. The predicted *LmjF*07.1060 protein is 55.2 kDa, and as such is substantially larger than *Ec*LPLA (37.9 kDa) and *H. sapiens* LT (*Hs*LT) (42.5 kDa) (see Table 3.5). Both *LmjF*07.1060 and *Hs*LT have approximately 37 amino acid *N*-terminal extensions compared to *Ec*LPLA, which contain residues involved in mitochondrial targeting (see Figure 3.4 and Table 3.5). However, *LmjF*07.1060 has a *C*-terminus that is 137 amino acids longer than that of *Ec*LPLA, and 127 amino acids longer than that of *Hs*LT (see Figure 3.5). The *LmjF*07.1060 protein is predicted to be mitochondrial by MitoProt and TargetP (with confidence values of 63 % and 76 %, respectively) (see Table 3.5). In order to analyse sequence identities, only the *N*-terminal domains of LPLA proteins were

aligned (see Figure 3.4), since inter-species sequence variation within LPLA *C*-terminal domains is common. Firstly, *Ec*LPLA and *LmjF*07.1060 share 25 % and 26 % sequence identity, respectively, with *Hs*LT (see Table 3.5). *LmjF*07.1060 has 29 % sequence identity to *E. coli* LPLA (see Table 3.5).

ClustalW alignment of the *N*-termini of putative trypanosomatid LPLA proteins, *Ec*LPLA, *Bt*LT and *Hs*LT were carried out in order to determine whether *LmjF*07.1060 is most likely to be LPLA or LT. The key motif ^{108}RRx$_2$GGGxV(F/Y)HD^{119} that forms the lipoyl-AMP binding pocket in all LPLA and LT proteins (Kim do *et al.*, 2005; McManus *et al.*, 2005) is conserved in *LmjF*07.1060 and all other trypanosomatid homologues (see Figure 3.4). Additionally, the invariant (Fujiwara *et al.*, 2007; Reche, 2000) Lys^{172} is present in *LmjF*07.1060 and all other trypanosomatid homologues (see Figure 3.4). Lastly, in *Ec*LPLA, the motif ^{177}GITSVR^{182} which has been shown to bind to LA is instead ^{217}SVGSVR^{222} in *LmjF*07.1060, and is similar in other trypanosomatid putative LPLA proteins. In *Hs*LT and *Bt*LT the motifs are instead ^{204}ATASIPS^{210} and ^{204}ATASTPA^{210}, respectively. Although the trypanosomatid motifs share more sequence identity and similarity to the *Ec*LPLA motif than to the *Hs*LT and *Bt*LT motif, it is not possible to categorically label *LmjF*07.1060 as LPLA instead of LT.

It is apparent that alignment of the *N*-termini does not result in an obvious conclusion as to whether *LmjF*07.1060 is LPLA or LT, since lipoyl-AMP binding and transferase functions that are shared by LPLA and LT appears to be a highly conserved property of their *N*-termini (Fujiwara *et al.*, 2007; Reche, 2000). The recent analysis of the structure of *Bt*LT infers that although the *C*-terminal domains of *Ec*LPLA and *Bt*LT have the same overall fold, they are rotated approximately 180° relative to one another, which is thought to permit LPLA but not LT, to accommodate ATP and thus activate LA (Fujiwara *et al.*, 2007). Unfortunately, since trypanosomatid *C*-terminal domains are 120 – 140 residues longer than the *C*-terminal domains of *Ec*LPLA and *Hs*LT and *Bt*LT proteins (see Figure 3.5), modelling the *Lmj*LPLA *C*-terminal domain was not an option. BLAST analyses indicate that the *C*-terminal extensions of trypanosomatid putative LPLA proteins have no known homologues, although they share approximately 22 % to one another (see Figure 3.5).

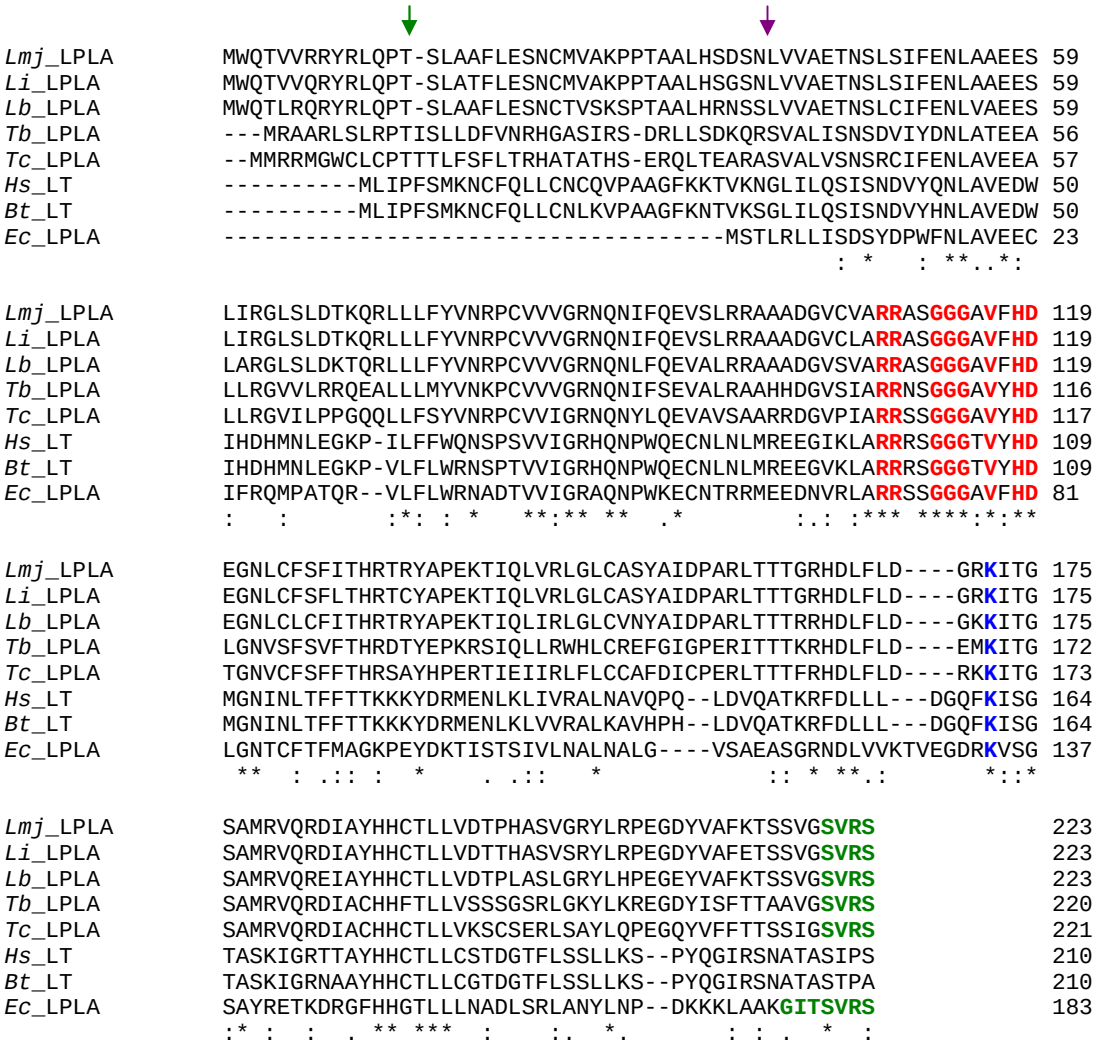


Figure 3.4 Alignment of LPLA and LT *N*-terminal domains
ClustalW alignment of the *N*-terminal domains of *L. major* LPLA (*Lmj*LPLA) (accession number *LmjF*07.1060) was carried out with homologues in *L. infantum* (*Li*LPLA) (accession *LinJ*07_V3.1230), *L. braziliensis* (*Lb*LPLA) (accession *Lbr*M07_V2.1130), *T. brucei* (*Tb*LPLA) (accession *Tb*927.8.630), *T. cruzi* (*Tc*LPLA) (accession unknown) and *E. coli* (*Ec*LPLA) (accession number AAC77339), and with *H. sapiens* LT (*Hs*LT) (accession number NP_057013) and *B. taurus* LT (*Bt*LT) (accession number NP_77720) (see Table 3.5 also). Only the *N*-terminal domains were aligned, since trypanosomatid LPLA *C*-termini are considerably longer than LPLA and LT in other organisms. The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues represent the most highly conserved amino acids within the **RR**x$_2$**GGG**x**V**(F/Y)**HD** motif of a disordered loop, which forms the hydrophobic pocket in which the lipoyl-domain of lipoyl-AMP is buried (Kim do *et al.*, 2005). In *E. coli* LPLA, LA binds to the motif ^{177}**GITSVR**^{182} (green letters), within a second disordered loop (McManus *et al.*, 2005). A lysine residue strictly conserved in all LPLA, BPL and LT enzymes is highlighted in blue (Kim do *et al.*, 2005). Bold letters represent strictly conserved residues, 'x' represents any amino acid, and 'h' represents any hydrophobic amino acid. The targeting peptide for mitochondrial transit is predicted by programme TargetP as being 12 amino acids (green arrow, LPLA$_{T1}$). However, in order to produce soluble recombinant protein, an *N*-terminal 36 amino acid truncated version of LPLA was constructed (purple arrow, LPLA$_{T2}$) (see section 3.5.3).

```
Lmj_LPLA        SPVTSLAESGCIADGP---------GAVAALKGNMADFFLAEGDRVLNAATPWELDAVEL 51
Li_LPLA         SPVTSLAESGCIAGGP---------GAVAALKRNMADFFLAEGDRVLDAATPWELDAVKL 51
Lb_LPLA         SPVTTLAESVHIASGQ---------GAMASLKRNMAEFFLTEGDRVLEAAAPWELDVREL 51
Tb_LPLA         SRTTTLKEAGVLTDSASADP-------VNFILRSMAGFFVEHSWDILAHKAPWETNPLNF 53
Tc_LPLA         SSVTTLQMAGVLPEGYCSDENINGEDVVHCTQRSLADFFVRHVGVITAHSAPWEIDPSIF 60
                * .*:*  :  :. .            :     .:* **: .   :    :*** :   :

Lmj_LPLA        R-------QRFTTARHRCANAPLFALDVVGAVAADMPFIEREGRQPARGDAATLGEAVHK 104
Li_LPLA         R-------QSFATSRHRCENAPLFALDVVGAVAADMPFFEGEGRRPTRGDAATLGEAVRK 104
Lb_LPLA         R-------QSFATARHCCADTPLFSLDVVGAVAADMSFIEGEGRRAASGDLATLGEAVHK 104
Tb_LPLA         G--DNVDTVESKRTRECRETAAVFTLDTTKAIADGTVMIDGENRRPCAGASKTVREECVR 111
Tc_LPLA         LSKRNEETKGSEEKTLHPSNESVFLLDVVGALRENTRIVDGEGRRPAAGSSKTVLEEVDR 120
                                    .:* **.. *:  .  :.: *.*:.  *   *: *   :

Lmj_LPLA        AASKEWAYAMPTFTSTVFLSDSELRRRLKTLSVRQHVALLCSLAEEELLAVLQRCVLEGS 164
Li_LPLA         TASKDWAYAMPSFTSTVLLSGSELRRRLKTLSMWQDVASLCSLAEEELLAVLQRCVLEGS 164
Lb_LPLA         AASKDWAYAMPAFTSTVLLSSGELQRRLQALSVWPDVVRLSSLAEEQLLAALQQCVFADL 164
Tb_LPLA         LQSLGWLFNMPKFETRVAITLADILASEETFSKHAALP---PSLIAWLLQSCTRVDIITL 168
Tc_LPLA         LRSADWMFSMPKFTTCVAVTAKDLLAWENDTIVRSCVP---PGVVTYLMQGRTRFDITTV 177
                  *  * : ** * : * ::  ::    :       :    .     *:    :  :

Lmj_LPLA        EGPEAVAG-TEAGLHLVTTVEHRLVTSVVVRRVAPTTATASAVAPSSASSGGWVQRYLSA 223
Li_LPLA         EGLEEVAAGTEAGLHLITTVEHRLVTSVVVRRAASTTSTASAVAPSSGSSGGWVQRCLSA 224
Lb_LPLA         VGCGEVVAETEVGLHLLTTVEHRLVTSVTVRRAPSTDAATRAVAPLGGYSGGWVERYLSA 224
Tb_LPLA         IENRRVTSITAYWIKQGEPAPEQPWFSCLLRYLVEGRYVDNVFADMGG------KCSVLA 222
Tc_LPLA         VECRHITKLTACWAENAAGDVDEMWCAAILRVLLEGVRVDDPTVDVSE------ENAVLV 231
                     :.  *    .      ..   :  :*       .    .  .       :  : .

Lmj_LPLA        LLVGHPCDAAIEGLECVGDTTAVVQGLVHEAGSLCPDLPDAVG--DAALLMVARVLLDVW 281
Li_LPLA         LLVGQPCDAAVEGLERTGDTTAVMQGLVHEAGSLCLDLPDLAG--DAALLMVTRVLLDVW 282
Lb_LPLA         LLVGHACDAAVEGLESLGDITAVVQGLLHETGSLCPELPDVAR--DAALLMVARVLLNVW 282
Tb_LPLA         AAIDLECSQIISGLPAIISQLSAGGSRNGEAVGSVISLAEERERSHDVIQRFLQAVLDVW 282
Tc_LPLA         AALQLECRSLMDGMPLSAREENIACDENGSG-----SLDEQ-----SALLLFMRAVLSVW 281
                  :   *   :.*:           .   .      .* :       .:  . :.:*.**

Lmj_LPLA        RDKNVFDIAT--                                                291
Li_LPLA         RDKNVFDIAT--                                                292
Lb_LPLA         RDKNVFDLSR--                                                292
Tb_LPLA         RRKNVFYPISY-                                                293
Tc_LPLA         RVKNVFIPVISQ                                                293
                * ****
```

Figure 3.5 Alignment of trypanosomatid LPLA *C*-terminal domains
ClustalW alignment of the *N*-terminal domains of *L. major* LPLA (*Lmj*LPLA) with homologues in *L. infantum* (*Li*LPLA) (accession *LinJ*07_V3.1230), *L. braziliensis* (*Lb*LPLA) (accession *Lbr*M07_V2.1130), *T. brucei* (*Tb*LPLA) (accession *Tb*927.8.630) and *T. cruzi* (*Tc*LPLA) (accession unknown) was carried out. Only the *C*-termini of trypanosomatid LPLA proteins were aligned as they are circa 120-140 residues longer than the *C*-termini of *Ec*LPLA and *Bt*LT. The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.).

3.5 Expression and functionality of lipoylating proteins

In order to determine whether the lipoylating proteins identified *in silico* were functional, and in order to produce polyclonal antibodies against these proteins, it was necessary to clone the genes for *LIPA*, *LIPB* and *LPLA* into a protein expression plasmid – pASK-IBA3 – adding a Strep-tag to the 3' end.

Expression of eukaryotic proteins in *E. coli* lines can be difficult, especially if the genes in question have hydrophobic targeting sequences. All three *L. major*

lipoylating proteins were predicted to be mitochondrial by MitoProt and TargetP (see Table 3.5), and the programs predicted the number of amino acids comprising the potential targeting peptide. Based on these predictions, constructs were made corresponding to full-length genes or truncated versions (see arrows in Figure 3.2, Figure 3.3 and Figure 3.4). All proteins were successfully expressed, purified and polyclonal antibodies generated (see Section 3.5.4).

3.5.1 LIPA

3.5.1.1 Cloning and expression of LIPA

DNA of full-length (*FL*) *LIPA* and three 5' truncated versions (*T1*, *T2* and *T3*) (see Figure 3.3) were cloned into the pASK-IBA3 plasmid, resulting in constructs *LIPA*$_{FL}$IBA3, *LIPA*$_{T1}$IBA3, *LIPA*$_{T2}$IBA3 and *LIPA*$_{T3}$IBA3, as described in Section 2.5.6.1 (see Figure 3.6A for *LIPA*$_{T3}$IBA3 cloning). Constructs encoding LIPA$_{T1}$ and LIPA$_{T2}$ result in LIPA proteins that are truncated at the *N*-terminus by 5 amino acids and 15 amino acids (see Figure 3.3); the respective constructs were designed as such based upon the mitochondrial targeting peptide cleavage sites predicted by TargetP and MitoProt, respectively. The *LIPA*$_{T3}$ construct was designed such that the LIPA$_{T3}$ protein would be truncated by 63 amino acids at the *N*-terminus (see Figure 3.3). The rationale was that LIPA$_{T3}$ would lack all of the associated hydrophobic targeting sequence and as such possibly be expressed to a higher degree in the soluble fraction. Relative to the predicted start of the *Ec*LIPA protein, LIPA$_{T3}$ lacks 14 amino acids (see Figure 3.3). However, the aim was to purify soluble protein and not to characterise the *Lmj*LIPA enzyme, and as such this truncation was deemed acceptable.

Trial expressions identified that the optimal conditions for protein expression were using the *LIPA*$_{T3}$IBA3 construct in BLR (DE3) *E. coli*, shaking at 37 °C 250 rpm for one hour, post induction. *LIPA*$_{FL}$IBA3, *LIPA*$_{T1}$IBA3 and *LIPA*$_{T2}$IBA3 constructs expressed in the same *E. coli* line did not result in the production of sufficient quantities of soluble protein for protein purification at any of the conditions tested (see Section 2.5.6.2) (data not included).

Purification of LIPA$_{T3}$-Strep protein via affinity chromatography on a Strep-tactin column was subsequently carried out. SDS-PAGE analysis resulted in the identification of a band potentially corresponding to LIPA$_{T3}$-Strep (40.6 kDa) in

elution fractions two, three and four (see Figure 3.6B). In addition, contaminating bands were observed at approximately 60 kDa and 65 kDa, which were consistently co-purified with LIPA$_{T3}$-Strep. Western blotting with α-Strep antibody confirmed the circa 40 kDa band to be LIPA$_{T3}$-Strep, and some cross-reactivity is observed in elution fraction 3 (see Figure 3.6C).

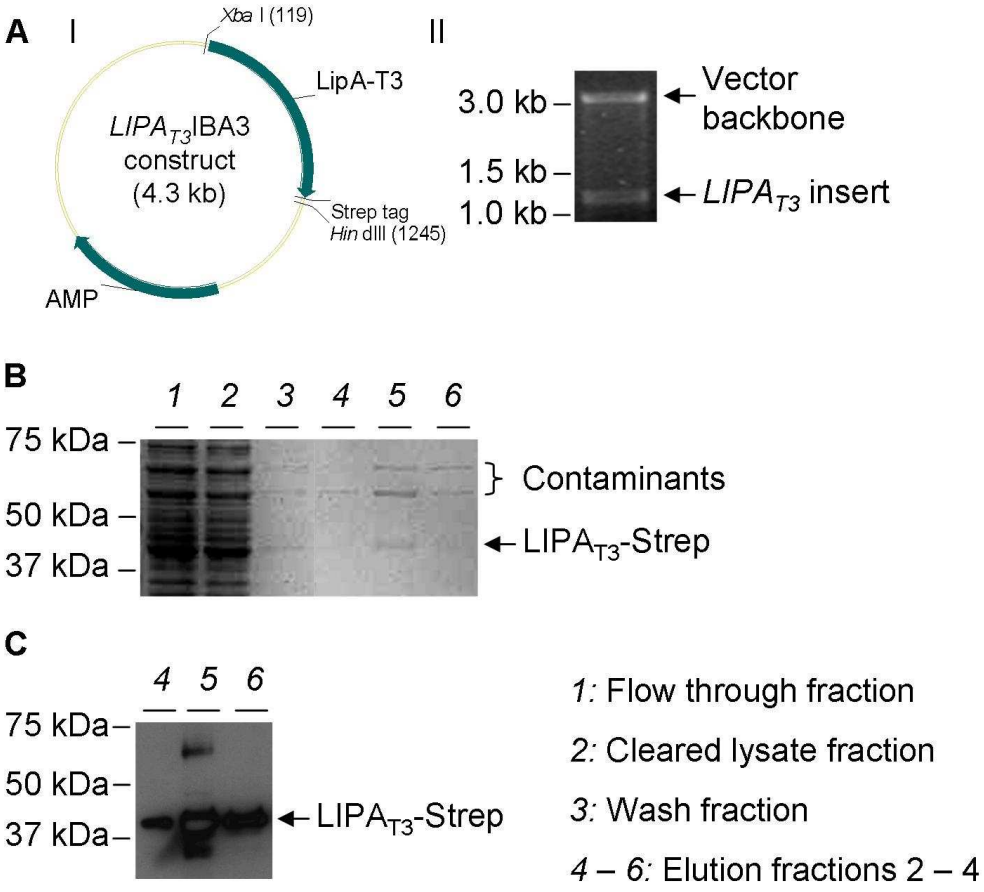


Figure 3.6 Cloning, expression and purification of LIPA
A, LIPA$_{T3}$ (see Figure 3.3) was amplified from gDNA with primer pair Lm28-29 (see Section 2.1.4). The insert was cloned into expression vector pASK-IBA3, generating construct *LIPA$_{T3}$*IBA3 (I). Correct cloning was verified by *Xba*I / *Hind*III digest, releasing a 1.1 kb insert corresponding to *LIPA$_{T3}$*, and vector backbone of 3.1 kb (II). *B, LIPA$_{T3}$*IBA3 was expressed in *E. coli* BLR (DE3) cells, and LIPA$_{T3}$-Strep protein purified via affinity chromatography on a Strep-tactin column. 10 µl of fractions collected during purification were subjected to 10 % SDS-PAGE and stained with Coomassie-blue. *C,* Western blotting with α-Strep antibody at 1/7,500 on elution fractions 2 – 4 verifies the 40.6 kDa band as purified LIPA$_{T3}$-Strep.

3.5.1.2 Functionality of LIPA

An *E. coli* line lacking the endogenous *lipA* gene (KER176) was transformed with *LIPA$_{T1}$*IBA3, *LIPA*$_{T2}$IBA3 or *LIPA*$_{T3}$IBA3 constructs, or uncut plasmid pASK-IBA3 (negative control). All were plated on minimal medium without- or with supplementation of 10 µM LA. In the presence of LA, growth was apparent in all

cases. Without LA, growth was only supported by the *LIPA*$_{T1}$IBA3 construct (see Figure 3.7).

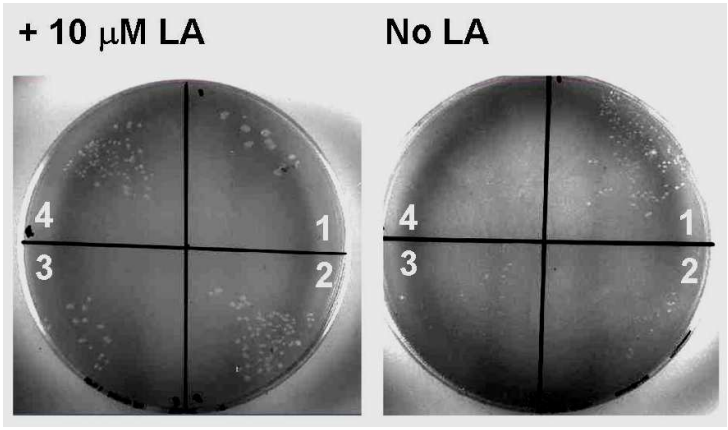


Figure 3.7 Functionality of LIPA
Functionality of *L. major* LIPA was assessed by complementation of *lipA* deficient bacteria, KER176 (Reed & Cronan, 1993). Bacteria were transformed with *LIPA*$_{T1}$IBA3 (1), *LIPA*$_{T2}$IBA3 (2), *LIPA*$_{T3}$IBA3 (3) (see Figure 3.3) or pASK-IBA3 uncut plasmid negative control (4), and plated on minimal medium without- or with supplementation of 10 μM LA. Plates were incubated at 37 ℃ for 48 – 72 h and growth recorded by photographing plates.

3.5.2 LIPB

3.5.2.1 Cloning and expression of LIPB

DNA of full-length (*FL*) *LIPB* and one 5' truncated version (*T1*) (see Figure 3.2) were cloned into pASK-IBA3 plasmid, resulting in constructs *LIPB*$_{FL}$IBA3 and *LIPB*$_{T1}$IBA3, as described in Section 2.5.6.1 (see Figure 3.8A for *LIPB*$_{FL}$IBA3 cloning). The construct encoding LIPB$_{T1}$-Strep results in LIPB protein that is truncated at the *N*-terminus by 34 amino acids (see Figure 3.2); the construct was designed as such based upon the mitochondrial targeting peptide cleavage site predicted by MitoProt. TargetP did not predict a cleavage site which explains why only one truncated construct was designed.

Trial expressions identified the optimal conditions for protein expression with the *LIPB*$_{FL}$IBA3 construct in BLR (DE3) *E. coli* cells, shaking at 37 °C 250 rpm for two hours, post induction. Expression of *LIPB*$_{T1}$IBA3 was not tested, since it was possible to express a relatively high level of soluble LIPB$_{FL}$-Strep protein.

Subsequently, LIPB$_{FL}$-Strep protein was purified via affinity chromatography on a Strep-tactin column. SDS-PAGE analysis resulted in the identification of a band potentially corresponding to LIPB$_{FL}$-Strep (30.7 kDa) in elution fractions three, four

and five (see Figure 3.8B). In addition, a contaminating band was observed at approximately 60 kDa, which was consistently co-purified with $LIPB_{FL}$-Strep. The contaminant was excised from a Coomassie-Blue-stained SDS-PAGE and shown by tandem MS / MS analysis to be the *E. coli* GroEL large subunit (accession number AAC77103) (Dr R. Burchmore, University of Glasgow). Western blotting with α-Strep antibody confirmed the 30 kDa band to be $LIPB_{FL}$-Strep (see Figure 3.8C). An additional cross-reacting band is apparent at circa 48 kDa, although the nature of the band is unknown.

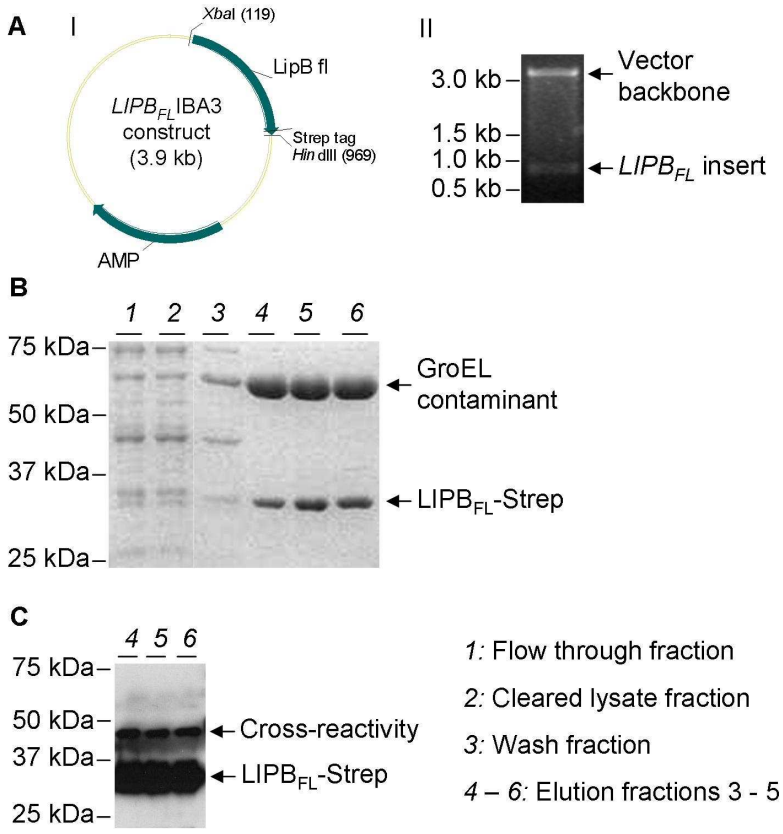


Figure 3.8 Cloning, expression and purification of LIPB
A, $LIPB_{FL}$ (see Figure 3.2) was amplified from gDNA with primer pair Lm30-33 (see Section 2.1.4). The insert was cloned into expression vector pASK-IBA3, generating construct *$LIPB_{FL}$*IBA3 (I). Correct cloning was verified by *Xba*I/*Hind*III digest, releasing a 0.9 kb insert corresponding to *$LIPB_{FL}$*, and vector backbone of 3.1 kb (II). *B, $LIPB_{FL}$*IBA3 was expressed in *E. coli* BLR (DE3) cells, and $LIPB_{FL}$-Strep protein purified via affinity chromatography on a Strep-tactin column. 10 µl of fractions collected during purification were subjected to 10 % SDS-PAGE and stained with Coomassie-blue. *C,* Western blotting with α-Strep antibody at 1/7,500 on elution fractions 3 – 5 verifies the 31 kDa band as purified $LIPB_{FL}$-Strep.

3.5.3 LPLA

3.5.3.1 Cloning and expression of LPLA

DNA of full-length (*FL*) *LPLA* and two 5' truncated versions (*T1* and *T2*) (see Figure 3.4) were cloned into pASK-IBA3 plasmid, resulting in constructs *LPLA*$_{FL}$IBA3, LPLA$_{T1}$IBA3 and LPLA$_{T2}$IBA3, as described in Section 2.5.6.1 (see Figure 3.9A for LPLA$_{T2}$IBA3 cloning). Constructs encoding LPLA$_{T1}$Strep and LPLA$_{T2}$Strep result in LPLA proteins that are truncated at the *N*-terminus by 12 amino acids and 36 amino acids (see Figure 3.4); the respective constructs were designed as such based upon the mitochondrial targeting peptide cleavage site predicted by TargetP and to align with the start of *Ec*LPLA, respectively.

Trial expressions identified the optimal conditions for protein expression with the *LPLA*$_{T2}$IBA3 construct in BLR (DE3) *E. coli* cells, shaking at 15 °C 250 rpm overnight, post induction. Under all conditions tested (see Section 2.5.6.2), *LPLA*$_{FL}$IBA3 and *LPLA*$_{T1}$IBA3 constructs did not produce soluble protein to the same degree as did the *LPLA*$_{T2}$IBA3 construct (data not shown).

Purification of LPLA$_{T2}$-Strep protein via affinity chromatography on a Strep-tactin column was subsequently carried out. SDS-PAGE analysis resulted in the identification of a band potentially corresponding to LPLA$_{T2}$-Strep (52.4 kDa) in elution fractions three and four (see Figure 3.9B). In addition, contaminating bands were observed at approximately 60 and 65 kDa, which were consistently co-purified with LPLA$_{T2}$-Strep. Western blotting with α-Strep antibody confirmed the circa 52 kDa band to be LPLA$_{T2}$-Strep (see Figure 3.9C).

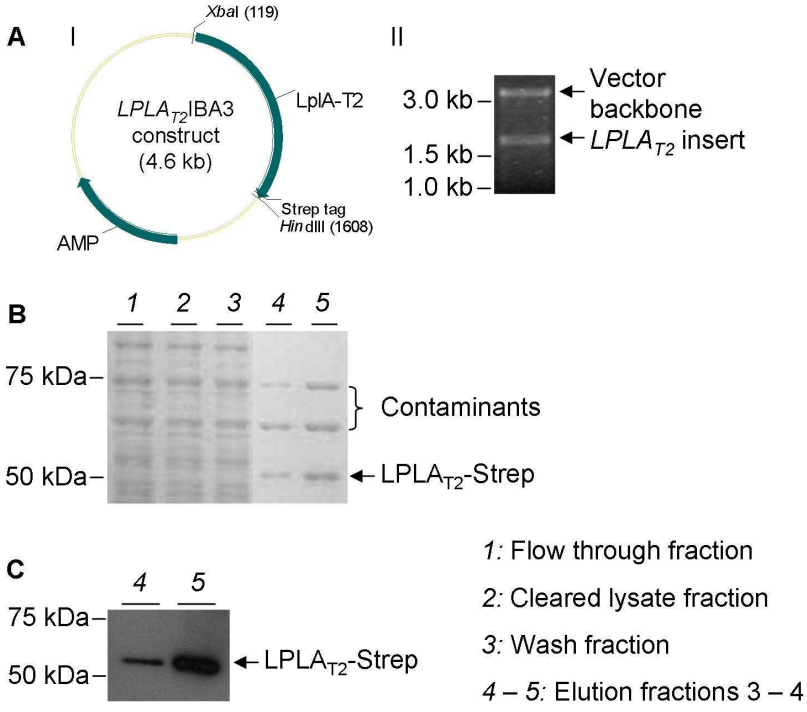


Figure 3.9 Cloning, expression and purification of LPLA
A, LPLA$_{T2}$ (see Figure 3.4) was amplified from gDNA with primer pair Lm36-37 (see Section2.1.4). The insert was cloned into expression vector pASK-IBA3, generating construct *LPLA$_{T2}$*IBA3 (I). Correct cloning was verified by *Xba*I/*Hind*III digest, releasing a 1.5 kb insert corresponding to *LPLA$_{T2}$*, and vector backbone of 3.1 kb (II). *B, LPLA$_{T2}$*IBA3 was expressed in *E. coli* BLR (DE3) cells, and LPLA$_{T2}$-Strep protein purified via affinity chromatography on a Strep-tactin column. 10 µl of fractions collected during purification were subjected to 10 % SDS-PAGE and stained with Coomassie-blue. *C,* Western blotting with α-Strep antibody at 1/7,500 on elution fractions 3 – 4 verifies the 52 kDa band as purified LPLA$_{T2}$-Strep.

3.5.3.2 Functionality of LIPB and LPLA

In order to test the functionality of *L. major* LIPB and LPLA enzymes, the *E. coli* line KER184 which lacks the endogenous *lipB* gene, was transformed with *LIPB$_{FL}$*IBA3, *LPLA$_{T1}$*IBA3, *LPLA$_{T2}$*IBA3 or *P. falciparum LPLA*1$_{FL}$IBA3 (positive control) constructs, or uncut plasmid pASK-IBA3 (negative control) (see Figure 3.10). All were plated on minimal medium without- or with supplementation of 10 µM LA.

Growth was supported in all cases except with *LPLA$_{T2}$*IBA3. Growth of *LPLA$_{T1}$*IBA3 was considerably lower than that of *P. falciparum LPLA*1$_{FL}$IBA3 positive control. *LIPB$_{FL}$*IBA3 complemented growth of KER184 in the absence of LA to a higher extent than did *LPLA$_{T1}$*IBA3. As explained in Section 3.5.3.1, under all conditions tested, expression of the *LPLA$_{T2}$*IBA3 construct produced more soluble protein than did *LPLA$_{FL}$*IBA3 or *LPLA$_{T1}$*IBA3. Although expression was not analysed in KER184 cells, a possible explanation for the unexpected results with

regards to complementation of KER184 is that $LPLA_{T2}$-Strep protein is non-functional and overexpression is toxic to *E. coli*. In contrast, $LPLA_{T1}$-Strep is likely to be at least partially functional because it supports colony growth of KER184 *E. coli* in LA-depleted medium (see Figure 3.10). Potential reasons why $LPLA_{T1}$-Strep may be functional yet $LPLA_{T2}$-Strep is toxic to KER184 cell growth will be discussed in Section 6.3.3.1.

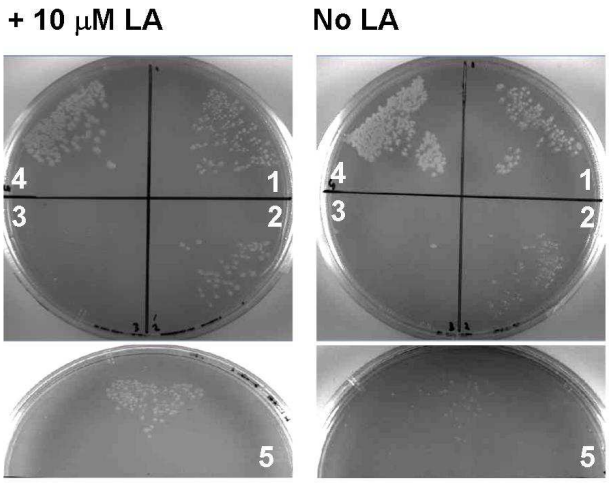


Figure 3.10 Functionality of LIPB and LPLA
Functionalities of *L. major* LIPB and LPLA were assessed by complementation of *LIPB* deficient bacteria, KER184 (Morris *et al.*, 1995). Bacteria were transformed with $LIPB_{FL}$IBA3 (1) (see Figure 3.2), $LPLA_{T1}$IBA3 (2), $LPLA_{T2}$IBA3 (3) (see Figure 3.4), *P. falciparum LPLA*1_{FL}IBA3 positive control (4) or pASK-IBA3 uncut plasmid negative control (5), and plated on minimal medium without- or with supplementation of 10 μM LA. Plates were incubated at 37 ℃ for 48 – 72 h and growth recorded by photographing plates.

3.5.4 Polyclonal antibody generation and purification

Two rabbits were inoculated with SDS-PAGE-purified $LIPA_{T3}$-Strep, $LIPB_{FL}$-Strep or $LPLA_{T2}$-Strep antigen (see Sections 2.5.6 and 2.5.7). Western blotting with the antisera resulted in extensive background signal, however it was possible to discern which antisera recognised the appropriate antigen. In all cases, the antisera also detected co-purified contaminants from the recombinant protein elution fractions. It was possible to purify specific antibodies raised against $LIPA_{T3}$-Strep and $LIPB_{FL}$-Strep antigens from the antisera. Unfortunately, the yield of α-*Lmj*LPLA antibodies was poor. Albeit multiple attempts to purify α-*Lmj*LPLA antibodies from the two antisera, it was not possible to reliably use the antibodies for western blot analysis of recombinant protein or *L. major* protein lysate.

3.5.4.1 Purification of α-*Lmj*LIPA antibodies

α-*Lmj*LIPA antibodies were purified from the antiserum via a two-step procedure. Firstly, α-*Lmj*LIPA antibodies were separated from antibodies recognising contaminating proteins in the purified $LIPA_{T3}$-Strep elution fractions, by incubating 5 ml of α-*Lmj*LIPA antiserum on agarose beads cross-linked to $LPLA_{T2}$-Strep recombinant protein. Given that the contaminants in $LPLA_{T2}$-Strep purification were the same size (and thus probably the same proteins) as those in $LIPA_{T3}$-Strep purification, the rationale was that contaminating antibodies in α-LIPA antiserum would bind to the column, whereas α-*Lmj*LIPA-specific antibodies would be collected in the flow-through.

Subsequently, α-*Lmj*LIPA antibodies were purified using a $LIPA_{T3}$-Strep recombinant protein column. Proteins in the antiserum that resulted in high background in western blots would be lost in the flow-through fraction, and then α-*Lmj*LIPA-specific antibodies could be eluted from the column (see Figure 3.11).

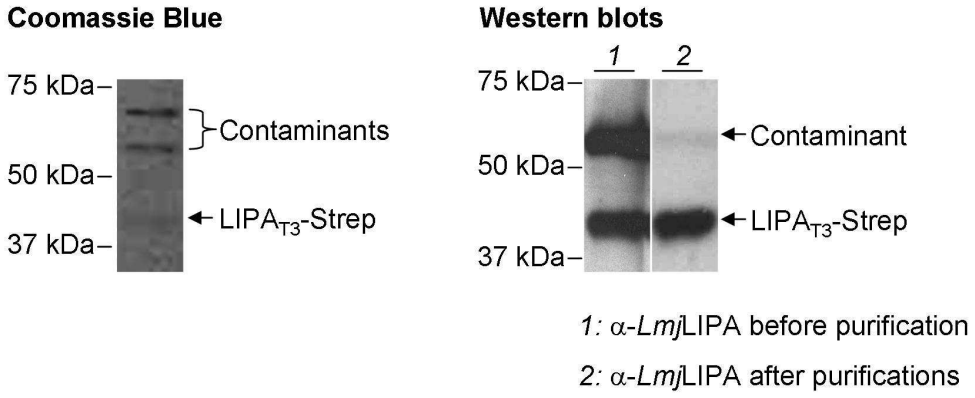


Figure 3.11 Purification of α-*Lmj*LIPA antibodies from α-*Lmj*LIPA antiserum
In order to purify α-*Lmj*LIPA antibodies from α-*Lmj*LIPA antiserum, two affinity chromatography steps were carried out. Purified recombinant $LIPA_{T3}$-Strep was subjected to SDS-PAGE and then stained with Coomassie Blue. The overall yield of the elution fraction shown is 0.2 μg μl^{-1}, and 10 μl was loaded on the gel. $LIPA_{T3}$-Strep band is barely visible, and as such it was not possible to determine the quantity loaded. The same quantity of material was used for western blot analyses. Western blotting with α-*Lmj*LIPA antiserum before purification was carried out at 1/20,000 (*lane 1*). After two rounds of purification, western blotting was carried using α-*Lmj*LIPA antibody at 1/500 (*lane 2*).

3.5.4.2 Purification of α-*Lmj*LIPB antibodies

Two rabbits were inoculated with the SDS-PAGE-purified antigen (see Sections 2.5.6 and 2.5.7). One of the two antisera was found to recognise $LIPB_{FL}$-Strep recombinant protein by western blotting, although the background signal was high.

α-*Lmj*LIPB antibodies were purified from the antiserum using a $LIPB_{FL}$-Strep recombinant protein column. Compounds in the antiserum that resulted in high background in western blot were lost in the flow-through fraction, and then α-*Lmj*LIPB-specific antibodies were eluted from the column (see Figure 3.12).

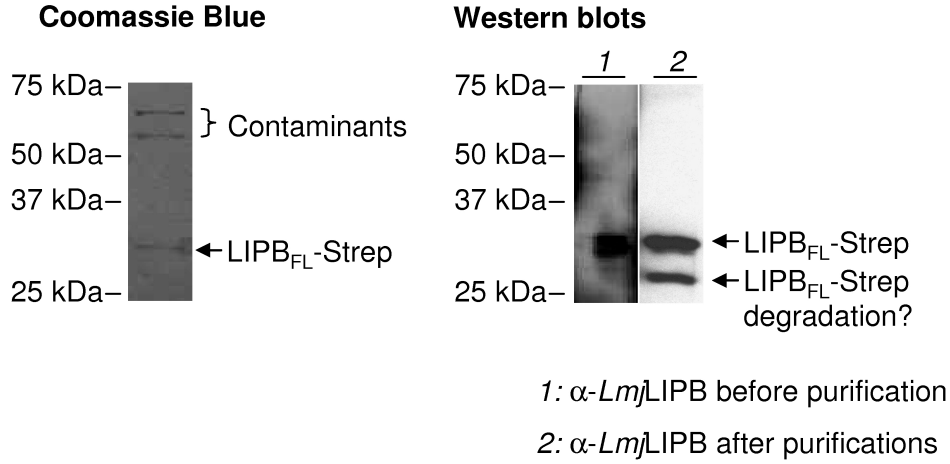


Figure 3.12 Purification of α-LmLIPB antibodies from α-LmLIPB antiserum
In order to purify α-*Lmj*LIPB antibodies from α-*Lmj*LIPB antiserum, one affinity chromatography step was carried out. Purified recombinant $LIPB_{FL}$-Strep was subjected to SDS-PAGE and then stained with Coomassie Blue. The overall yield of the elution fraction shown is 0.2 μg μl^{-1}, and 10 μl was loaded on the gel, and since $LIPB_{FL}$-Strep corresponds to approximately 1/3 of the total protein, 666 ng of $LIPB_{FL}$-Strep protein is present in 10 μl of the elution fraction. The same quantity of material was used for western blot analyses. Western blotting with α-*Lmj*LIPB antiserum before purification was carried out at 1/50,000 (*lane 1*). After one round of purification, western blotting was carried using α-*Lmj*LIPB antibody at 1/1,000 (*lane 2*).

3.6 Effects of LA analogues on promastigotes

Having established that *L. major* possesses active lipoylating enzymes, and that α-KADH and GCC components are lipoylated throughout promastigote development, I next determined the effect of LA analogues on promastigote cell growth and lipoylation patterns. The analogues octanoic acid (OA) and 8'-bromooctanoic acid (8-BOA) were used, since data has already been published in other parasites – *T. gondii* and *P. falciparum* – with regards to their effects on growth and lipoylation patterns (Allary *et al.*, 2007; Crawford *et al.*, 2006). In these experiments, parasites

were grown in lipid-depleted medium in order to minimise competition of exogenous LA with LA analogues.

The hypothesis was that a lower concentration of 8-BOA than OA would be required to detrimentally affect parasite growth, and that addition of just 1 μM LA would rescue the phenotype, as is the case in *T. gondii* and *P. falciparum* (Allary *et al.*, 2007; Crawford *et al.*, 2006). The rationale behind this result is that octanoylation of apoproteins is a natural cellular process ordinarily catalysed by LIPB, and the octanoyl-moiety can be converted to an active lipoyl-group by the action of LIPA. However, 8-BOA is an analogue of LA which when ligated to apoproteins, renders them inactive. The toxicity of a similar LA analogue, selenolipoic acid (SeLA), has been demonstrated in *E. coli* (Fujiwara *et al.*, 1997b; Jordan & Cronan, 2002; Morris *et al.*, 1994; Reed *et al.*, 1994).

In order to remain consistent with previously published data, *L. major* promastigotes were grown in lipid-depleted medium for two sub-passages (equating to 1/100 dilution of lipids from the original HOMEM + 10 % FCS medium) before adding LA analogues alone or with 1 μM LA supplementation. Alamar blue assays were carried out in order to determine the effects of LA analogues on *L. major* promastigote growth.

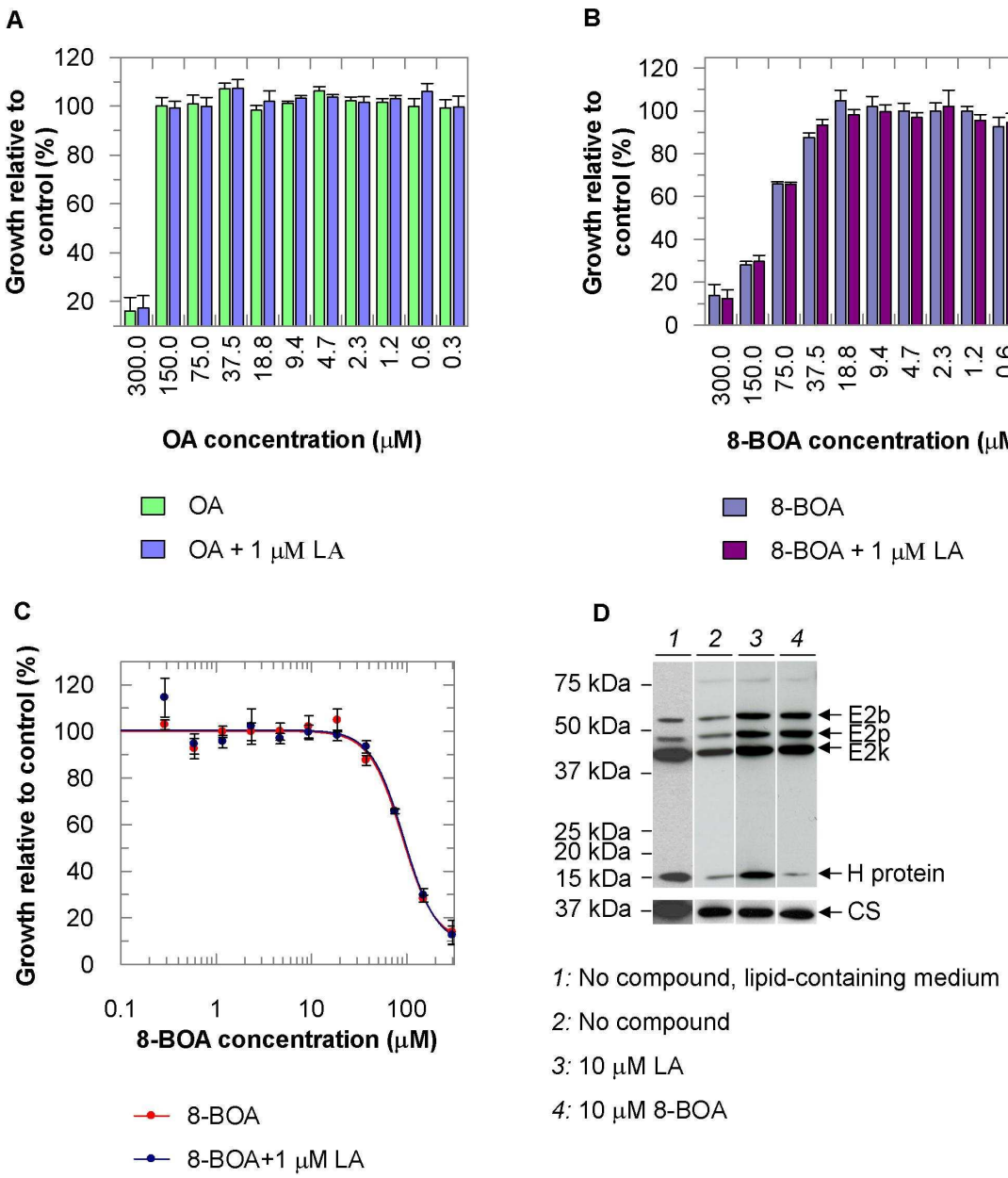


Figure 3.13 Growth and lipoylation patterns of promastigotes incubated with different concentrations of LA analogues
Alamar blue assay was used in order to determine the effect of LA analogues, OA (*A*) and 8-BOA (*B* and *C*), on promastigote cell growth. Fluorescence at OD_{590nm} was recorded and all values normalised against negative EtOH (solvent) controls to give percentage growth relative to the control. *D,* In order to determine the effect of LA analogues on lipoylation patterns, promastigotes were grown in lipid-depleted medium, and supplemented with either 10 μM LA or 10 μM 8-BOA. Stationary phase promastigotes (120 h) were lysed and the soluble protein fraction isolated. 10 μg of each lysate was subjected to SDS-PAGE. Western blotting was carried out using α-LA antibody at 1/6,000 and α-CS antibody to assess loading, at 1/5,000. *Lane 1* represents a typical lipoylation pattern at 120 h growth in normal HOMEM + 10 % FCS medium (see Figure 3.1B *lane 4*, for derivation of *lane 1* in the current figure).

Figure 3.13A demonstrates that OA concentrations up to 150 μM do not affect promastigote growth. 300 μM OA reduces growth to 16 % relative to the negative control. In *T. gondii*, the IC_{50} for OA is approximately 500 μM, yet 1 μM LA completely restored growth to wild-type levels (Crawford *et al.*, 2006). In *L. major*, 1 μM LA does not alleviate the effects of 300 μM OA (Figure 3.13A). Addition of

LA or 8-BOA to the growth medium did not cause an alteration in lipoylation intensity/patterns (Figure 3.13D), and as such it is difficult to discern whether these compounds are not efficiently taken up, or whether parasites rely on biosynthesis of LA and as such do not incorporate LA or 8-BOA into α-KADHs or the GCC.

Figure 3.13B & C illustrate that the IC_{50} for 8-BOA is 90.5 μM (± 9.2 μM) and the IC_{50} for 8-BOA with 1 μM LA supplementation is 92.3 μM (± 10.8 μM). These values are relatively high, compared to those published for *T. gondii*, whereby 10 μM 8-BOA inhibits cell growth by 95 %, and 1 μM LA is sufficient to completely rescue the growth defect (Crawford *et al.*, 2006). In *P. falciparum* red blood cell stage, 8-BOA at 25 μM – 400 μM results in growth defects, and addition of 2 μM LA significantly rescued the phenotype (Allary *et al.*, 2007). Possible explanations for the discrepancy between the results observed in *T. gondii* and *P. falciparum* versus *L. major* will be discussed in Section 6.3.3.2.

3.7 Summary

- *In silico* searches show that *L. major* possesses all potential components of α-KADHs and the GCC, although some potentially interesting differences are apparent.

- Lipoylation of α-KADH E2-subunits and H-protein occurs during promastigote growth, and lipoylation patterns are consistent throughout: E2k is predominantly lipoylated, and during stationary phase E2p, E2b and H-protein become lipoylated. Purified metacyclic promastigotes only lipoylate E2k. Amastigotes may possess lipoylated E2p, E2k, H-protein and not E2b.

- *L. major* has genes corresponding to *LIPA* and *LIPB* from the LA biosynthesis pathway, and both contain motifs that are pivotal to activity of these enzymes.

- *L. major* has a gene for *LPLA* from the salvage pathway. The *Lmj*LPLA protein possesses all of the key motifs within the *N*-terminus that are required for transferase activity; however these motifs are also conserved in the mammalian LT enzyme. *Lmj*LPLA has a circa 140 amino acid *C*-terminal extension relative to the *C*-terminal domains of *Ec*LPLA and mammalian LT enzymes, the nature of which is unclear.

- *L. major LIPA*, *LIPB* and *LPLA* were cloned into expression construct pASK-IBA3, and the proteins were expressed and purified.

- Antibodies were generated for all three proteins, and α-*Lmj*LIPA and α-*Lmj*LIPB antibodies were successfully purified for use in western blot analyses.

- The functionality of lipoylating genes from *L. major* was tested by functional complementation of *E. coli* lacking *lipA* (KER176) or *lipB* (KER184) genes. *Lmj*LIPA, *Lmj*LIPB and *Lmj*LPLA rescued bacterial growth. *Lmj*LPLA did not complement KER184 bacteria as efficiently as did the *Pf*LPLA1 positive control, perhaps indicating that *Lmj*LPLA does not utilise octanoyl-ACP as efficiently as does *Pf*LPLA1, and it could also be due to the extended *C*-terminal extension not found in *E. coli*.

- In lipid-depleted medium, analogues of LA – OA and 8-BOA – affected *L. major* promastigote cell growth at higher concentrations than those published for *T. gondii* and *P. falciparum*, and supplementation with 1 μM LA did not rescue the growth phenotype. Addition of 10 μM LA or 8-BOA did not cause an alteration in lipoylation patterns.

4 Genetic manipulation of *LPLA*

4.1 Introduction

L. major encodes a putative LPLA enzyme, which was shown to be functional by complementation of *E. coli* deficient in *lipB*, as described in Section 3.4.3. In the study reported in this chapter, the aims were to determine the subcellular localisation of LPLA and the role that LPLA plays in *L. major*. The former question was addressed by expressing a LPLA-green fluorescent protein (GFP) fusion protein in promastigotes, and establishing the subcellular localisation of LPLA-GFP by live cell fluorescence microscopy. In order to elucidate the function of LPLA in promastigotes, attempts were made to ablate both gene copies by homologous recombination. Additionally, wild-type LPLA and an active site mutant of LPLA were ectopically expressed in promastigotes, with the aim of perturbing the normal balance of lipoylation, to investigate whether this results in a decrease in parasite fitness.

4.2 Localisation of LPLA

To generate a *L. major* line expressing a *C*-terminal fusion protein of LPLA with GFP, the full-length *LPLA* gene was amplified by PCR using primer pair Lm23-24 (see Section 2.1.4). The resulting 1.5 kb fragment was digested with *Nde*I/*Bgl*II before cloning into *Nde*I/*Bgl*II-digested pGL1132 vector (Tetaud *et al.*, 2002) to yield construct *LPLA-GFP* (see Figure 4.1A). The destination vector pGL1132 contains the required control elements that allow constitutive expression of the proteins in *L. major* as well as the neomycin phosphotransferase (NEO) selectable marker (see Section 2.2.6) (Cruz & Beverley, 1990; Tetaud *et al.*, 2002). The *LPLA-GFP* construct was transfected into wild-type promastigotes (WT) to give rise to the WT[*LPLA-GFP*] line, which was grown in the presence of 50 μg ml^{-1} G418 to select for parasites that possess ectopic copies of the *LPLA-GFP* plasmid (see Section 2.2.8).

In order to verify that the LPLA-GFP fusion protein is expressed in the WT[*LPLA-GFP*] line, western blotting with α-GFP monoclonal antibody was carried out. Figure 4.1B shows that a 79.5 kDa band is found in the WT[*LPLA-GFP*] line but not in wild-type, which corresponds to the LPLA-GFP fusion protein. The size of

LPLA after processing of its target peptide can be calculated by subtraction of the known size of GFP (27.1 kDa) from that observed for the LPLA-GFP band (79.5 kDa), which equals 52.4 kDa. This is an intermediate between the $LPLA_{T1}$ (53.7 kDa) and $LPLA_{T2}$ (50.3 kDa) constructs created for expression in *E. coli* (see Sections 3.4.3 and 3.5.3). The calculated size of the *T. brucei* elongation factor 1-α (*Tb*EF1-α) loading control was 52.4 kDa, which is in accordance with previously published data that reported *Tb*EF1-α to be 53 kDa (Kaur & Ruben, 1994).

Fluorescence microscopy to detect the LPLA-GFP reporter, in combination with Mitotracker CMXRos mitochondrial stain, revealed that LPLA-GFP is a mitochondrial protein (see Figure 4.1C). This result is in accordance with subcellular targeting predictions using MitoProt and TargetP (see Section 3.4.3).

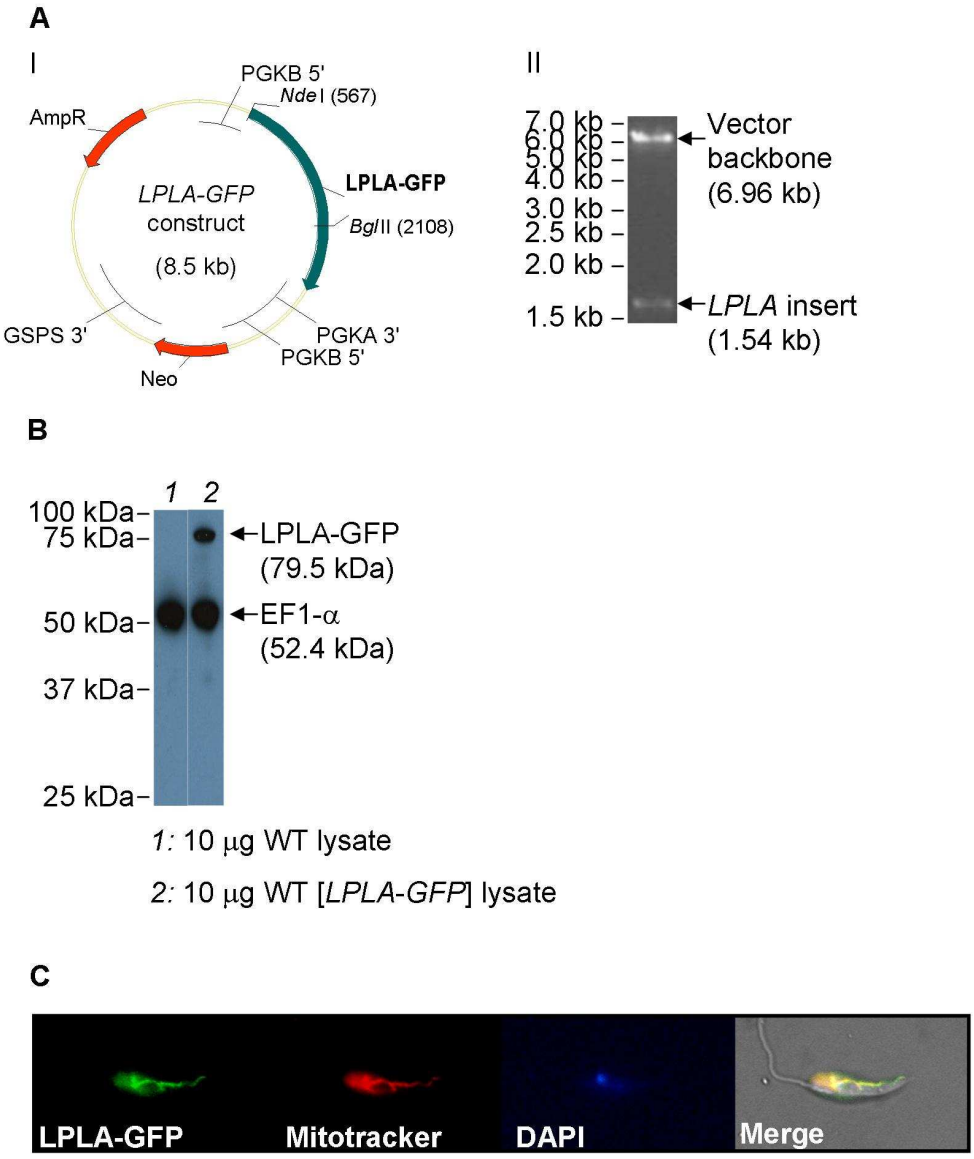


Figure 4.1 Subcellular localisation of LPLA-GFP
A, The *LPLA* gene (see Section 3.4.3) was amplified from gDNA with primer pair Lm23-24 (see Section 2.1.4). The insert was cloned into expression vector pGL1132, generating construct *LPLA-GFP* (I). Correct cloning was verified by *Nde*I/*Bgl*II digest, releasing a 1.54 kb insert corresponding to *LPLA*, and vector backbone of 6.96 kb (II). *B*, The *LPLA-GFP* construct was expressed in promastigotes to give rise to the WT[*LPLA-GFP*] line, and expression verified by western blotting of 10 μg wild-type protein lysate and 10 μg LPLA-GFP protein lysate with α-GFP antibody at 1/5,000. Equal loading was verified by probing with α-*Tb*EF1-α antibody at 1/10,000.The calculated sizes of LPLA-GFP and *Tb*EF1-α are 79.5 kDa and 52.4 kDa, respectively (see Section 2.5.5) for method of calculating molecular sizes from SDS-PAGE). *C*, The mitochondrion of WT[*LPLA-GFP*] promastigotes was stained with Mitotracker CMXRos, and images of live cells were attained by fluorescence microscopy with DIC (for phase), FITC (for LPLA-GFP), rhodamine (for Mitotracker) or DAPI filters.

4.3 Knockout studies

In order to study the *in vivo* role of LPLA, a gene replacement strategy was employed. Knockout cassettes *LPLA-SAT* and *LPLA-HYG* were generated and transfected into promastigotes, as described in Sections 2.2.4–2.2.9.

Figure 4.2 is representative of the results obtained from *LPLA* knockout attempts, a summary of which is provided in Table 4.1. Two independent heterozygous lines for *LPLA* – Δ*lplA*$_{HYG}$*/LPLA* and Δ*lplA*$_{SAT}$*/LPLA* – were created by replacement of one *LPLA* allele with either *LPLA-HYG* or *LPLA-SAT* constructs, respectively. Based on growth data generated in duplicate, Δ*lplA*$_{HYG}$*/LPLA* and Δ*lplA*$_{SAT}$*/LPLA* lines reached significantly lower cell densities at stationary phase than did the wild-type control (see Figure 4.3).

Subsequently, attempts were made to replace the remaining *LPLA* allele in Δ*lplA*$_{HYG}$*/LPLA* and Δ*lplA*$_{SAT}$*/LPLA* lines. Δ*lplA*$_{HYG}$*/LPLA* was transfected with the *LPLA-SAT* knockout cassette on three independent occasions (see Table 4.1); a total of nine clones proliferated during the first cloning step in 96-well plates. However, upon sub-passaging into fresh medium, all nine clones failed to grow, and no genotypic data were retrieved.

On the other hand, three independent transfections of Δ*lplA*$_{SAT}$*/LPLA* with the *LPLA-HYG* knockout cassette resulted in viable clones. Correct integration of the second knockout cassette was achieved for 30/63 clones analysed (see Table 4.1). However, as indicated by Southern blot analysis of one such clone, Δ*lplA*$_{SAT}$*/lplA*$_{HYG}$*/LPLA*, the endogenous gene is still present (see Figure 4.2, lane 4). In fact, as shown by FACS analysis (Figure 4.4) the Δ*lplA*$_{SAT}$*/lplA*$_{HYG}$*/LPLA* line is tetraploid, meaning that non-dividing cells possess four copies of each chromosome, instead of the normal two copies in diploid cells. This indicates that the double integrant Δ*lplA*$_{SAT}$*/lplA*$_{HYG}$*/LPLA* still possesses two copies of the endogenous *LPLA* allele, despite having correct integration of *LPLA-SAT* and *LPLA-HYG* knockout cassettes. This is the case for all clones isolated, as determined by PCR (data not shown). It can be determined that for Δ*lplA*$_{SAT}$*/lplA*$_{HYG}$*/LPLA*, genome duplication must have occurred during the selection process for *LPLA-HYG* integration, since the parent line, Δ*lplA*$_{SAT}$*/LPLA*, is diploid (see Figure 4.4). In addition, *LPLA* Δ*lplA*$_{SAT}$*/lplA*$_{HYG}$*/LPLA* seems to have growth defects compared to wild-type, including slower logarithmic development, and premature entry into stationary phase (see Figure 4.3).

In order to provide more convincing evidence that *LPLA* is an essential/very important gene, a complementation construct, *LPLA*-pRB, was generated and transfected into Δ*lplA*$_{SAT}$*/LPLA* and Δ*lplA*$_{HYG}$*/LPLA* lines to give rise to

Δ*lplA*$_{SAT}$/*LPLA*::*P*$_{rRNA}$*LPLA* and Δ*lplA*$_{HYG}$/*LPLA*::*P*$_{rRNA}$*LPLA* lines, respectively (see Sections 2.2.4–2.2.9). Clones were tested for *LPLA*-pRB integration by Southern blot analysis. The pRB construct is designed to integrate into any of the six *18S SSU rRNA* gene copies present on chromosome 27 (Misslitz *et al.*, 2000). Therefore, depending upon the site of *LPLA*-pRB integration, a Southern blot probed with a *LPLA* gene probe on *Nru*I-digested gDNA would result in a band size of 8.4 kb, 6.3 kb, 6.4 kb, 9.0 kb, 7.4 kb or 5.5 kb (see Figure 4.2B). Figure 4.2C shows a 6.3 kb/6.4 kb band for Δ*lplA*$_{HYG}$/*LPLA*::*P*$_{rRNA}$*LPLA.* However, the band size for Δ*lplA*$_{SAT}$/*LPLA*::*P*$_{rRNA}$*LPLA* is larger than 9 kb. In this case, *LPLA*-pRB could have integrated non-specifically into the *L. major* genome. Alternatively, *LPLA*-pRB could have integrated correctly, and subsequently duplicated within the *18S SSU rRNA* locus, resulting in a band size double that of 5.5 kb or 6.3 kb/6.4 kb. However, the nature of the > 9 kb band was not investigated further.

Finally, Δ*lplA*$_{SAT}$/*LPLA*::*P*$_{rRNA}$*LPLA* and Δ*lplA*$_{HYG}$/*LPLA*::*P*$_{rRNA}$*LPLA* lines were transfected with *LPLA-HYG* or *LPLA-SAT* knockout cassettes, respectively, in order to create null mutants of the endogenous *LPLA* allele (Δ*lplA*::*P*$_{rRNA}$*LPLA*). Analysis of a total of 43 clones from both transfections (see Table 4.1) revealed only one endogenous gene knockout, whereas the remaining clones retained the *LPLA* gene. Figure 4.2C shows one clone with gene ablation, Δ*lplA*$_{SAT}$/*lplA*$_{HYG}$::*P*$_{rRNA}$*LPLA*. Another clone has correct second-round integration but still retains at least one endogenous gene copy (Δ*lplA*$_{HYG}$/*lplA*$_{SAT}$/*LPLA*::*P*$_{rRNA}$*LPLA*). These results indicate that replacement of the *LPLA* allele is possible when *LPLA* is ectopically expressed.

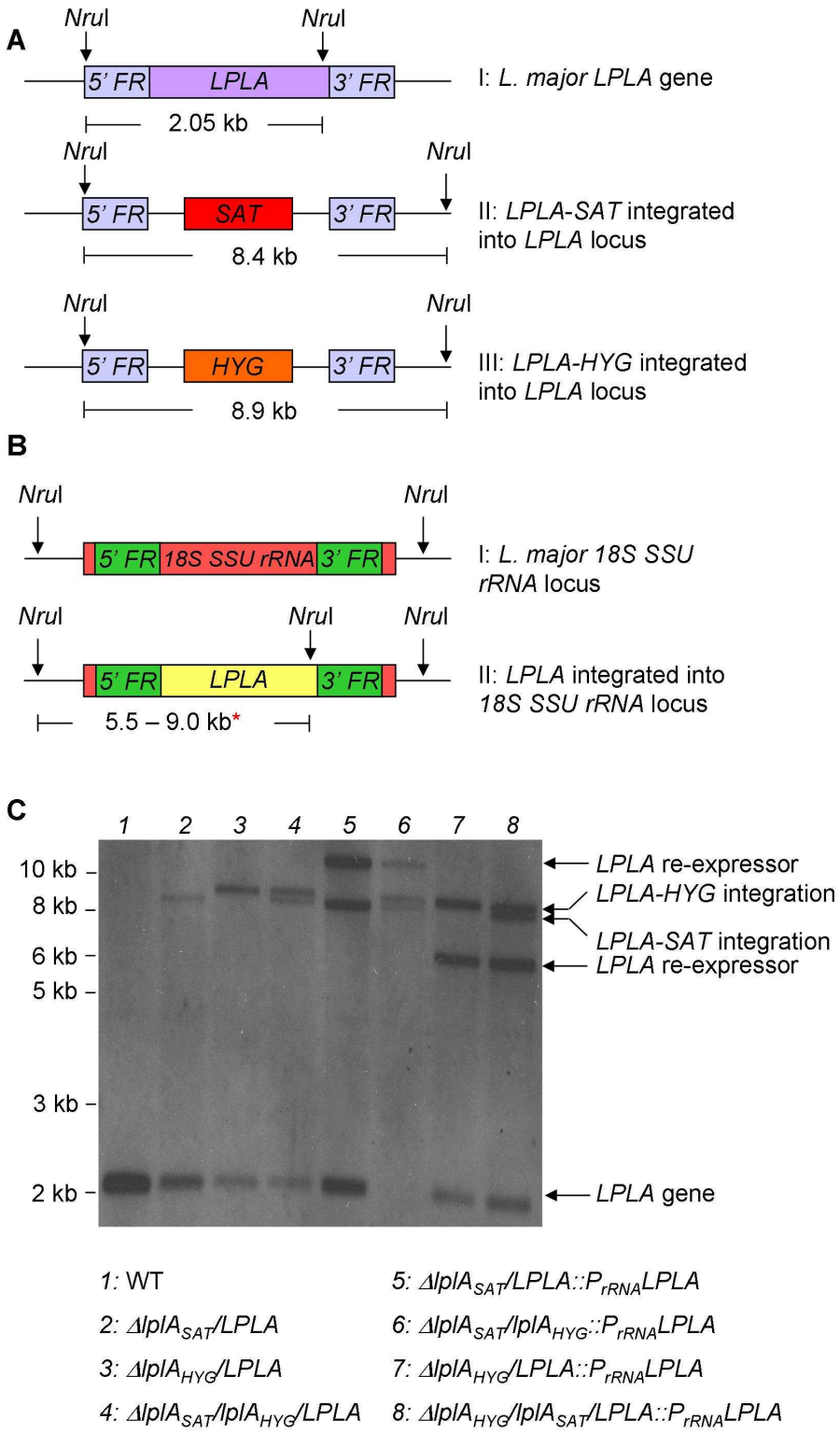


Figure 4.2 Knockout attempts of *LPLA* gene, and complementation to permit gene ablation
A, Schematic representation of the endogenous *LPLA* locus (I), and the locus after *LPLA-SAT* (II) or *LPLA-HYG* (III) integration. *B*, Complementation of a knockout necessitated replacement of one of the six *18S SSU rRNA* gene copies (I) with the *LPLA* gene, under puromycin selection (II). Sizes in kb represent the expected band sizes to be observed by Southern blot (*C*). *Depending upon which of the six *18S SSU rRNA* gene copies is replaced by the *LPLA* re-expressor, *Nru*I-digested gDNA and Southern blot with *LPLA* gene probe may yield either 8.4 kb, 6.3 kb, 6.4 kb, 9 kb, 7.4 kb or 5.5 kb bands. *C*, Southern blot of 2 μg gDNA digested with *Nru*I and probed with *HYG*-, *SAT*- and *LPLA* gene probes, with expected bands shown in A and B. '*LPLA*-re-expressor' represents *LPLA-pRB* integration into the *18S SSU rRNA* locus.

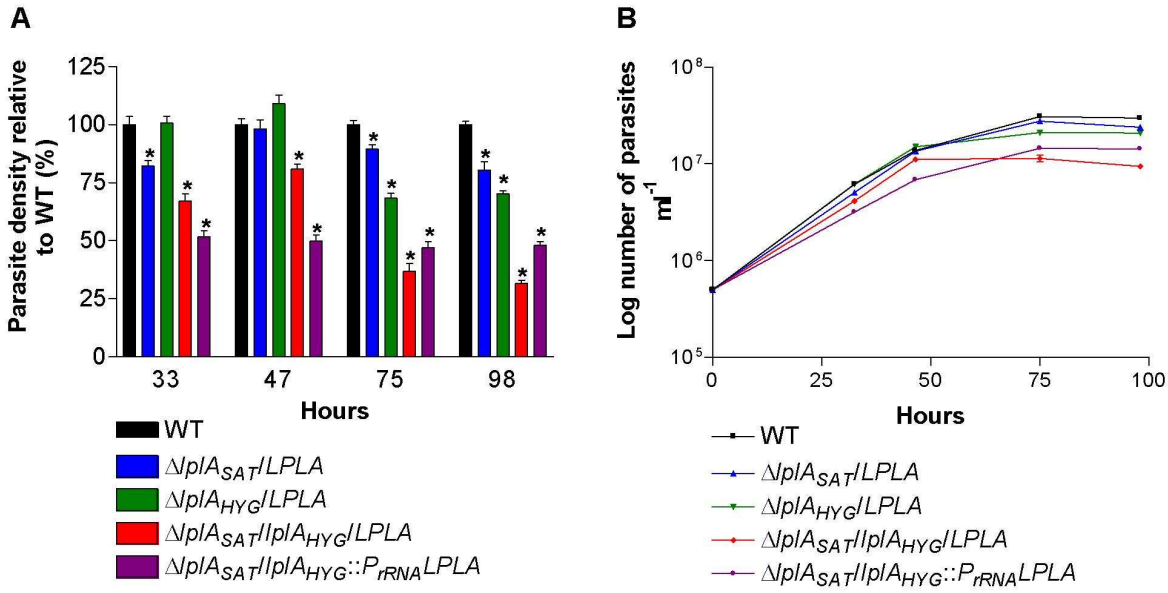


Figure 4.3 Growth of *LPLA* knockout mutants
Growth analysis of different *LPLA* knockout mutants was carried out by diluting all cultures to 5 x 10^5 parasites ml^{-1}, and recording parasite density circa every 24 h. All lines were cultured in duplicate. *A,* Average and standard deviation values were calculated for each line at individual time points. Values were normalised to wild-type values and plotted as a bar chart. One-way ANOVA with Tukey post-tests were performed at each time point. Values that are significantly different from wild-type values (p-value < 0.05) are depicted with the '*' symbol. *B,* Growth of parasite lines represented by a line graph with logarithmic *Y*-axis.

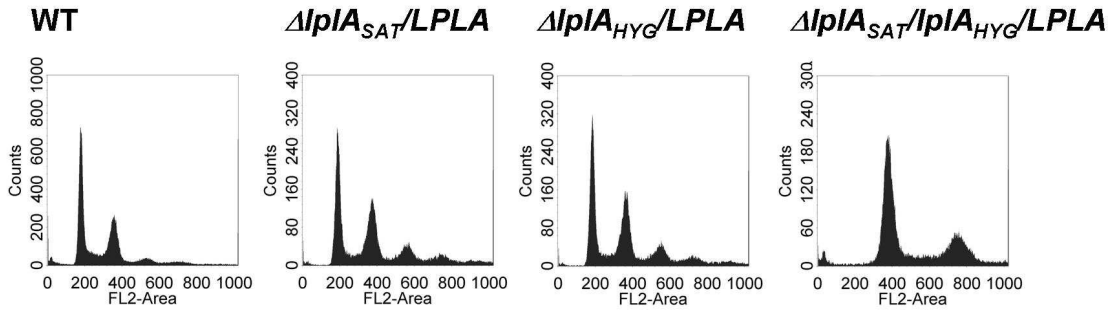


Figure 4.4 Ploidy of *LPLA* knockout mutants
FACS was used in order to determine the ploidy of different *LPLA* knockout mutants. Cells were fixed in 70 % methanol and incubated with propidium iodide. Incorporation of propidium iodide into DNA was detected with a Becton Dickinson FACS Calibur using FL2-A (detecting fluorescence emissions between 585 – 642 nm). Data were interpreted using the CellQuestPro software. Histograms show cell count relative to FL2 fluorescence (arbitrary units). Peaks at 200-, 400- and 800 FL2 arbitrary units represent diploid, tetraploid and octaploid DNA contents, respectively.

	Total numbers analysed			
Transfections	**Clones**	**Transfections**	**Integration**	**Gene knockout**
Δ*lplA* $_{HYG}$/*LPLA* + *LPLA*-SAT	9*	3	N/A	N/A
Δ*lplA* $_{SAT}$/*LPLA* + *LPLA*-HYG	63	3	30/63	0/63
Δ*lplA* $_{HYG}$/*LPLA* ::P_{rRNA} *LPLA* + *LPLA*-SAT	23	2	2/2	0/23
Δ*lplA* $_{SAT}$/*LPLA* ::P_{rRNA} *LPLA* + *LPLA*-HYG	20	2	3/3	1/20

Table 4.1 Summary table of *LPLA* knockout attempts
The information provided illustrates the total number of second round knockout attempts on *LPLA* heterozygotes, and those carried out using heterozygotes re-expressing *LPLA* within the *18S SSU rRNA* locus. All values were obtained through PCR-based knockout screens. *Only nine clones were attained from three transfections of Δ*lplA*$_{HYG}$/*LPLA* + *LPLA-SAT*, and all were non-viable after sub-passaging into larger volumes of HOMEM + 10 % FCS. As such, gDNA was not attained from these lines, and therefore, integration or gene knockout determination is not applicable (N/A).

4.4 Overexpression of LPLA-His

Given that it was not possible to study the function of *LPLA* by gene knockout, an alternative approach was taken; LPLA-His expression/overexpression in promastigotes. Wild-type *LPLA* (*LPLA*) and *LPLA* containing a single point mutation resulting in a H118A amino acid mutation (*LPLA*^{H118A}), were sub-cloned into the pGL1137 plasmid from the pGL1132 plasmid (see Section 4.2). When expressed in *L. major*, the constructs result in the translation of LPLA proteins with *C*-terminal 6x His-tags (see Section 2.2.6). The reason for expressing LPLA-His proteins instead of untagged versions was in order to be able to detect ectopic protein expression, since (as explained in Section 3.5.4) antibodies that detect *Lmj*LPLA were not available.

As discussed in Section 3.4.3, LPLA is a member of a larger family encompassing LPLA, LT and BPL proteins, all of which share a structurally- and catalytically similar *N*-terminal domain (Fujiwara *et al.*, 2005; Fujiwara *et al.*, 2007; Kim do *et al.*, 2005; McManus *et al.*, 2005; Wilson *et al.*, 1992). The rationale for introducing the mutation LPLA^{H118A} was based on the fact that in all LPLA and LT proteins His^{118} (*Lmj*LPLA numbering) is one of the key residues involved in forming the hydrophobic crevice to which the lipoyl-AMP intermediate binds, and is found within the strictly conserved ^{108}RRx$_2$GGGxV(F/Y)**H**D^{119} motif (see Figure 3.4) (Fujiwara *et al.*, 2005; Kim do *et al.*, 2005; McManus *et al.*, 2005). Specifically, the side-chain of His^{118} is thought to form van der Waals interactions with the lipoyl-moiety of the lipoyl-AMP intermediate (Fujiwara *et al.*, 2007). Published data highlight the importance of the ATP-binding pocket in *E. coli* BPL (*Ec*BPL) (Choi-Rhee *et al.*, 2004; Cronan, 2005; Kwon & Beckett, 2000; Reche, 2000; Wilson *et*

al., 1992). *Ec*LplA and *Ec*BPL enzymes operate in a similar fashion; via a two-step biotin/LA activation and transfer mechanism (see Section 3.4.3) (Fujiwara *et al.*, 2005; Wilson *et al.*, 1992). Ordinarily in *E. coli*, the post-translational ligation of biotin (biotinylation) is highly specific to a single apoprotein, biotin carboxyl carrier protein (BCCP) (Beckett & Matthews, 1997; Cronan, 1989). However, a mutant within the ATP-binding domain, *Ec*BPL^{R118G}, biotinylates substrates other than BCCP (Choi-Rhee *et al.*, 2004; Cronan, 2005). The most likely explanation for this "promiscuous biotinylation" was found to be due to the activated biotin intermediate (biotinoyl-AMP) escaping more easily from the *Ec*BPL^{R118G} active site and biotinylating lysine residues on non-specific substrates by chemical (not enzymatic) acylation, in a concentration-dependent manner (of both enzyme and substrate) (Choi-Rhee *et al.*, 2004; Cronan, 2005). Given that LplA and BPL enzymes share a similar mechanism of catalysis, the question was asked as to whether ectopic expression of LPLA^{H118A} in *L. major* promastigotes would result in growth and/or lipoylation pattern phenotypes.

WT[*LPLA-His*] and WT[*LPLA^{H118A}-His*] lines were analysed by determining parasite growth rates relative to those of wild-type and WT[uncut pGL1137] negative controls. In addition, soluble protein was isolated during logarithmic growth and stationary phase and western blotting carried out with α-His-, α-LA-, α-HASPB- and α-*Lmj*LIPA antibodies, and α-cysteine synthease (α-CS) antibody to verify equal loading. At the final time point in stationary phase, in addition to harvesting mixed promastigotes, metacyclic promastigotes were isolated by a peanut agglutination protocol (see Section 2.2.2). All parasite lines were cultured in duplicate, and each experiment was independently repeated. Repetitions yielded very similar results; a representative set of which are shown.

Firstly, LPLA-His expressing lines grown on 10 μg ml^{-1} or 50 μg ml^{-1} G418 drug selection pressures exhibited growth phenotypes; relative to the WT[uncut pGL1137] negative control, LPLA-His expressing lines enter stationary phase at a significantly earlier density ($p<0.05$) (see Figure 4.5). In terms of growth, both LPLA-His lines exhibit similar phenotypes, irrespective of G418 concentration (see Figure 4.5). Secondly, western blot analyses show that LPLA-His and LPLA^{H118A}-His proteins are expressed at similar levels, irrespective of whether the mutant lines were cultured in the presence of 10 μg ml^{-1} or 50 μg ml^{-1} G418 drug selection pressure (see Figure 4.6). The calculated molecular size of LPLA-His and

$LPLA^{H118A}$-His proteins from SDS-PAGE was 56.1 kDa (see Section 2.5.5 for basis of calculations). Given that the LPLA proteins are fused to a His-tag that is 3.1 kDa, the actual molecular size of the two proteins is 53.0 kDa (56.1 kDa minus 3.1 kDa), which is similar to the predicted size of processed LPLA from the LPLA-GFP fusion protein (52.4 kDa) (see Section 4.2). It is likely therefore that the wild-type LPLA-His and mutant LPLA-His proteins are *N*-terminally processed for mitochondrial transit. Based on the sizes of LPLA calculated from LPLA-GFP (see Section 4.2) and LPLA-His data, it can be estimated that the mitochondrial targeting peptide for LPLA is between 19 – 24 amino acids, instead of 12 amino acids ($LPLA_{T1}$) predicted by TargetP or 36 amino acids ($LPLA_{T2}$) (see Figure 3.4 for LPLA alignment).

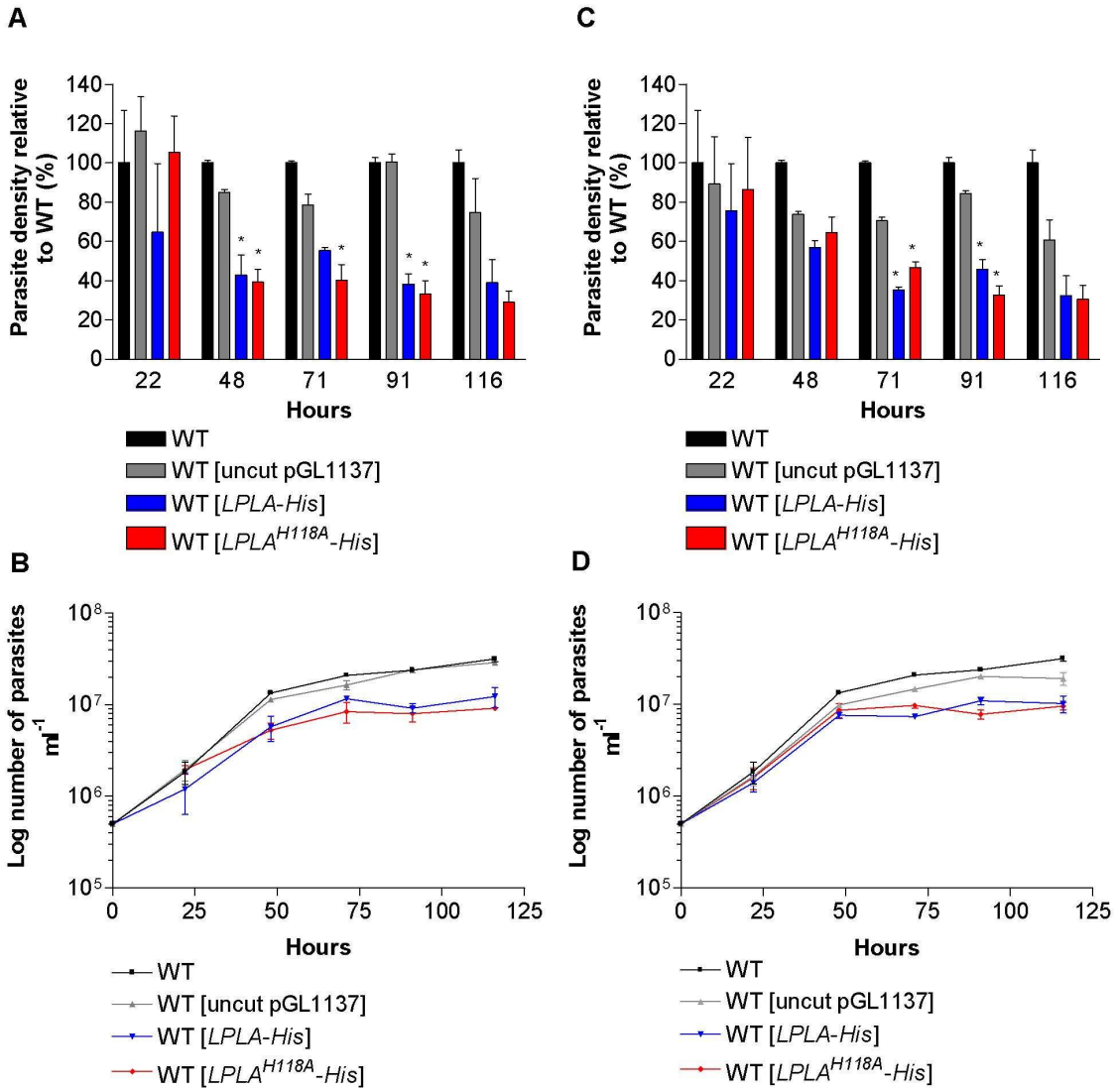


Figure 4.5 Growth of LPLA-His expressing lines
Growth analysis of LPLA-His expressing lines was carried out by diluting all cultures to 5 x 10^5 parasites ml^{-1}, and recording parasite density circa every 24 h. All lines were cultured in duplicate. WT[uncut pGL1137], WT[*LPLA-His*] and WT[*LPLA*^{H118A}*-His*] lines were maintained on 10 µg ml^{-1} (*A & B*) or 50 µg ml^{-1} (*C & D*) G418 drug pressure. *A & C,* Average and standard deviation values were calculated for each parasite line at individual time points. Values were normalised to wild-type and plotted as a bar chart. One-way ANOVA with Tukey post-tests were performed at each time point, with uncut pGL1137 values as control. Values that are significantly different from uncut pGL1137 values (*p*-value < 0.05) are depicted with the '*' symbol. *B,* Growth of parasite lines represented by a line graph with logarithmic *Y*-axis.

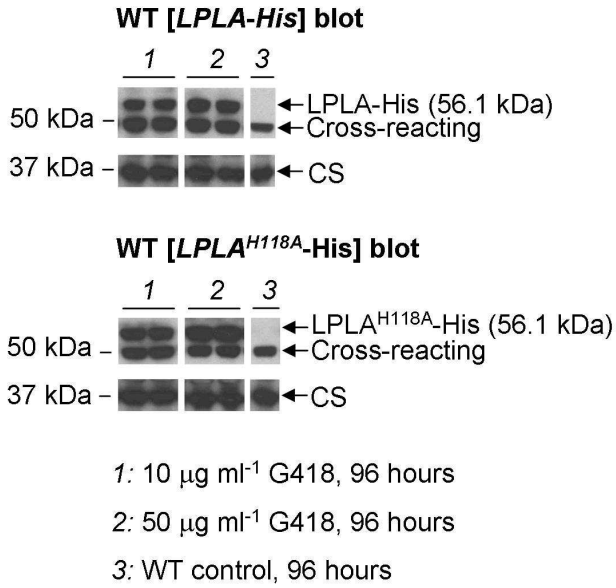


Figure 4.6 Protein expression levels of LPLA-His expressing lines
10 μg soluble protein samples from wild-type, WT[*LPLA-His*] and WT[*LPLA^{H118A}-His*] lines maintained on 10 μg ml^{-1} or 50 μg ml^{-1} G418 drug pressure, from 96 h growth (see Figure 4.5), were subjected to SDS-PAGE and western blot analysis with α-HIS antibody at 1/10,000. The two blots were subsequently stripped of antibody and re-probed with α-CS antibody at 1/7,500 in order to check for equal loading. All lines were grown in duplicate, and both replicates are tested by western blot. The calculated size for both LPLA-His and LPLA^{H118A}-His is 56.1 kDa (see Section 2.5.5 for method of calculating molecular sizes from SDS-PAGE).

As illustrated in Section 3.3, the lipoylation patterns of α-KADH E2 subunits and the H-protein of the GCC can be traced, in a repeatable way, throughout promastigote development *in vitro*. As such, the question was asked as to whether lipoylation patterns would be affected in WT[*LPLA-His*] and/or WT[*LPLA^{H118A}-His*] lines. G418 concentration has no obvious effect on lipoylation pattern; as such, a representative set of data are shown, and the protein used for western blotting was derived from mutant lines grown in the presence of 50 μg ml^{-1} G418. Figure 4.7 shows lipoylation patterns of wild-type and WT[uncut pGL1137] negative control lines, and WT[*LPLA-His*] and WT[*LPLA^{H118A}-His*] lines. Protein loading is comparable in all lanes of each blot, lipoylation patterns of replicate treatments are consistent, and lipoylation patterns of the WT[uncut pGL1137] control are similar to those in wild-type. Figure 4.7 shows that the lipoylation pattern in WT[*LPLA-His*] is comparable to that in the WT[uncut pGL1137] line. The only difference between lipoylation patterns in these two lines compared to wild-type is that E2k, E2p and E2b are lipoylated in wild-type at 116 h mixed promastigote culture, yet only E2k is lipoylated in WT[*LPLA-His*] and WT[uncut pGL1137] lines (see Figure 4.7).

In contrast, lipoylation of all three α-KADH E2 subunits and H-protein is up-regulated in the WT[*LPLA^{H118A}-His*] line at all time points. A very large increase in

lipoylation of the H-protein is apparent at all time points except in purified metacyclic promastigotes. This is in contrast to LPLA-His expression (see Figure 4.7) and LIPB-His expression (see Section 5.4.2), which do not result in an increase in H-protein lipoylation relative to the control lines. Perhaps the most intriguing finding is that metacyclic promastigotes in the WT[*LPLA*^{H118A}*-His*] line possess lipoylated E2k, E2p and E2b subunits, whereas all other lines only have lipoylated E2k and H-protein (see Figure 4.7).

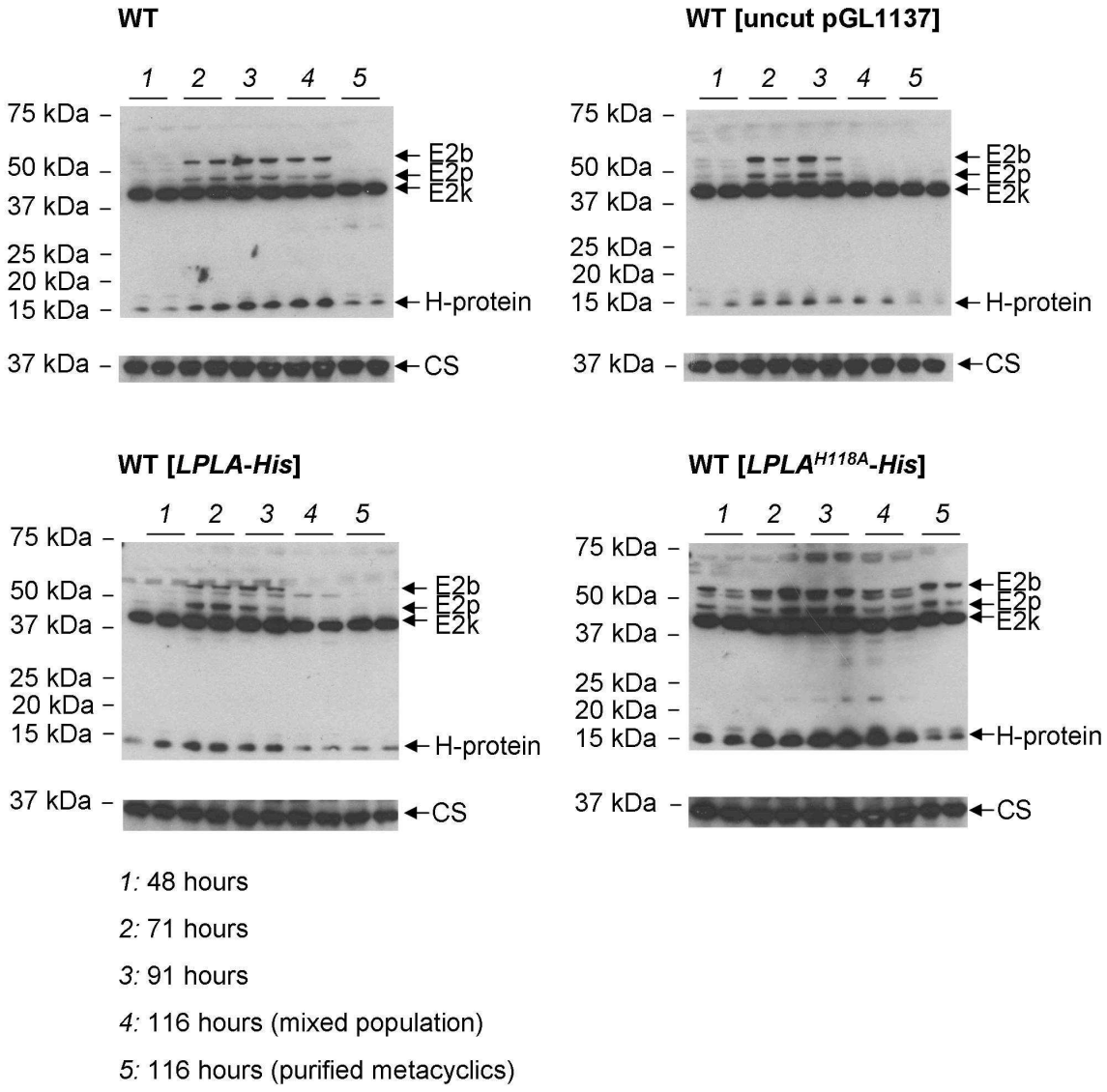


Figure 4.7 Lipoylation patterns of LPLA-His expressing lines
Approximately 1 e^8 cells were lysed at four different time points from cultures maintained on 50 μg ml^{-1} G418 drug pressure (see Figure 4.5). Soluble protein was harvested and subjected to SDS-PAGE. Western blots were carried out with α-LA antibody at 1/6,000 to determine lipoylation patterns, and with α-CS antibody at 1/7,500 to check for equal loading. All lines were grown in duplicate, and both replicates are tested by western blot.

Purified metacyclic promastigotes at the 116 h time point from wild-type and LPLA-His expressing lines grown on different G418 selection pressures were tested by

western blotting for a metacyclic-specific marker, hydrophilic acylated surface protein (HASPB, kind gift from Professor D.F. Smith) (Flinn *et al.*, 1994). Figure 4.8 shows that LPLA-His expressing lines have significantly less HASPB protein than does the wild-type control (p<0.05), suggesting that LPLA-His expression could affect metacyclogenesis. A possible explanation is that since LPLA-His expressing lines enter stationary phase at an earlier time point (see Figure 4.5), and that this time point is most correlated to the production of HASPB-positive metacyclic promastigotes, 116 h was not a suitable time to probe LPLA-His and LPLA^{H118A}-His protein lysates with α-HASPB antibody. This option was ruled out by western blotting with α-HASPB antibody at all time points, which revealed that levels of HASPB expression are consistently lower in LPLA-His expressing lines compared to wild-type (data not shown).

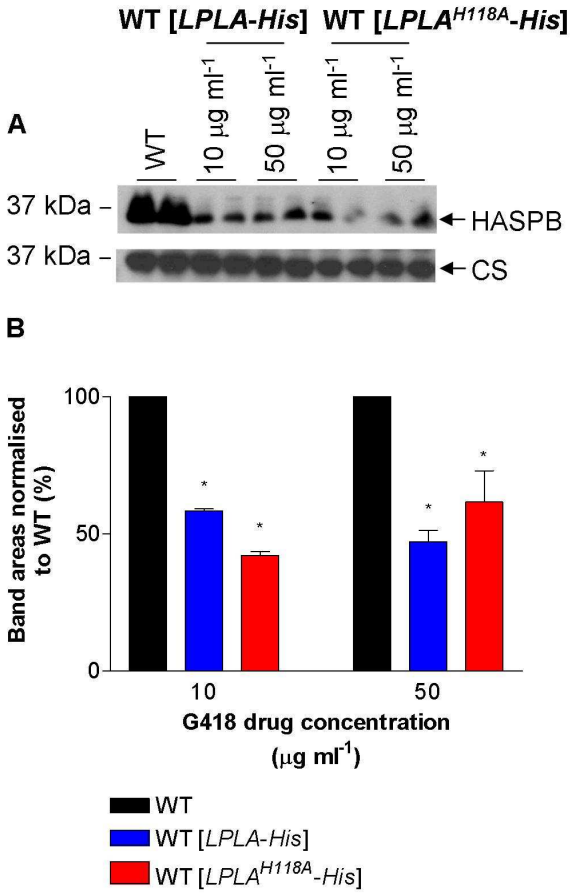
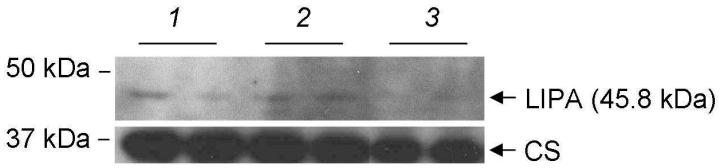


Figure 4.8 Metacyclogenesis of LPLA-His expressing lines
A, 10 μg soluble protein of metacyclic promastigotes (at 116 h time point) purified from LPLA-His expressing lines maintained on 10 μg ml^{-1} or 50 μg ml^{-1} G418 drug pressure, was subjected to SDS-PAGE. Western blotting was subsequently carried out with α-HASPB antibody at 1/4,000 and α-CS antibody at 1/7,500, to assess loading. All lines were grown in duplicate, and both replicates are tested by western blot. *B,* Band areas (shown in *A*) were determined using Scion Image program, averages and standard deviations were calculated and a bar chart plotted. Subsequently, one-way ANOVA with a Tukey post-test was performed to determine whether band areas in WT[*LPLA-His*] and WT[*LPLA^{H118A}-His*] lines were significantly different from those in wild-type, where p-value < 0.05 is indicated by the '*' symbol.

In order to determine whether expression of LPLA-His and LPLA^{H118A}-His proteins resulted in up-regulation of the LA biosynthesis pathway, western blotting was carried out using the α-*Lmj*LIPA antibody (see Section 3.5.4.1). Given that the most notable differences in lipoylation pattern (with regards to the WT[*LPLA*^{H118A}-*His*] line) were observed in metacyclic promastigotes (see Figure 4.7), only protein lysates from this time point were probed with α-*Lmj*LIPA by western blotting. Figure 4.9 shows that in all parasite lines, a band of 45.8 kDa is apparent. The fact that this band corresponds to LIPA is corroborated by the fact that western blotting with α-*Lmj*LIPA antibody on protein lysate from the WT[*LIPA-GFP*] line reveals a 45.3 kDa band, which is most likely to represent degradation of LIPA-GFP into LIPA and GFP (see Figure 5.1). According to prediction programmes TargetP and MitoProt, the expected molecular size of LIPA after cleavage of the *N*-terminal mitochondrial targeting peptide is 44.1 kDa ($LIPA_{T1}$) or 45.5 kDa ($LIPA_{T2}$), respectively (see Section 3.4.2). The average of the bands observed at 45.3 kDa and 45.8 kDa gives 45.5 kDa, which would be in line with the prediction made by MitoProt. Figure 4.9 shows that LIPA band intensity is similar in wild-type and WT[*LPLA-His*] lines, yet perhaps lower in the WT[*LPLA*^{H118A}-*His*] line. However, the level of CS protein in protein lysates from the WT[*LPLA*^{H118A}-*His*] line is lower than in the other two lines, and as such it is probable that the level of LIPA protein is similar in all lanes. This implies that any phenotypes, such as alterations in lipoylation patterns, are probably due to expression of LPLA-His/LPLA^{H118A}-His proteins and not because of up-regulation of LA biosynthesis.



1: WT metacyclic promastigote lysates

2: WT [*LPLA-His*] metacyclic promastigote lysates

3: WT [*LPLA*^{H118A}-*His*] metacyclic promastigote lysates

Figure 4.9 LIPA expression levels in LPLA-His expressing lines
Western blotting of protein lysates of purified metacyclic promastigotes from the 116 h time point (only lines grown on 50 μg ml^{-1} G418 selection pressure) using α-*Lmj*LIPA antibodies at 1/20 to determine LIPA expression levels in LPLA-His expressing lines relative to the wild-type control. Blots were also probed with α-CS antibody at 1/7,500 to check for equal loading. All lines were grown in duplicate, and both replicates are tested by western blot.

4.5 Summary

- LPLA-GFP is a mitochondrial protein, as shown by fluorescence microscopy in which LPLA-GFP (FITC filter) and Mitotracker CMXRos (Rhodamine filter) overlay.

- *LPLA* is essential/very important to promastigote survival, since 0/72 clones from six independent transfections resulted in gene ablation.

- Both *LPLA* alleles can be replaced, although all clones with double allele replacement still possessed at least one copy of the endogenous *LPLA* gene; apparently achieved by the parasites becoming tetraploid.

- A permissive knockout strategy using an integrating *LPLA* re-expressor construct allowed complete gene replacement. The *LPLA* knockout with re-expressor grew significantly slower than wild-type and entered stationary phase at a substantially lower cell density.

- Expression of *C*-terminally His-tagged LPLA and active site mutant LPLA^{H118A} resulted in the onset of stationary phase markedly earlier than in control lines.

- The lipoylation pattern in WT[*LPLA-His*] was very similar to the usual wild-type pattern, whereas expression of LPLA^{H118A}-His resulted in higher lipoylation of all three α-KADH E2 subunits and the H-protein of the GCC at all time points. Intriguingly, metacyclic promastigotes in the WT[*LPLA^{H118A}-His*] line lipoylated E2p and E2b in addition to E2k and H-protein, which was not observed in wild-type, WT[uncut pGL1137] or WT[*LPLA-His*] lines.

- Expression of LPLA-His or LPLA^{H118A}-His resulted in the isolation of a lower proportion of HASPB-positive non-procyclic promastigotes, suggesting that LPLA-His expression could affect metacyclogenesis.

- There is no noticeable increase in LIPA protein in LPLA-His expressing lines relative to the wild-type control, indicating that the changes in lipoylation patterns observed in the WT[*LPLA^{H118A}-His*] line are most likely a result of LPLA^{H118A}-His expression and not due to up-regulation of LA biosynthesis. This increase in

lipoylation could be due to promiscuous lipoylation of the LPLA^{H118A}-His mutant enzyme.

5 Biosynthesis of lipoic acid

5.1 Introduction

L. major encodes putative LIPA and LIPB enzymes, which were shown to be active by functional complementation of *E. coli* deficient in *lipA* and *lipB*, respectively, as described in Section 3.5. In the study reported in this chapter, the aims were to determine the sub-cellular localisations of LIPA and LIPB and to discern the roles that each plays in *L. major*. The former question was addressed by expressing a LIPA-GFP fusion protein in promastigotes, and establishing the sub-cellular localisation of LIPA-GFP by live cell fluorescence microscopy. An alternative approach was taken to determine the sub-cellular localisation of LIPB, which involved a pre-fractionation technique using digitonin titration. In order to elucidate the function of LIPA in promastigotes, attempts were made to ablate both gene copies by homologous recombination. In order to study the function of LIPB, wild-type LIPB with a *C*-terminal His-tag was ectopically expressed in promastigotes, with the aim of perturbing the normal balance of lipoylation, which would perhaps incite a decrease in parasite fitness.

5.2 Localisation of LIPA

To generate a *L. major* line expressing a *C*-terminal fusion protein of LIPA with GFP, the full length *LIPA* gene was amplified by PCR using primer pair Lm19-20 (see Section 2.1.4). The resulting 1.3 kb fragment was digested with *Nde*I/*Kpn*I before cloning into *Nde*I/*Kpn*I digested pGL1132 vector (Tetaud *et al.*, 2002) to yield the construct *LIPA-GFP* (see Figure 5.1A). The destination vector pGL1132 contains the required control elements that allow constitutive expression of the proteins in *L. major* as well as the neomycin phosphotransferase (NEO) selectable marker (see Sections 2.2.6 and 2.2.7) (Cruz & Beverley, 1990; Tetaud *et al.*, 2002). The *LIPA-GFP* construct was transfected into wild-type to give rise to the WT[*LIPA-GFP*] line, which was grown in the presence of 10 μg ml^{-1} G418 to select for parasites that possess ectopic copies of the *LIPA-GFP* plasmid (see Section 2.2.7). It is interesting to note that unlike the WT[*LPLA-GFP*] line, which grew efficiently when selecting with 50 μg ml^{-1} G418 (see Section 4.2), the WT[*LIPA-GFP*] line was only able to proliferate when the concentration of G418 was

lowered to 10 μg ml^{-1}; suggesting that expression of LIPA-GFP is not as well tolerated as that of LPLA-GFP.

In order to verify that the LIPA-GFP fusion protein was expressed in the WT[*LIPA-GFP*] line, western blotting with α-GFP monoclonal antibody and with α-*Lmj*LIPA antibody was carried out. Figure 5.1B shows that both antibodies recognised a band with a calculated molecular size of 72.2 kDa. Unlike in the WT[*LPLA-GFP*] line where no GFP degradation was apparent, a clear 26.9 kDa band was detected in WT[*LIPA-GFP*] protein lysate using α-GFP antibody, with an expected GFP band size of 27.1 kDa (see Figure 5.1B). Probing WT[*LIPA-GFP*] protein lysate with α-*Lmj*LIPA antibody resulted in a 45.3 kDa band on a western blot. Probing wild-type protein lysate with α-*Lmj*LIPA antibody resulted in the recognition of a 45.8 kDa band (see Section 4.4, Figure 4.9). According to prediction programmes TargetP and MitoProt, the expected molecular size of *Lmj*LIPA after cleavage of the *N*-terminal mitochondrial targeting peptide would be 45.5 kDa (*Lmj*$LIPA_{T1}$) or 44.1 kDa (*Lmj*$LIPA_{T2}$), respectively (see Section 3.4.2). The average of the bands observed at 45.3 kDa and 45.8 kDa is 45.5 kDa, which would be in line with the prediction made by TargetP, and would equate to a targeting peptide of five amino acids.

Fluorescence microscopy to detect the LIPA-GFP reporter, in combination with Mitotracker CMXRos mitochondrial stain, revealed that LIPA-GFP is a mitochondrial protein (see Figure 5.1C). This result is in accordance with sub-cellular targeting predictions using MitoProt and TargetP (see Section 3.4.2).

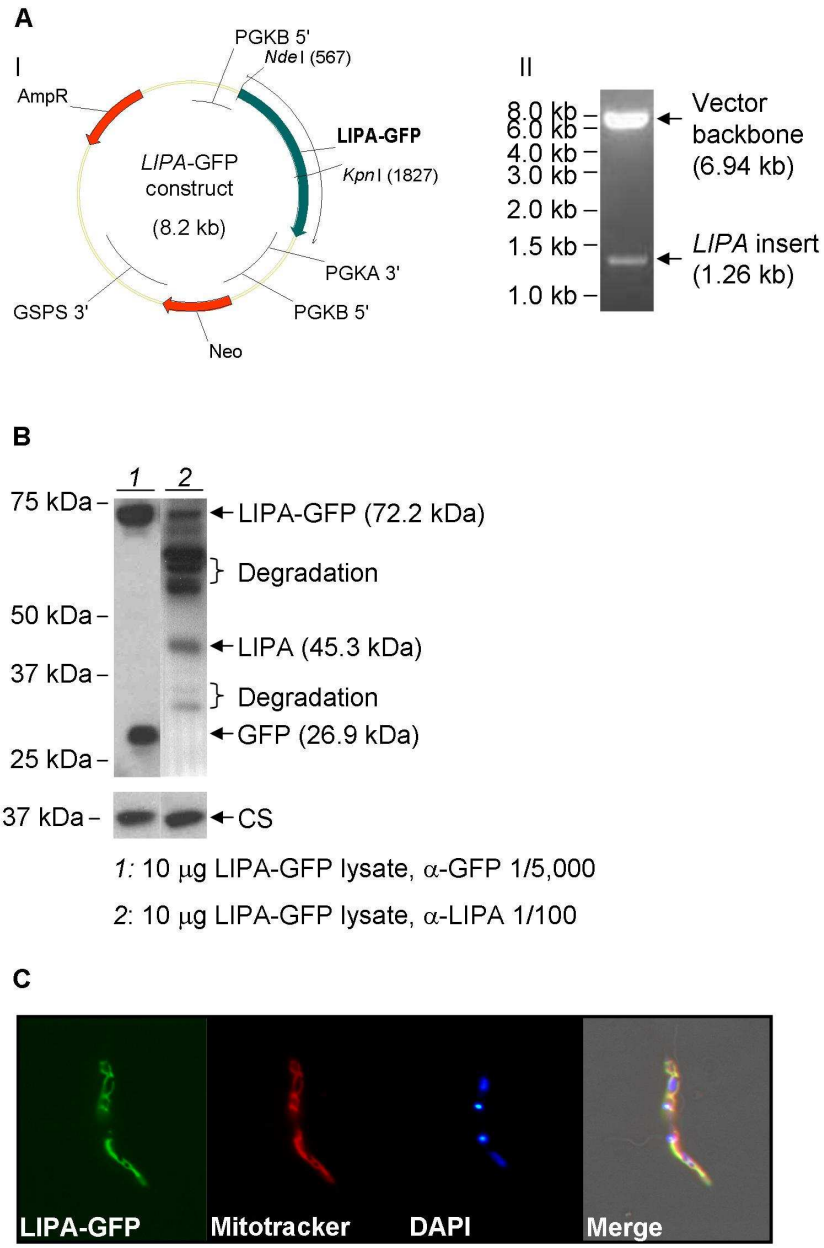


Figure 5.1 Sub-cellular localisation of LIPA-GFP
A, The *LIPA* gene (see Section 3.4.2) was amplified from gDNA with primer pair Lm19-20 (see Section 2.1.4). The insert was cloned into expression vector pGL1132, generating construct *LIPA-GFP* (I). Correct cloning was verified by *Nde*I/*Kpn*I digest, releasing a 1.26 kb insert corresponding to *LIPA*, and vector backbone of 6.94 kb (II). *B*, The *LIPA-GFP* construct was expressed in promastigotes to give rise to the WT[*LIPA-GFP*] line, and expression verified by western blotting of 10 μg LIPA-GFP protein lysate with α-GFP antibody at 1/5,000 or α-*Lmj*LIPA antibody at 1/20. Equal loading was verified by probing with α-CS antibody at 1/7,500. The band observed for LIPA-GFP has a calculated size of 72.2 kDa. The calculated sizes of LIPA and GFP are 45.3 kDa and 26.9 kDa, respectively (see Section 2.5.5 for method of calculating molecular weights from SDS-PAGE). *C*, The mitochondrion of WT[*LIPA-GFP*] promastigotes was stained with Mitotracker CMXRos, and images of live cells were attained by fluorescence microscopy with DIC (for phase), FITC (for *LIPA-GFP*), rhodamine (for Mitotracker) or DAPI filters.

5.3 Knockout studies of *LIPA*

In order to study the *in vivo* role of LIPA, a gene replacement strategy was employed. Knockout cassettes *LIPA-SAT* and *LIPA-HYG* were generated, and *LIPA-SAT* was transfected into promastigotes, as described in Section 2.2.4.

Figure 5.2 is representative of the results obtained from *LIPA* knockout attempts, a summary of which is provided in Table 5.1. Two independent heterozygous lines for *LIPA* (Δ*lipA*$_{SAT}$/*LIPA*) were created by replacement of one *LIPA* allele with the *LIPA-SAT* knockout cassette (see Figure 5.2C, lane 2 for one of the two heterozygotes). Based on growth data generated in duplicate from one of the heterozygous lines (other heterozygote exhibits the same phenotype; data not shown), Δ*lipA*$_{SAT}$/*LIPA* reached a significantly lower cell density at stationary phase than did the wild-type control (see Figure 5.3)

Attempts were made to replace the remaining *LIPA* allele with the *LIPA-HYG* knockout cassette, and in all three transfections carried out 10 μM LA was included in the medium at all steps of the selection process. Therefore, if LA salvage and LA biosynthesis pathways are redundant, the theory was that supplementation with LA could permit *LIPA* gene ablation. In total, 61 clones from three independent transfections of Δ*lipA*$_{SAT}$/*LIPA* lines transfected with the *LIPA-HYG* knockout cassette were analysed by PCR for the presence of the gene and for *LIPA-HYG* integration. 33/61 clones had correct *LIPA-HYG* integration yet 0/61 had gene knockout (See Table 5.1, PCR data not shown). Southern blot analysis of one such double integrant (Δ*lipA*$_{SAT}$/*lipA*$_{HYG}$/*LIPA*) with *LIPA*, *SAT* and *HYG* gene probes showed that this clone indeed has *LIPA-SAT* and *LIPA-HYG* integration, and still possesses the endogenous gene (see Figure 5.2, lane 3). This clone was analysed by FACS, which illustrated that Δ*lipA*$_{SAT}$/*lipA*$_{HYG}$/*LIPA* is tetraploid, meaning that non-dividing cells possess four copies of each chromosome, instead of the normal two copies in diploid cells (see Figure 5.4). In addition, Δ*lipA*$_{SAT}$/*lipA*$_{HYG}$/*LIPA* seems to have growth defects compared to wild-type, including slower logarithmic development, and premature entry into stationary phase (see Figure 5.3).

In order to provide more convincing evidence that *LIPA* is an essential/very important gene, a complementation construct, *LIPA-pRB*, was generated and

transfected into one of the two Δ*lipA*$_{SAT}$/*LIPA* lines to give rise to the Δ*lipA*$_{SAT}$/*LIPA*::*P*$_{rRNA}$*LIPA* line (see Sections 2.2.4-2.2.9). As discussed in Section 4.3, the *pRB* construct is designed to integrate into any of the six *18S SSU rRNA* gene copies present on chromosome 27 (Misslitz *et al.*, 2000). Therefore, depending upon the site of *LIPA-pRB* integration, a Southern blot probed with a *LIPA* gene probe on *Nru*I-digested gDNA would result in a band size of 13.2 kb, 11.0 kb, 11.2 kb, 13.8 kb, 12.2 kb or 10.3 kb (see Figure 5.2B). Figure 5.2C (lane 4) shows that Southern blotting with *LIPA* gene probe on *Nru*I-digested gDNA from the Δ*lipA*$_{SAT}$/*LIPA*::*P*$_{rRNA}$*LIPA* line resulted in a band size >10 kb. It is not possible to interpret which of the six *18S SSU rRNA* gene copies has been replaced by *LIPA-pRB* integration, since the DNA ladder does not surpass 10 kb.

Subsequently, two independent transfections were carried out in order to attempt the replacement of the remaining endogenous *LIPA* allele in the Δ*lipA*$_{SAT}$/*LIPA*::*P*$_{rRNA}$*LIPA* line, with the *LIPA-HYG* knockout cassette. A total of 23 clones were analysed by PCR for the presence of the endogenous *LIPA* gene as well as for integration of the *LIPA-HYG* knockout cassette (data not shown); it was found that 14/23 clones had correct integration of *LIPA-HYG*, yet 0/23 clones were lacking the endogenous *LIPA* gene (see Table 5.1). Southern blot analysis indicated that one such clone possessed an unexpected band above 10 kb instead of the expected band for *LIPA-HYG* integration (see Figure 5.2C, lane 5). The reason why this clone (and presumably other clones) appears to have correct *LIPA-HYG* integration by PCR (using primers amplifying from outside of the 5' FR used in the construct to within the *HYG* gene) and not by Southern blot is unclear.

	Total numbers analysed			
Transfections	**Clones**	**Transfections**	**Integration**	**Gene knockout**
lipA $_{SAT}$/*LIPA* + *LIPA* -HYG	61	3	33/61	0/61
lipA $_{SAT}$ /*LIPA* ::*P* $_{rRNA}$ *LIPA* + *LIPA*- HYG	23	2	14/23	0/23

Table 5.1 Summary table of *LIPA* knockout attempts
The information provided illustrates the total number of second round knockout attempts on *LIPA* heterozygotes, and those carried out using heterozygotes re-expressing *LIPA* within the *18S SSU rRNA* locus. All values were obtained through PCR-based knockout screens.

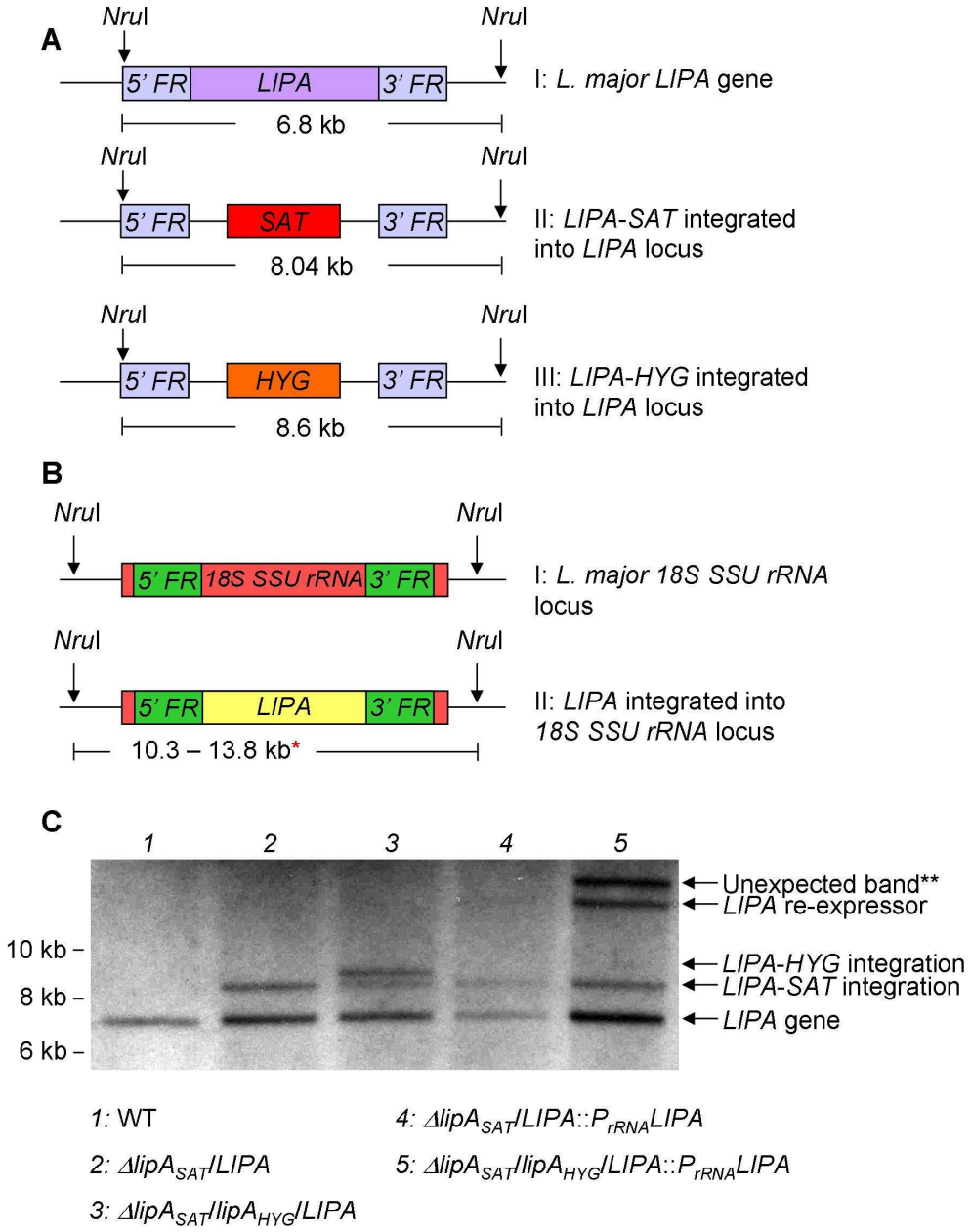


Figure 5.2 Knockout attempts of *LIPA* gene, and complementation in an attempt to permit gene ablation

A, Schematic representation of the endogenous *LIPA* locus (I), and the locus after *LIPA-SAT* (II) or *LIPA-HYG* (III) integration. *B*, Complementation of a knockout necessitated replacement of one of the six *18S SSU rRNA* gene copies (I) with the *LIPA* gene, under puromycin selection (II). Sizes in kb represent the expected band sizes to be observed by Southern blot (*C*). *Depending upon which of the six *18S SSU rRNA* gene copies is replaced by the *LIPA* re-expressor, *Nru*I-digested gDNA and Southern blot with *LIPA* gene probe may yield either 13.2 kb, 11.0 kb, 11.2 kb, 13.8 kb, 12.2 kb, 10.3 kb bands. *C*, Southern blot of 2 µg gDNA digested with *Nru*I and probed with *HYG*-, *SAT*- and *LIPA* gene probes, with expected bands shown in A and B. '*LPLA*-re-expressor' represents *LPLA-pRB* integration into the *18S SSU rRNA* locus. The '**' symbol represents an unknown band, which could correspond to unexpected *LIPA-HYG* integration (see text for explanation).

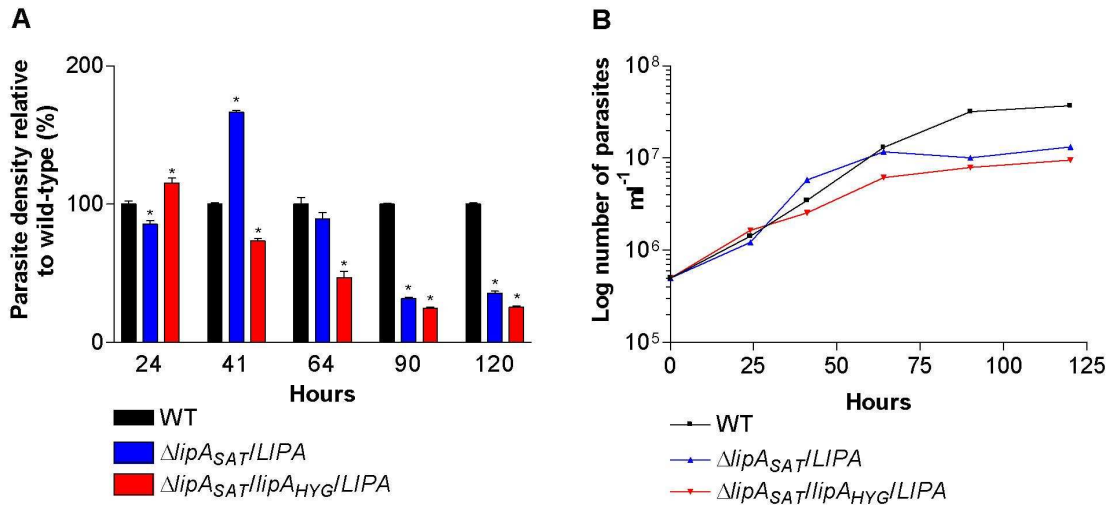


Figure 5.3 Growth of *LIPA* knockout mutants
Growth analysis of different *LIPA* knockout mutants was carried out by diluting all cultures to 5 x 10^5 parasites ml^{-1}, and recording parasite density circa every 24 h. All lines were cultured in duplicate. *A,* Average and standard deviation values were calculated for each line at individual time points. Values were normalised to wild-type values and plotted as a bar chart. One-way ANOVA with Tukey post-tests were performed at each time point. Values that are significantly different from wild-type values (*p*-value < 0.05) are depicted with the '*' symbol. *B,* Growth of parasite lines represented by a line graph with logarithmic *Y*-axis.

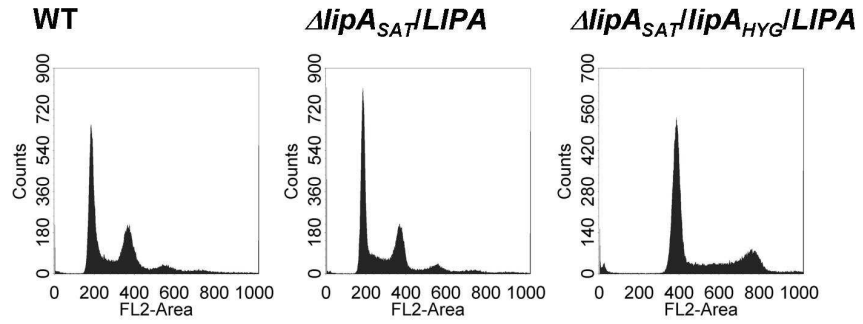


Figure 5.4 Ploidy of *LIPA* knockout mutants
FACS was used in order to determine the ploidy of different *LIPA* knockout mutants. Cells were fixed in 70 % methanol and incubated with propidium iodide. Incorporation of propidium iodide into DNA was detected with a Becton Dickinson FACS Calibur using FL2-A (detecting fluorescence emissions between 585 – 642 nm). Data were interpreted using the CellQuestPro software. Histograms show cell count relative to FL2 fluorescence (arbitrary units). Peaks at 200-, 400- and 800 FL2 arbitrary units represent diploid, tetraploid and octaploid DNA contents, respectively.

5.4 Overexpression of LIPB-His

5.4.1 Localisation of LIPB-His

Figure 5.1 illustrates that the LIPA-GFP fusion protein is targeted to the mitochondrion, which infers that LIPA is a mitochondrial protein. Given that LipA has been shown to be dependent upon the action of LipB in *E. coli* (Zhao *et al.*, 2005), and due to the fact that the two proteins co-localise to the apicoplast

compartment in *P. falciparum* (Wrenger & Muller, 2004) and in *T. gondii* (Thomsen-Zieger *et al.*, 2003), it was hypothesised that *Lmj*LIPA and *Lmj*LIPB would also be found within the same organelle. In order to determine whether *Lmj*LIPB is a mitochondrial protein, an alternative approach to the GFP reporter system was adopted; determination of the sub-cellular localisation of a LIPB-His fusion protein in promastigotes by cellular pre-fractionation.

To generate a *L. major* line expressing a *C*-terminal fusion protein of *LIPB* with His-tag, the full length *LIPB* gene was amplified by PCR using primer pair Lm21-22 (see Section 2.1.4). The resulting 0.8 kb fragment was digested with *Nde*I/*Kpn*I before cloning into *Nde*I/*Kpn*I digested pGL1137 vector to yield construct *LIPB-His* (see Figure 5.5A). The destination vector pGL1137 is derived from pGL1132, and as such contains the required control elements that allow constitutive expression of the proteins in *L. major* as well as the neomycin phosphotransferase (NEO) selectable marker (see Section 2.2.6 and 2.2.7) (Cruz & Beverley, 1990; Tetaud *et al.*, 2002). The *LIPB-His* construct was transfected into wild-type to give rise to the WT[*LIPB-His*] line.

In order to determine the sub-cellular localisation of LIPB, parasites from wild-type and WT[*LIPB-His*] lines were permeabilised with increasing quantities of digitonin in order to pre-fractionate (that is, to concentrate) different cellular compartments. Previous digitonin-titration experiments in *L. mexicana* showed that cytosolic-, glycosomal- or mitochondrial protein markers could be detected by western blot when cells were treated with 0.3 mg ml^{-1}, 1.5 mg ml^{-1} or 3.0 mg ml^{-1} to 5.0 mg ml^{-1} digitonin, respectively (Leroux *et al.*, 2006). Similar experiments have been reported in *T. brucei* (Coustou *et al.*, 2005; Marche *et al.*, 2000), *T. cruzi* (Bouvier *et al.*, 2006; Maugeri *et al.*, 2003; Miranda *et al.*, 2008), *Trypanoplasma borelli* (*T. borelli*) (Wiemer *et al.*, 1995) and *L. donovani* and *L. infantum* (Foucher *et al.*, 2006).

The rationale for using the pre-fractionation technique was in order to obtain a semi-pure- and concentrated mitochondrial protein lysate fraction that could be probed with α-*Lmj*LIPB antibody. The different protein lysate fractions were also probed with α-LA antibody, which served as a mitochondrial protein marker control. Figure 5.5B shows that no protein is detected when wild-type protein lysate is probed with α-*Lmj*LIPB antibody, indicating that endogenous *Lmj*LIPB is either expressed at levels that are undetectable to the α-*Lmj*LIPB antibody, or the

protein is not expressed in promastigotes. Figure 5.5B illustrates that the LIPB-His protein is detected by the α-*Lmj*LIPB antibody, yet only in fractions in which protein was solubilised by 3.0 mg ml^{-1} or 5.0 mg ml^{-1} digitonin and not in the 1.5 mg ml^{-1} digitonin fraction. As shown in Figure 5.5B, this pattern coincides with that observed when WT[*LIPB-His*] protein lysate is probed with α-LA antibody, whereby lipoylated E2k, E2p, E2b and H-protein are concentrated in 3.0 mg ml^{-1} and 5.0 mg ml^{-1} digitonin fractions. There is some leakage of E2k and H-protein from cells treated with 1.5 mg ml^{-1} digitonin. Overall, the results indicate that LIPB-His is a mitochondrial protein.

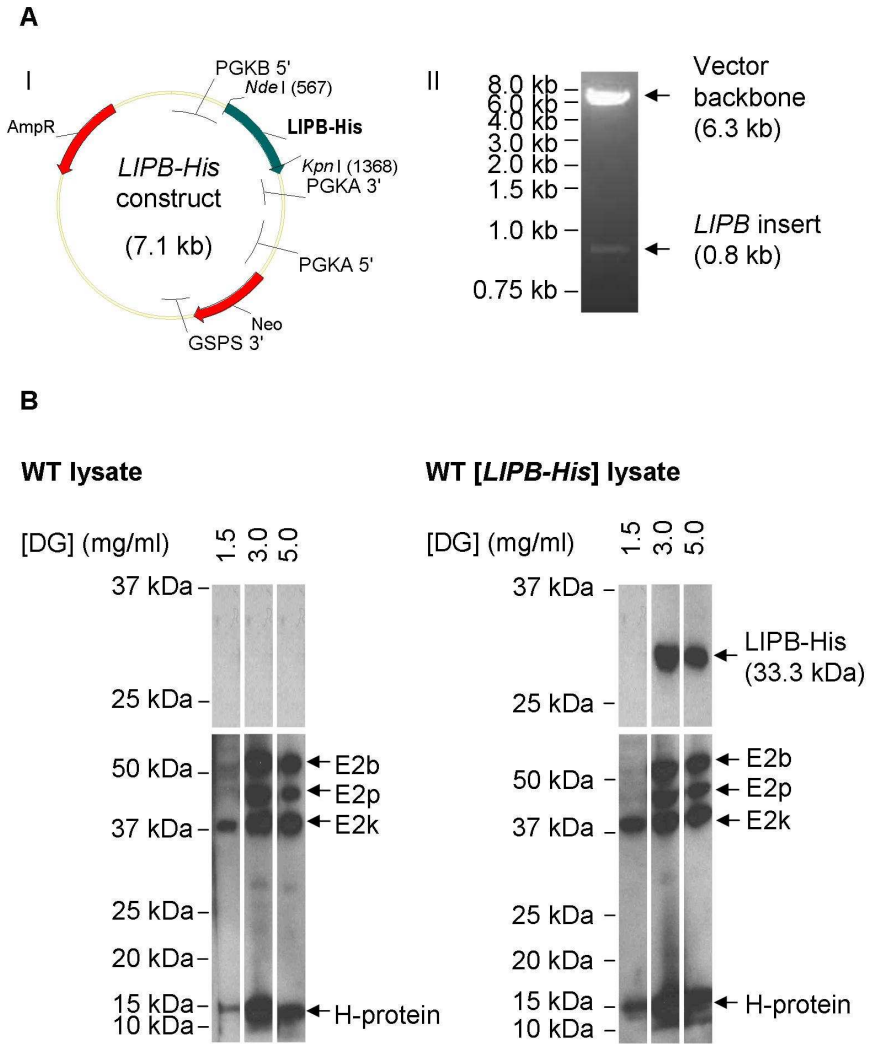


Figure 5.5 Sub-cellular localisation of LIPB-His in the WT[*LIPB-His*] line
A, The *LIPB* gene (see Section 3.4.1) was amplified from gDNA with primer pair Lm21-22 (see Section 2.1.4). The insert was cloned into expression vector pGL1137, generating construct *LIPB-His* (I). Correct cloning was verified by *Nde*I/*Kpn*I digest, releasing a 0.8 kb insert corresponding to *LIPB*, and vector backbone of 6.3kb (II). *B*, The *LIPB-His* construct was expressed in promastigotes to give rise to the WT[*LIPB-His*] line. In order to determine the sub-cellular localisation of LIPB, wild-type and WT[*LIPB-His*] lines from stationary phase were sequentially treated with increasing concentrations of digitonin ([DG]) varying from 1.5 – 5.0 mg ml^{-1} (see Section 2.2.3.3). The three protein lysate fractions for each line were subjected to SDS-PAGE and western blotting was carried out using α-*Lmj*LIPB antibody (top panel) at 1/1,000 or α-LA antibody (bottom panel) at 1/6,000. LIPB-His was detected in LIPB-His lysate and was calculated to be 33.3 kDa (see Section 2.5.5 for method of calculating molecular weights from SDS-PAGE).

LIPB is expected to be expressed in promastigotes given that LIPA was shown to be expressed (see Figure 4.9). Given that LIPB is not detected in concentrated mitochondrial protein lysate from promastigotes (see Figure 5.5B), combined with the knowledge that *E. coli* LipB is expressed at levels that are undetectable (Jordan & Cronan, 2003), an experiment was carried out to determine the concentration of LIPB-His protein in the WT[*LIPB-His*] overexpressing line. Quantities of $LIPB_{FL}$-Strep recombinant protein ranging from 0.2 – 3.2 ng and 10 μg of WT[*LIPB-His*] protein lysate were subjected to SDS-PAGE and western blotting with α-*Lmj*LIPB antibody (see Figure 5.6A). Band density was measured

using Quantity One software and a standard curve produced with band intensity as a function of quantity of $LIPB_{FL}$-Strep recombinant protein (as described in Section 2.5.5) (see Figure 5.6B). Using the standard curve, it was determined that LIPB-His protein is present at a concentration of 50 pg per μg of WT[*LIPB-His*] soluble protein lysate. This equates to approximately 27,000 copies of LIPB-His polypeptide per cell (given that there are 3.7×10^5 cells per 1 μg of soluble protein lysate). Considering that *E. coli* express < 10 copies of LipB per cell (Jordan & Cronan, 2003), and that *Lmj*LIPB is not detectable in concentrated promastigote mitochondrial protein lysate, it can be concluded that the WT[*LIPB-His*] line drastically overexpresses the LIPB-His protein relative to endogenous LIPB.

In addition to analysing band density, the molecular sizes of $LIPB_{FL}$-Strep and LIPB-His were calculated (see Section 2.5.5 for method of calculating molecular sizes from SDS-PAGE). The molecular sizes of $LIPB_{FL}$-Strep and LIPB-His are 30.8 kDa and 32.2 kDa, respectively (see Figure 5.6A). The expected molecular size of $LIPB_{FL}$-Strep is 30.7 kDa, and since the protein will not be processed for mitochondrial transit in *E. coli*, the observed 30.8 kDa band corresponds well to the predicted size. The expected molecular size of full-length LIPB-His is 32.3 kDa, and based on TargetP mitochondrial targeting prediction (see Figure 3.2), the size of LIPB-His after cleavage of the mitochondrial targeting peptide would be 28.8 kDa. The fact that the observed LIPB-His protein is 32.2 kDa indicates that the mitochondrial targeting peptide is likely to be considerably shorter than that predicted by TargetP.

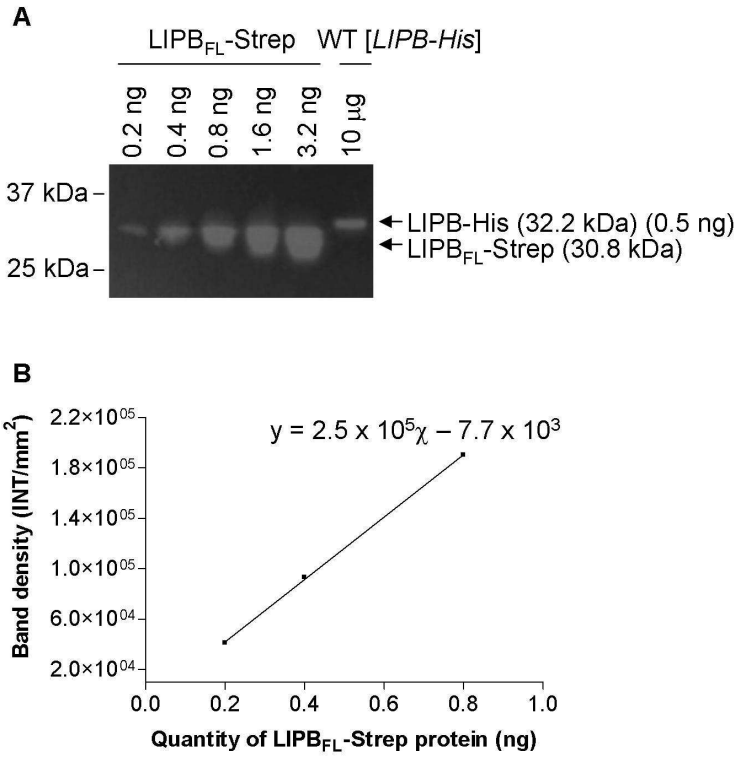


Figure 5.6 Quantification of the level of LIPB-His expression in the WT[*LIPB-His*] line
A, In order to quantify the amount of LIPB-His in the WT[*LIPB-His*] line, mid-log growth phase promastigote cells were harvested and 10 µg of soluble protein lysate subjected to SDS-PAGE. Different quantities of recombinant LIPB$_{FL}$-Strep protein were loaded on the same gel. Western blotting was carried out with α-*Lmj*LIPB antibody at 1/1,000 and chemiluminescence signal detected using a BioRad ChemiDoc XRS machine. The calculated sizes for LIPB-His and LIPB$_{FL}$-Strep are 32.2 kDa and 30.8 kDa, respectively (see Section 2.5.5 for method of calculating molecular sizes from SDS-PAGE). *B,* Band density was measured using the Quantity One software program and the linear phase was plotted on a graph. Using the linear equation, the quantity of LIPB-His protein in 10 µg of WT[*LIPB-His*] soluble protein lysate was determined to be 0.5 ng (as depicted in *A*).

5.4.2 Effects of LIPB-His overexpression on promastigotes

As discussed in Section 5.4.1, the WT[*LIPB-His*] line dramatically overexpresses the LIPB-His protein, relative to endogenous *Lmj*LIPB expression in wild-type promastigotes. The effects of LIPB-His overexpression were analysed by employing the same assays used to study the effects of LPLA-His and LPLA^{H118A}-His expression (see Section 4.4). That is, the WT[*LIPB-His*] line was analysed by determining parasite growth rate relative to those of wild-type and WT[uncut pGL1137] controls. In addition, soluble protein was isolated during logarithmic growth and stationary phase and western blotting carried out with α-His, α-LA, α-HASPB and α-*Lmj*LIPA antibodies, and α-CS antibody to verify equal loading. All parasite lines were cultured in duplicate, and each experiment was independently repeated. Repetitions yielded very similar results; a representative set of which are shown.

Firstly, whether grown on 10 μg ml^{-1} or 50 μg ml^{-1} G418 drug selection pressure, the WT[*LIPB-His*] line exhibited a growth phenotype; relative to the WT[uncut pGL1137] control, WT[*LIPB-His*] entered stationary phase at a significantly earlier density ($p<0.05$) (see Figure 5.7). Secondly, western blot analysis with α-His antibody showed that LIPB-His protein expression was not affected by culturing parasites in the presence of different concentrations of G418 (see Figure 5.8).

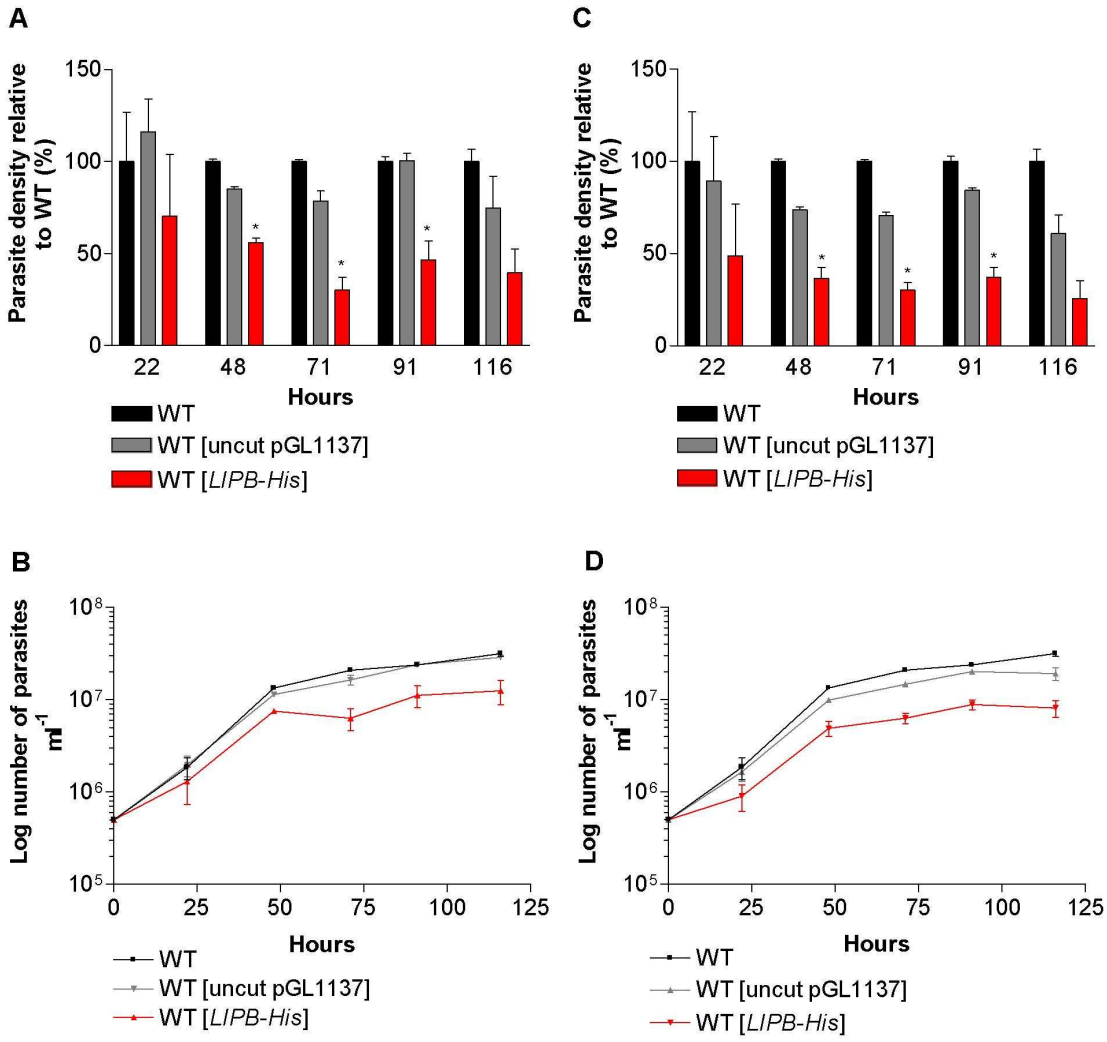


Figure 5.7 Growth of the WT[*LIPB-His*] line
Growth analysis of the LIPB-His over-expressing line was carried out by diluting all cultures to 5 x 10^5 parasites ml^{-1}, and recording parasite density circa every 24 h. All lines were cultured in duplicate. WT[uncut pGL1137] and WT[*LIPB-His*] lines were maintained on 10 μg ml^{-1} (*A* & *B*) or 50 μg ml^{-1} (*C* & *D*) G418 drug pressure. *A & C,* Average and standard deviation values were calculated for each parasite line at individual time points. Values were normalised to wild-type and plotted as a bar chart. One-way ANOVA with Tukey post-tests were performed at each time point, with uncut pGL1137 values as control. Values that are significantly different from uncut pGL1137 values (*p*-value < 0.05) are depicted with the '*' symbol. *B,* Growth of parasite lines represented by a line graph with logarithmic *Y*-axis.

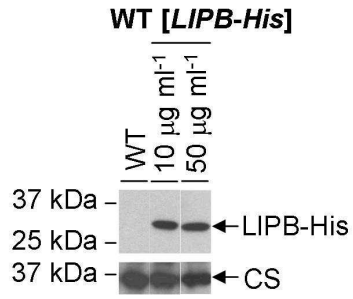


Figure 5.8 LIPB-His expression levels in the WT[*LIPB-His*] line
10 µg soluble protein samples from wild-type or the WT[*LIPB-His*] line maintained on 10 µg ml^{-1} or 50 µg ml^{-1} G418 drug pressure, from 96 h growth (see Figure 5.7), were subjected to SDS-PAGE and western blot analysis with α-HIS antibody at 1/10,000. The blot was subsequently stripped of antibody and re-probed with α-CS antibody at 1/7,500 in order to check for equal loading.

Figure 5.9 illustrates the lipoylation patterns of wild-type and the WT[*LIPB-His*] lines cultured in the presence of 50 µg ml^{-1} G418 (the lipoylation patterns do not differ in LIPB-His overexpressing lines grown in the presence of either 10 µg ml^{-1} or 50 µg ml^{-1} G418, and therefore only the latter data are shown). It should be noted that lipoylation of E2p and E2b are not as pronounced as that which is normally observed in wild-type parasites during stationary phase (72 h onwards) (see Figure 3.1). Throughout promastigote growth, lipoylation patterns in the WT[*LIPB-His*] line are very similar to those observed in the wild-type control. The only notable difference is that whereas metacyclic promastigotes in wild-type do not possess lipoylated E2p and E2b subunits, lipoylation of these acceptor proteins is apparent in WT[*LIPB-His*] protein lysate (see Figure 5.9), which is similar to the phenotype observed in the WT[*LPLA*^{H118A}*-His*] line but not in the WT[*LPLA-His*] line (see Section 4.4, Figure 4.7).

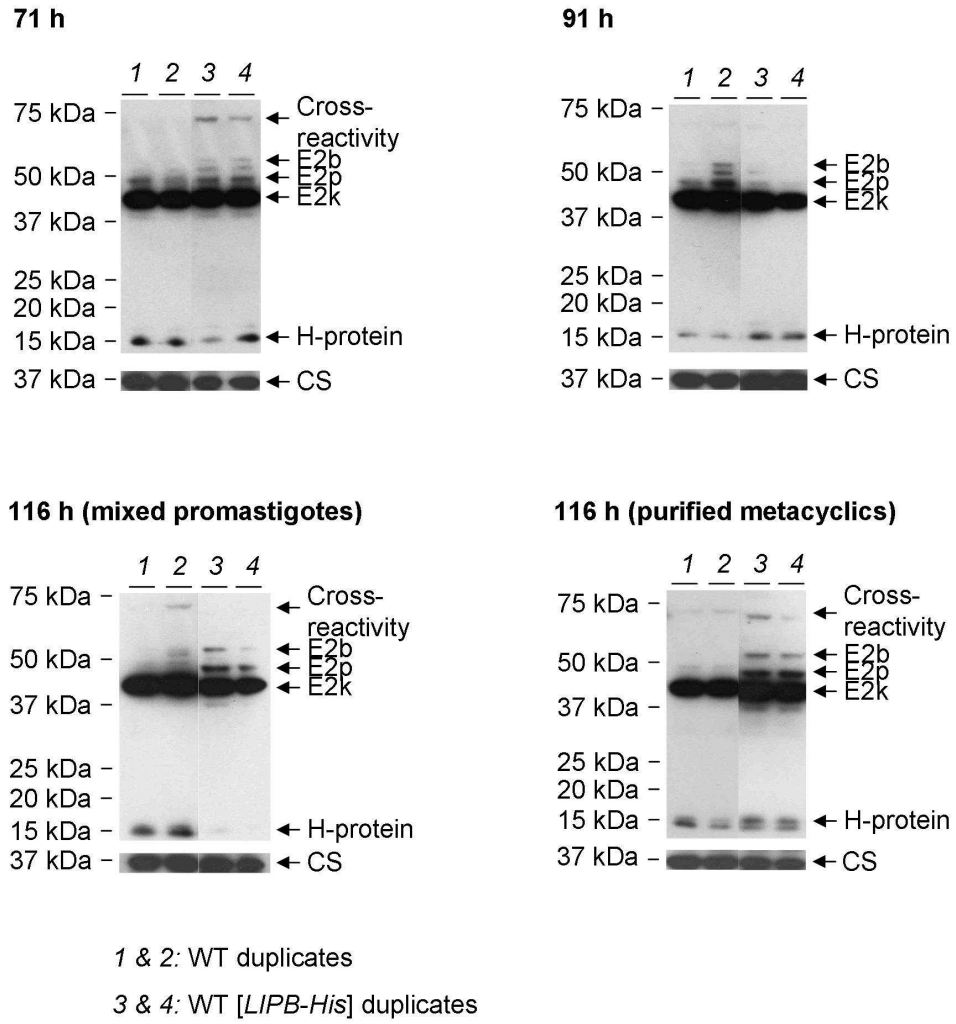


Figure 5.9 Lipoylation patterns of the WT[*LIPB-His*] line
Approximately 1 e^8 cells were lysed at four different time points from cultures maintained on 50 μg ml^{-1} G418 drug pressure. Soluble protein was harvested and subjected to SDS-PAGE. Western blots were carried out with α-LA antibody at 1/6,000 to determine lipoylation patterns, and with α-CS antibody at 1/7,500 to check for equal loading. All lines were grown in duplicate, and both replicates are tested by western blot.

In terms of the capacity to undergo metacyclogenesis as determined by the detection of the metacyclic protein marker HASPB at 116 h growth, the WT[*LIPB-His*] line grown in the presence of either concentration of G418 showed a marked decrease, indicating that metacyclogenesis is possibly impaired (see Figure 5.10). A possible explanation is that since the WT[*LIPB-His*] line enters stationary phase at an earlier time point (see Figure 5.7), and this time point is most correlated to the production of HASPB-positive metacyclic promastigotes, 116 h was not a suitable time point to probe WT[*LIPB-His*] protein lysates with α-HASPB antibody. This possibility was ruled out by western blotting with α-HASPB antibody at all time points, which revealed that levels of HASPB expression are consistently lower in WT[*LIPB-His*] lines compared to wild-type (data not shown).

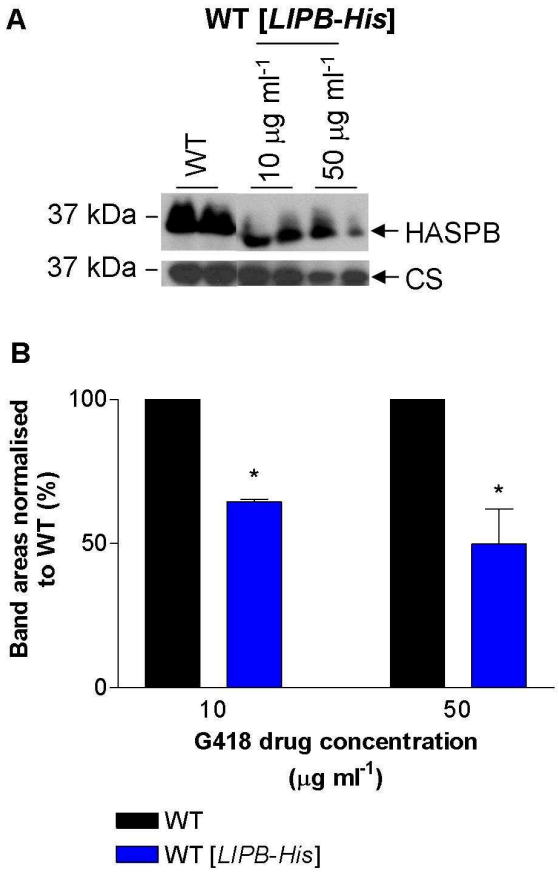
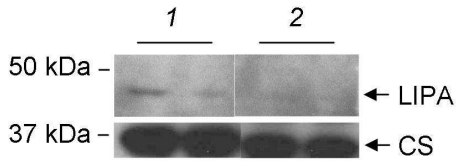


Figure 5.10 Metacyclogenesis of the WT[*LIPB-His*] line
A, 10 μg soluble protein from metacyclic promastigotes (at 116 h time point) purified from the LIPB-His overexpressing line maintained on 10 μg ml^{-1} or 50 μg ml^{-1} G418 drug pressure, was subjected to SDS-PAGE. Western blotting was subsequently carried out with α-HASPB antibody at 1/4,000 and with α-CS antibody at 1/7,500, to assess loading. All lines were grown in duplicate, and both replicates were tested by western blot. All treatments were carried out in duplicate *B,* Band areas (shown in *A*) were determined using Scion Image program, averages and standard deviations were calculated and a bar chart plotted. Subsequently, one-way ANOVA with a Tukey post-test was performed to determine whether band areas in the WT[*LIPB-His*] line were significantly different from those in wild-type, where *p*-value < 0.05 is indicated by the '*' symbol.

Given that the WT[*LIPB-His*] line demonstrates a high overexpression of LIPB-His relative to endogenous LIPB, it was hypothesised that LIPA might consequently be up-regulated also. In order to test the hypothesis, wild-type and WT[*LIPB-His*] protein lysates from metacyclic promastigotes were probed with α-*Lmj*LIPA antibody by western blotting. The reason for using this time point to analyse the levels of LIPA protein was due to the fact that the most notably difference in lipoylation pattern between the two lines is observed at this point. Figure 5.11 shows that the level of LIPA protein is not higher in the WT[*LIPB-His*] line grown on 50 μg ml^{-1} G418 drug selection. This result indicates that the expression levels of endogenous LIPA and LIPB are low, and probably tightly regulated, and that overexpression of LIPB-His does not result in a concomitant increase in LIPA.



1: WT metacyclic promastigote lysates

2: WT [*LIPB-His*] metacyclic promastigote lysates

Figure 5.11 LIPA expression levels in the WT[*LIPB-His*] line
Western blotting of protein lysates from purified metacyclic promastigotes from the 116 h time point (only lines grown on 50 μg ml^{-1} G418 selection pressure) using α-*Lmj*LIPA antibodies at 1/20 to determine LIPA expression levels in the WT[*LIPB-His*] line relative to the wild-type control. Blots were also probed with α-CS antibody at 1/7,500 to check for equal loading. All lines were grown in duplicate, and both replicates are tested by western blot.

5.5 Summary

- LIPA-GFP is a mitochondrial protein, as shown by fluorescence microscopy in which LIPA-GFP (FITC filter) and Mitotracker CMXRos (Rhodamine filter) overlay.

- *LIPA* is essential/very important to promastigote survival, since none of the 61 clones from three independent transfections lacked the gene.

- Both *LIPA* alleles can be replaced, although all clones with double allele replacement still possess at least one copy of the endogenous *LIPA* gene.

- A permissive knockout strategy using an integrating *LIPA* re-expressor construct did not lead to gene knockout in any of the 23 clones examined.

- A LIPB-His fusion protein was shown to localise to the mitochondrial fraction by a digitonin pre-fractionation experiment.

- Endogenous LIPB is expressed at sub-picomolar levels, since it could not be detected by western blotting of concentrated mitochondrial soluble protein lysate from wild-type parasites with α-*Lmj*LIPB antibody.

- Overexpression of LIPB-His resulted in the onset of stationary phase markedly earlier than in control lines.

- The lipoylation pattern in WT[*LIPB-His*] was very similar to the typical wild-type pattern, except that metacyclic promastigotes in the WT[*LIPB-His*] line lipoylated E2p and E2b in addition to E2k and H-protein.

- Overexpression of LIPB-His resulted in the isolation of a lower proportion of HASPB-positive non-procyclic promastigotes, suggesting that LIPB-His overexpression could affect metacyclogenesis.

- There is no noticeable increase in LIPA protein levels in the WT[*LIPB-His*] line relative to the wild-type control.

6 Discussion

6.1 Introduction

LA metabolism has been most extensively studied in *E. coli*, and a wealth of data have also been generated in other organisms including the bacteria *M. tuberculosis* (Ma *et al.*, 2006; Rachman *et al.*, 2006; Sassetti *et al.*, 2003) and *L. monocytogenes* (Keeney *et al.*, 2007; O'Riordan *et al.*, 2003), the archaebacterium *T. acidophilum* (Kim do *et al.*, 2005; McManus *et al.*, 2005), the mammals *H. sapiens* (Feng *et al.*, 2009a), *M. musculus* (Morikawa *et al.*, 2001; Yi & Maeda, 2005; Yi *et al.*, 2009) and *B. taurus* (Fujiwara *et al.*, 2007; Witkowski *et al.*, 2007), the plants *A. thaliana* (Ewald *et al.*, 2007), *Oryza sativa* L. (Kang *et al.*, 2007) and *P. satisvum* (Wada *et al.*, 1997), the fungus *N. crassa* (Wada *et al.*, 1997), the yeast *S. cerevisiae* (Hiltunen *et al.*, 2009) and *Kluyveromyces lactis* (Chen, 1997), and the protozoan parasites *P. falciparum* (Allary *et al.*, 2007; Gunther *et al.*, 2005; Gunther *et al.*, 2007; Gunther *et al.*, 2009a; Gunther *et al.*, 2009b; Wrenger & Muller, 2004) and *T. gondii* (Crawford *et al.*, 2006; Mazumdar *et al.*, 2006; Thomsen-Zieger *et al.*, 2003).

Some of these studies have investigated enzymes involved in LA metabolism, whereas others have researched type II FAS, and many have concluded that LA is essential for aerobic respiration, mitochondrial genome stability (in eukaryotes) and cell survival. The principal aim of this thesis was to investigate LA metabolism in *L. major*, because there is little information regarding the role of the cofactor in this or any other trypanosomatid.

The first aim of this thesis was to identify the components of multienzyme complexes that require to be lipoylated for activity, by taking an *in silico* approach. This work was carried out using the *L. major* genome, the sequence of which is available (Ivens *et al.*, 2005) in an online database (GeneDB). Subsequently, experimental evidence was gathered to show that these complexes are lipoylated *in vivo*. Furthermore, the question was asked as to which complexes are preferentially lipoylated at different stages of the *L. major* life cycle. After establishing that lipoylation of apo-E2 and apo-H-protein occurs, I set out to identify the genes involved in LA metabolism in *L. major*, and again an *in silico* approach was taken to address the question. Subsequently, the function(s) of *L.*

major gene products involved in LA metabolism were studied using several assays; namely, complementation of *E. coli* lines lacking certain genes involved in LA metabolism, observing the effect of LA analogues on LA metabolism, as well as gene replacement and protein overexpression. The results obtained will be discussed in this order below, with reference to published data.

6.2 α-KADHs and the GCC

6.2.1 *In silico* predictions

From bacteria to mammals, α-KADHs are central to energy metabolism (Cronan *et al.*, 2005; Mooney *et al.*, 2002; Perham, 2000), and the GCC is important for the provision of one-carbon (C1) units, which can be used to synthesise purines, thymidylate, serine, methionine and formylmethionyl- transfer RNA (tRNA) (Cossins & Chen, 1997; Douce *et al.*, 2001).

There are a plethora of published data with regards to the α-KADH and GCC multienzyme complexes, in terms of subunit compositions and crystal structures. These complexes are presumed to exist in *Leishmania*, based upon biochemical evidence, and the first aim of this thesis was to identify genes in the *L. major* genome encoding all of the putative subunits comprising PDH, α-KGDH, BCKDH and GCC complexes, and to compare the amino acid sequences of the *L. major* predicted proteins with the known protein sequences in *E. coli* and *H. sapiens* (see Table 3.1). It should be noted that genes encoding all of the subunits comprising the GCC in *L. major* have previously been identified (Scott *et al.*, 2008), and the role of the P-protein was evaluated by a gene disruption strategy. The results indicated that the GCC is active in *L. major*, and that the promastigote and amastigote forms lacking the P-protein exhibit a slow growth phenotype (Scott *et al.*, 2008).

Potential genes encoding all of the subunits comprising the three α-KADHs and the GCC were identified in the *L. major* genome. Sequence alignments were carried out, and most of the components were found to contain conserved residues that have been shown to be important in catalytic activity or in substrate binding. Importantly, the PDH E2 (E2p), KGDH E2 (E2k), BCKDH E2 (E2b) and H-protein from the GCC, all possess lipoyl-domains containing a strictly conserved

lysine residue that has been shown to be essential for lipoylation (see Section 3.2). All of the predicted E2 proteins also contain a strictly conserved histidine residue at the *C*-terminus, which is required for acyltransferase activity. Lastly, all of the components of α-KADHs and the GCC were predicted with high confidence by TargetP and MitoProt to possess mitochondrial-targeting peptides, which would be in line with the fact that these complexes also localise to mitochondria in yeast and human.

Given that the *in silico* results were explained in depth in Section 3.2, and that one would predict that *L. major* possesses all genes necessary to encode functional components of α-KADHs and the GCC, the aim of this section is to highlight the potentially interesting differences that were mentioned in Section 3.2.

Firstly, *L. major* has genes encoding two potential KGDH E1 (E1k) proteins (referred to as *Lmj*E1k-A and *Lmj*E1k-B), whereas *E. coli*, for example, encodes a single E1k, which forms a homodimer (Frank *et al.*, 2007) (see Section 3.2.1.1). Interestingly, recent sequence analyses have resulted in the identification of a second human E1k, which is referred to as E1k-like (E1k-L) (Bunik & Degtyarev, 2008). The E1k-L protein shares all motifs characteristic of E1k enzymes (Bunik & Degtyarev, 2008), and is the most common isomer in human heart tissue (Bunik *et al.*, 2008). In addition, it was shown that another E1k-like protein exists in human – dehydrogenase E1 and transketolase domain-containing 1 (DHTDK1) – and is more like bacterial E1k enzymes, since it lacks the three pivotal Ca^{2+}-binding motifs that are characteristic of vertebrate E1k proteins (Bunik *et al.*, 2008; Lawlis & Roche, 1981; Rutter *et al.*, 1989). *Lmj*E1k-A and *Lmj*E1k-B share only 36 % sequence identity with each other, and although they contain most of the required motifs necessary for E1k activity, such as lipoyl-domain-binding motifs and TPP-binding motifs, they lack all three conserved motifs required for binding of Ca^{2+}. In terms of known motifs involved in α-KGDH activity, it would appear unlikely that *Lmj*E1k-A and *Lmj*E1k-B complement one another, and as such *Lmj*E1k-A and *Lmj*E1k-B are more likely to be E1k isomers, rather than E1k heteromers that are dependent upon one another for full E1k function. This could have implications in terms of substrate specificity of the α-KGDH, and it is tempting to hypothesise that *L. major* possesses two E1k subunits that are able to recognise different substrates or that the two proteins are expressed in different life cycle/cell cycle

stages, although experimental evidence would be required to make any conclusions.

The second potentially interesting observation from *in silico* analyses regards the E2p subunit. Querying the *L. major* genome database with the *E. coli* protein sequence for E2p resulted in the identification of two potential E2p subunits; *Lmj*E2p and *Lmj*E2p-like (*Lmj*E2p-L) (see Section3.2.2.2). It is most likely that the annotations provided by GeneDB are correct, since *Lmj*E2p possesses the strictly conserved *C*-terminal histidine residue, whereas *Lmj*E2p-L lacks this residue, and as such is unlikely to be catalytically active. Interestingly, both *Lmj*E2p and *Lmj*E2p-L possess lipoyl-domains containing the strictly conserved lysine residue, which is the target of lipoylation. Notably, *Lmj*E2p has a single lipoyl-domain with the conserved lipoylation motif ^{60}TDKA^{63}, and *Lmj*E2p-L potentially has two lipoyl-domains with lipoylation motifs ^{80}TDKA^{83} and ^{177}TDKA^{180}. *E. coli* E2p (*Ec*E2p) and *H. sapiens* E2p (*Hs*E2p) contain three lipoyl-domains (Reed & Hackert, 1990), although it has been demonstrated in *E. coli* that only one is required for activity and that lipoylation of all three domains does not increase acyltransferase activity (Guest *et al.*, 1985).

The fact that *Lmj*E2p-L has putative lipoyl-domains and is likely to be catalytically inactive indicates that it is most likely to correspond to E3-binding protein (E3BP), which is also found in mammals (De Marcucci & Lindsay, 1985) and some yeast (Maeng *et al.*, 1994), but not in bacteria. The function of E3BP is to bind to E3, and insert E3 into the E2 scaffold, since in these organisms, E3 cannot bind E2 directly (Sanderson *et al.*, 1996a; Sanderson *et al.*, 1996b). In terms of the evolutionary justification of E3BP, one reason might be to add an extra level of control of the PDH, whereby if E3BP is not expressed (or inhibited), E3 will not be able to bind to the PDH complex, thus rendering it inactive. Interestingly, in *T. brucei* BSF, in which the mitochondrion is poorly formed and the majority of enzymes involved in the TCA cycle are not expressed, E3 has been shown to localise to the inner surface of the plasma membrane (Danson *et al.*, 1987). This might appear to be a very strange observation, however there are published data implicating E3 and LA in the *E. coli* protein-dependent transport of galactose and maltose (Richarme & Heine, 1986), as well as the involvement of dithiols in hexose transport in 3T3-L1 adipocytes (Frost & Lane, 1985). However, the physiological relevance of the observed E3 localisation to the plasma membrane in *T. brucei* has not been investigated further. Another interesting and surprising study in *L. major* concluded

that of the proteins that are potentially secreted/excreted from promastigotes, *Lmj*E2p-L is one of them (Chenik *et al.*, 2006). This is possible, given that TargetP predicts with high confidence (85 %) that *Lmj*E2p-L possesses a secretory signal sequence. Also, MitoProt predicts with low confidence (23 %) that E3 has a mitochondrial targeting peptide, although TargetP predicts with high confidence that E3 possesses a mitochondrial targeting peptide. Is it possible that E3 and/or E3BP are dually-targeted to the mitochondrion and plasma membrane, and if so how would these proteins bind to the membrane, and what would be the physiological significance? To answer these questions is out of the scope of this thesis, yet these are potentially interesting questions which have not been addressed at all in recent studies regarding LA and α-KADHs.

6.2.2 Regulation of α-KADHs by phosphorylation

Mammals encode four PDH kinases (PDK), PDK-1-4 (Roche *et al.*, 2001; Roche & Hiromasa, 2007), and one BCKDH kinase (BCKDK) (Machius *et al.*, 2001; Wynn *et al.*, 2000), which act on PDH and BCKDH, respectively, to down-regulate activity of these complexes. In addition, they also possess two PDH phosphatases (PDP), which contain the same regulatory subunit (PDPr) but differ in the catalytic subunit (PDP-1c or PDP-2c) (Huang *et al.*, 1998; Lawson *et al.*, 1993; Lawson *et al.*, 1997; Teague *et al.*, 1982), as well as a single BCKDH phosphatase (BCKDP or PTMP) (Damuni *et al.*, 1984; Damuni & Reed, 1987; Joshi *et al.*, 2007). The role of PDP and BCKDP is to catalyse the dephosphorylation of their respective enzyme complexes, and as such repress the down-regulation caused by phosphorylation of PDH and BCKDH by PDK and BCKDK, respectively.

Querying the *L. major* genome database with human PDK and BCKDK protein sequences resulted in the identification of two potential PDK/BCKDK proteins, which were annotated in GeneDB as PDH kinase (lipoamide) (*LmjF*24.0010) and a phosphoprotein-like protein (*LmjF*20.0280) (see Table 3.2). Both predicted protein sequences, corresponding to *LmjF*24.0010 and *LmjF*20.0280 genes, are highly likely to contain mitochondrial targeting peptides, as predicted by TargetP and MitoProt. The predicted protein sequence of the *LmjF*20.0280 gene contains the strictly conserved N-, G1- and G2 boxes that constitute the ATP-binding domain that is a key feature of the ATPase/kinase superfamily (Bilwes *et al.*, 1999). However, the protein completely lacks the F-box, which is present in all known

human PDK and BCKDK enzymes (Bowker-Kinley & Popov, 1999; Wynn *et al.*, 2000). The *LmjF*24.0010 predicted protein lacks strictly conserved residues within the N-, G1- and G2 boxes, and as such is unlikely to correspond to PDK, contrary to the annotation found in GeneDB. Lastly, both *LmjF*24.0010 and *LmjF*20.0280 predicted proteins lack the *C*-terminal domain found in PDK, which is involved in binding the second lipoyl-domain (L2) of E2p (Kato *et al.*, 2005). Nevertheless, BCKDP in humans also lacks this *C*-terminal portion. In conclusion, of the two predicted PDK/BCKDK enzymes in the *L. major* genome database, it is not possible to conclude with any confidence that either acts to phosphorylate either the PDH or BCKDH. However, based on sequence comparisons with human PDK and BCKDK, it is more likely that *LmjF*20.0280 has the capacity to phosphorylate proteins, and experimental data would be required to show this, as well as the specificity of this kinase towards either (or both) of PDH or BCKDH.

If *L. major* possesses a kinase(s) to down-regulate certain α-KADHs, it would also be necessary to encode a phosphatase(s) that can release this inhibition by dephosphorylation of the α-KADH. Querying the *L. major* genome database with human PDP subunits and with BCKDP resulted in the identification of seven putative phosphatase genes. The seven corresponding predicted proteins have a high degree of conservation within motifs known to be important in members of the PPM family (Barford *et al.*, 1998; Bork *et al.*, 1996), which includes mammalian PDP and BCKDP. Of the seven predicted proteins, only two – *LmjF*30.0380 and *LmjF*32.1690 – are predicted with high confidence by TargetP and MitoProt to contain mitochondrial targeting peptides (see Section 3.2.6.2). Given that PDH and BCKDH are likely to be mitochondrial, these two predicted proteins are the best two candidates for being PDP/BCKDP.

Overall, sequence analyses did not give a clear answer as to whether *L. major* is likely to possess genes encoding kinase(s) and phosphatase(s) that are known in mammals to regulate PDH and BCKDH activities. It has been possible however to narrow down the potential list of genes that could play this role in *L. major*, and experimental evidence in the form of biochemical characterisation would be required to verify whether any of these predicted proteins possess kinase or phosphatase activity, and if they do, to what extent they possess α-KADH-specificity.

6.2.3 Lipoylation as a mechanism of regulating α-KADHs and the GCC

Bacteria do not possess kinases and phosphatases to regulate activity of α-KADHs, and also, a kinase (s) and phosphatase (s) that regulates α-KADH activity has not been documented. Also, given that mitochondrial proteins typically turn over slowly and the half-life of the PDH in rat is one week (Weinberg & Utter, 1980), there must be a post-translational mechanism other than phosphorylation, to regulate α-KADHs. Given that LA is pivotal to activity of α-KADHs, it is likely that controlling the relative levels of lipoylation of different α-KADHs is an important mechanism to regulate their activities. Indeed, it has been shown in human HEK293 T cells that lipoyl-moieties are rapidly turned over (Feng *et al.*, 2009a). Work in human (Hayakawa & Oizumi, 1988; Oizumi & Hayakawa, 1989), rat, *Lactobacillus casei* (Shirota) (Hayakawa *et al.*, 2006) and *Enterococcus faecalis* (Jiang & Cronan, 2005) has indicated the presence of an amidase enzyme called lipoamidase, which catalyses the removal of LA from E2 subunits and H-protein. It is interesting to note that querying the *L. major* genome database with the *E. faecalis* lipoamidase protein sequence resulted in the identification of one potential homologue; *LmjF*16.1360. In the *L. major* genome database, this gene is annotated as a hypothetical gene.

There are surprisingly few published data characterising the differential lipoylation intensities of the three α-KADHs and H-protein of the GCC under different stages of cell development or in different environmental conditions. The aim of this section of my thesis was to determine whether lipoylation patterns of the different α-KADHs and GCC change during promastigote growth *in vitro*, and in the amastigote stage, in order to establish whether the analysis of lipoylation patterns could be a useful diagnostic tool for studying metabolic responses.

In order to address this question, western blot analysis was carried out on soluble protein extracts from *L. major* using a polyclonal antibody that recognises protein-bound LA. This antibody has been used in previous studies in *L. monocytogenes* (O'Riordan *et al.*, 2003), *S. cerevisiae* (Schonauer *et al.*, 2008), *A. thaliana* (Ewald *et al.*, 2007), *P. falciparum* (Allary *et al.*, 2007; Gunther *et al.*, 2007), *T. gondii* (Crawford *et al.*, 2006), *T. bruce*i (Stephens *et al.*, 2007) and *H. sapiens* (Feng *et al.*, 2009a). In *L. major*, a total of four lipoylated protein species were identified.

Based on predicted sizes after cleavage of the mitochondrial targeting peptide by MitoProt and TargetP, three of the bands were designated as H-protein, E2k, E2p and E2b (see Section 3.3). These predictions are not perfect however. For example, the molecular mass of the predicted full-length E2b protein is 50.2 kDa, and yet the observed molecular mass of the putative E2b is 55 kDa (see Table 3.4). Also, E2k and E2p-L proteins were predicted to be a similar size after cleavage of the mitochondrial targeting peptide. It is much more likely however that the observed band of molecular mass 39.1 – 41.4 kDa is E2k and not E2p-L, since E2p-L is likely to be catalytically inactive and correspond to E3BP (see Section 6.2.1), and there have been no published studies using the α-LA antibody that have identified lipoylated E3BP, despite it possessing a lipoyl-domain. One possibility is that both proteins are lipoylated, and that the prominent band designated as E2k in western blots probed with α-LA antibody (see Figure 3.1) is masking a band of a similar size that could represent E2p-L. However, even when α-LA western blot films were exposed for less time to minimise the intensity of the putative E2k band, a second band was never observed. Nevertheless, in order to verify the predictions made above, it would be necessary to immunoprecipitate lipoylated proteins from *L. major* promastigote lysate, and to subsequently carry out mass spectrometry analysis. Please note that this is acknowledged as a limitation, however from henceforth the four lipoylated proteins will be referred to as H-protein, E2k (and not E2p-L), E2p and E2b, and potential reasons will be given to justify their differential lipoylation with regards to parasite energy metabolism.

The results generated by western blotting with α-LA antibody illustrated that lipoylation patterns change throughout *L. major* promastigote growth *in vitro*, and that the changes in lipoylation patterns are consistent from one sub-passage to the next. Unfortunately, I did not have antibodies recognising E2k, E2k, E2b or H-protein, and those raised against *P. falciparum* recombinant E2 proteins did not cross-react with *L. major* soluble protein extract (data not shown). As such, it was not possible to correlate lipoylation of E2 subunits with the levels of E2 protein, although the half-life of the PDH, for example, has been shown to be one week in rat (Weinberg & Utter, 1980). Nevertheless, since lipoylation is required for α-KADH activity, it is possible to make putative conclusions with regards to the relative activities of these different protein complexes at any time point.

My results highlight several interesting and notable differences in terms of lipoylation intensities in promastigotes. Firstly, the E2k is constitutively lipoylated at all time points studied, from early-log growth to late stationary phase (see Figure 3.1). α-KGDH catalyses the oxidative decarboxylation of α-ketoglutarate, one source of which is proline (another sources is glutamine) (see Figure 1.4). Given the importance of proline as a carbon source for energy production in the sandfly vector (Bringaud *et al.*, 2006), combined with the fact that *Leishmania* promastigotes grown *in vitro* preferentially use amino acids as an energy source (Cazzulo *et al.*, 1985) (see Section 1.2.2), could explain why E2k is highly lipoylated throughout promastigote *in vitro* culture.

The E2p and E2b subunits are lipoylated during late-log growth and stationary phase, although never to the extent to which E2k is lipoylated (see Figure 3.1). However, E2p and E2b are not lipoylated in mid-log phase. These results could indicate that proline/glutamine is preferentially used as a carbon source, and once proline/glutamine becomes depleted (although this was not directly shown), the parasites start using additional energy sources. Up-regulation of PDH would allow pyruvate from glycolysis to be oxidatively decarboxylated to acetyl-CoA, which would most likely be converted to acetate via the ASCT, instead of being fed into the TCA cycle (see Section 1.2.3 & Figure 1.4) (Van Hellemond *et al.*, 1998). There have not been any studies carried out with regards to the capacity of *Leishmania* to use branched-chain amino acids (BCAAs) as a carbon source, and these results are the first to show that a putative E2b is lipoylated, and that the BCKDH is thus likely to be active. BCAAs are converted to branched-chain ketoacids (BCKAs) by a branched-chain aminotransferase (BCAT), which is predicted to be present in the *L. major* genome database (*LmjF*27.2030). Genome data support the fact that *Leishmania* cannot synthesise branched-chain amino acids (Ivens *et al.*, 2005), and the fact that they are used as an energy source could indicate that the parasites make use of exogenously-supplied branched-chain amino acids as carbon sources. Indeed, in mammals, accumulating data suggests that BCKDH activity increases with exercise, as does oxidation of the BCAA leucine (Cynober & Harris, 2006; Norton & Layman, 2006). Although the midgut of the sandfly is likely to be sugar-rich, there is probably a limited supply of sugars as the parasites migrate towards the proboscis, and based on these results regarding lipoylation patterns, it could be that *Leishmania* are very adaptable to the carbon sources that they use to acquire energy. Interestingly, although mixed

culture promastigotes in late stationary phase lipoylate all three α-KADH E2 subunits, metacyclic promastigotes purified from the same culture flask only lipoylate E2k (see Figure 3.1). It is worth noting that the method used to isolate metacyclic promastigotes is a negative selection procedure, whereby peanut lectin binds to a specific plasma membrane-bound protein found on procyclic promastigotes (which is not found on the plasma membrane of metacyclic promastigotes), and as such procyclic promastigotes can be pelleted by low-speed centrifugation, whilst metacyclic promastigotes remain in the supernatant fraction. The metacyclic promastigote is truly defined by the ability to infect human/mouse macrophages, and this assay was not carried out in conjunction with lipoylation experiments, nevertheless the term 'metacyclic promastigote' will be used henceforth for simplicity. Metacyclic promastigotes are non-dividing. The fact that this sub-population of parasites do not lipoylate PDH or BCKDH indicates that they do not use acetyl-CoA (derived from pyruvate in glycolysis) or BCAAs as an energy source, and must therefore rely upon glycolysis and amino acid oxidation. A potential explanation for this observation is that metacyclic promastigotes do not expend as much energy as dividing procyclic promastigotes, and as such down-regulate PDH and BCKDH activities. This is intriguing because it potentially implies that lipoylation is a dynamic event, since 24 h beforehand, E2p and E2b were lipoylated (although it cannot be ruled out that the E2p and E2b proteins themselves are degraded). As mentioned already, *L. major* potentially encodes a lipoamidase which could cleave the amide bond between the lipoyl-moiety and the E2 subunit or H-protein. This could add an extra level of control to lipoylation, and it would be interesting to overexpress this protein, for example, to show whether it results in a decrease in lipoylation of the α-KADH and GCC.

It would be interesting in future studies to alter the proline/glutamine/glucose/threonine concentrations in the growth medium, and to determine how this affects lipoylation patterns, and whether the two can be correlated. If the lipoylation pattern would be diagnostic of the relative activities of different α-KADHs, analysing lipoylation patterns would represent a simple and effective method to obtain a global picture of metabolic switching of cells treated under different nutritional conditions. For example, detailed molecular studies were recently carried out in PCF *T. brucei* which illustrated that these parasites have the capacity to rapidly switch their metabolism in response to the carbon sources

available (Coustou *et al.*, 2008); it would be interesting to determine whether alterations in lipoylation patterns correlate with such metabolic switches.

In promastigotes, lipoylation of the H-protein was observed in a pattern that mirrored E2p and E2b lipoylation patterns; H-protein was lipoylated during late-log growth and stationary phase, yet not during mid-log growth or in metacyclic promastigotes. A homozygous disruption of the *L. major* gene encoding the P-protein (see Section 3.2.4.2) of the GCC was not lethal in the promastigote stage, as long as serine was present in the medium (Scott *et al.*, 2008). The authors of this work hypothesise that an enzyme called serine hydroxymethyltransferase (SHMT) is able to compensate for the loss of the GCC, since SHMT catalyses the reversible conversion of serine and THF (tetrahydrofolate) to glycine and 5,10-CH_2-THF (methylene tetrahydrofolate). Two isoforms of SHMT are found in *L. major* promastigotes that localise to the cytosol and to the mitochondrion (Gagnon *et al.*, 2006). The main role of the GCC is to provide C1 units in the form of 5,10-CH_2-THF, which are used by folate coenzymes to synthesise essential cellular compounds, including pyrimidines. Since metacyclic promastigotes are non-dividing, one would hypothesise that these parasites do not require as much pyrimidine for DNA synthesis as do procyclic promastigotes, which could partially explain why H-protein is not lipoylated in metacyclic promastigotes. Also, given that there is redundancy between the GCC and SHMT, perhaps only the SHMT need be active in metacyclic promastigotes.

In order to analyse the lipoylation pattern in the mammal-infective form, amastigotes were extracted from a mouse back lesion by dissecting out the lesion, crushing the tissue to release amastigotes, centrifuging the sample and then collecting the amastigote-containing supernatant fraction (work carried out by Denise Candlish, University of Glasgow). This is a crude method of cell isolation, and the amastigote-containing fraction inevitably contains host cell proteins. Western blotting of amastigote protein lysate with α-LA antibody revealed lipoylated proteins that potentially correspond to *L. major*-specific H-protein, E2k and E2p. However, lipoylation of the 55 kDa E2b band was not apparent in amastigotes. Interestingly, a lipoylated protein was observed that is of a very similar molecular mass to E2k, and I hypothesise that this could represent E2p-L, although it could be a cross-reacting mouse E2 subunit. The published molecular mass of mouse E2p after cleavage of the mitochondrial targeting peptide is 67 kDa, and no such band was observed in the α-LA western blot. However, it cannot

be ruled out that lipoylated mouse E2p or E2b proteins do not contaminate the amastigote sample preparation, and that the bands observed at 41.4 kDa and 47.8 kDa could represent these contaminating proteins. There are no published data regarding the molecular masses of mouse E2k and E2b after removal of the mitochondrial targeting peptide. Overall, it is more difficult to make any conclusions with regards to the lipoylation status of *L. major* amastigotes, and since *L. major* promastigotes cannot be differentiated into axenic amastigotes, this area remains elusive.

6.3 Enzymes involved in LA metabolism

6.3.1 *In silico* predictions

Having established that α-KADHs and the GCC are lipoylated in *L. major*, the next step was to identify the enzymes involved in LA metabolism. In all organisms, LA can be synthesised from the type II FAS product octanoyl-ACP by LipB and LipA (in *E. coli* and mammals), and/or LA can be salvaged by LplA (in *E. coli*) or by ACSM1 and LT (in mammals).

Crystal structure data on *Mt*LipB (Ma *et al.*, 2006) and biochemical analysis of *Ec*LipB (Zhao *et al.*, 2005) have shown that the LipB-catalysed reaction proceeds through an acyl-LipB intermediate, in which the octanoyl-moiety forms a thioester bond with the thiol of Cys^{169} (*Ec*LipB numbering). This residue is only found in LipB enzymes, and not in LplA, LT or BPL. The reason for this is because LplA, LT and BPL-catalysed reactions proceed through a non-covalently bound acyl-adenylate intermediate, whereas LipB forms a covalently bound acyl thioester intermediate, which requires the thiol group of Cys^{169}. LipB, LplA, LT and BPL enzymes share one common residue, Lys^{135} (*Ec*LipB numbering). In all of these enzymes, this residue has been proposed to be important in the transfer of the lipoyl-AMP/octanoyl intermediate to the target lysine residue on the lipoyl-domain of an apo-E2 or apo-H-protein. In LipB, Lys^{135} (*Ec*LipB numbering) is essential for formation of the octanoyl-LipB intermediate (Ma *et al.*, 2006), potentially by activating the octanoyl-moiety for thioester formation with Cys^{169} (*Ec*LipB numbering) (Kim do *et al.*, 2008; Ma *et al.*, 2006).

Putative genes encoding *Lmj*LIPB and *Lmj*LPLA were identified in the *L. major* genome database (see Section 3.4). *LmjF*36.3080 is the most likely candidate for *LmjLIPB* since the predicted gene product possesses the strictly conserved cysteine residue that defines LipB enzymes (Cys^{191} in *Lmj*LIPB). *LmjF*07.1060 is most likely to be *LPLA*, since the *N*-terminal domain of the predicted gene product has all of the key motifs that are essential for binding LA, ATP and lipoyl-AMP (see Section 3.4.3). Recent structural data indicate that despite LPLA and LT being similar in terms of amino acid sequence identity within the *N*-terminal domain, their *C*-terminal domains differ at the structural level, which has very important implications in their catalytic qualities. For example, *Bt*LT and *Hs*LT contain the motifs that are characteristic of LPLA, however are unable to activate LA due to steric hindrance of ATP access to the LA/ATP/lipoyl-AMP binding site, which is due to the fact that the *C*-terminal domain of LT is rotated 180 ℃ relative t o that of LplA (Fujiwara *et al.*, 2007). As such, although *LmjF*36.3080 has been named *Lmj*LPLA, it is possible that this gene in fact encodes LT and not LPLA. The *C*-terminal domains of LplA proteins from different organisms share little sequence homology. However, it is interesting to note the size difference of *Lmj*LPLA compared to LplA proteins in other organisms; *Lmj*LPLA has a *C*-terminal domain that is 137 amino acids longer than that of *Ec*LPLA, and 127 amino acids longer than that of *Hs*LT (see Figure 3.5). In fact, this anomalously long *C*-terminal domain is not restricted to *Lmj*LPLA, but is a common feature amongst all trypanosomatid putative LPLA proteins. The reason for possessing a significantly longer *C*-terminal domain is not known, and was not investigated during this thesis.

A gene (*LmjjF*19.0350) that potentially encodes *Lmj*LIPA was also identified in the *L. major* genome database. This predicted protein contains two motifs – $^{133}Cx_4Cx_5C^{144}$ and $^{165}Cx_3Cx_2C^{172}$ (see Figure 3.3) – that are known to be essential for the LipA-catalysed insertion of two sulphur atoms into the octanoyl-moiety of octanoyl-E2/octanoyl-H-protein (Cicchillo *et al.*, 2004a; Cicchillo *et al.*, 2004b; Cicchillo & Booker, 2005; Douglas *et al.*, 2006).

6.3.2 Localisation of lipoylating enzymes

All three predicted proteins, LIPB, LIPA and LPLA, were predicted with high confidence to contain mitochondrial targeting peptides by TargetP and MitoProt (see Section 3.4). Two approaches were taken in order to verify the subcellular

localisation of these proteins. Firstly, full-length LIPA-GFP and LPLA-GFP reporter constructs were expressed in promastigotes, giving rise to the respective lines WT[*LIPA-GFP*] and WT[*LPLA-GFP*], and live fluorescence microscopy was carried out using Mitotracker CMXRos as a marker for mitochondrial staining. In both lines, the green fluorescent fusion protein co-localised with Mitotracker CMXRos staining, indicating that LIPA-GFP and LPLA-GFP are mitochondrial. It is likely therefore that endogenous LIPA and LPLA proteins are mitochondrial.

The second approach involved the concentration of various organelles by treatment with digitonin, and subsequently probing the soluble protein from these different fractions with α-*Lmj*LIPB antibody. In the LIPB-His overexpressing line WT[*LIPB-His*], LIPB-His protein was found primarily in fractions that coincided with the highest levels of lipoylated protein. Given that α-KADHs and the GCC are predicted with high confidence to be mitochondrial, combined with the fact that these complexes are mitochondrial in all other eukaryotes, this was taken as good evidence that LIPB-His is mitochondrial. However, endogenous LIPB was not detected in any of the fractions when probed with α-*Lmj*LIPB antibody. This was the expected result, given that endogenous LipB protein detection by western blotting has not been documented in any other organism due to its very low expression levels (Jordan & Cronan, 2003; Vaisvila *et al.*, 2000).

Overall, I experimentally illustrated that the predictions of LIPB, LIPA and LPLA to be mitochondrial were correct. The next question to answer would be 'are these proteins functional?'

6.3.3 Functionality of lipoylating enzymes

6.3.3.1Bacterial complementation

In order to determine whether the *L. major* genes encoding putative enzymes involved in LA metabolism were functional, a complementation approach was taken. Two *E. coli* insertion mutant lines were used, KER176 and KER184, which contain inactive *lipA* and *lipB* genes respectively. These bacteria do not grow on minimal medium, unless supplemented with either LA or succinate and acetate (Vanden Boom *et al.*, 1991). However, the growth can be rescued to almost normal levels when complementing heterologous genes are expressed in these bacterial mutants.

My results indicate that *N*-terminally truncated versions of *L. major* LIPA-Strep proteins complement the growth defect of KER176 in the absence of LA, indicating that *L. major* encodes a functional LIPA protein (see Section 3.5.1.2). In addition, full-length LIPB-Strep complemented the growth defect of KER184 in the absence of LA (see Section 3.5.3.2), and thus it is possible to conclude that both enzymes comprising the LA biosynthesis pathway are likely to be functional in *L. major*.

When interpreting the complementation results of KER184 *E. coli*, it is important to appreciate that this line is 'leaky', because in the absence of LA a small amount of growth is still apparent. The reason for this residual colony formation is that LplA can catalyse the ligation of the octanoyl-moiety from octanoyl-ACP to apoproteins, albeit much less efficiently than LipB (Morris *et al.*, 1995). Only one of the two *N*-terminal truncated versions of LPLA-Strep, $LPLA_{T1}$-Strep, rescued the growth of KER184 in the absence of LA (see Section 3.5.3.2). Interestingly, in the presence of exogenous LA, the expression of $LPLA_{T1}$-Strep resulted in colony formation, whereas expression of $LPLA_{T2}$-Strep resulted in a decrease in colony formation relative to the negative control (empty pASK-IBA3 vector) (see Section 3.5.3.2). This phenomenon has been observed before, whereby expression of the *L. monocytogenes* LplA1 protein in an *E. coli* line lacking *lipA* and *lplA* (KER131) resulted in a reduction in colony formation in the presence of succinate and acetate, when compared to the empty vector negative control (Keeney *et al.*, 2007). The authors did not comment on this result (Keeney *et al.*, 2007), however one would interpret the result as signifying that expression of the recombinant *L. monocytogenes* LplA1 in KER131 *E. coli* has a negative impact upon cell growth in the bypass medium. The LPLA complementation results obtained in my thesis indicate that not only does $LPLA_{T2}$-Strep not support growth of KER184, but also causes a decrease in cell growth. The difference between $LPLA_{T1}$-Strep and $LPLA_{T2}$-Strep is that the former has an *N*-terminal truncation of 12 amino acids, whereas the latter has a 36 amino acid *N*-terminal truncation. One possible explanation for the no-growth phenotype is that the shortened *N*-terminus of $LPLA_{T2}$-Strep renders this protein catalytically inactive, yet still permits the protein to bind LA. This sequestration of LA could have a dominant negative effect, whereby the inactive $LPLA_{T2}$-Strep would compete for LA with the endogenous *E. coli* LplA. Indeed, purified recombinant $LPLA_{T2}$-Strep was never found to be active when using *P. falciparum* E2 subunits and H-protein as apoproteins in an *in vitro* lipoylation assay, and western blotting with α-LA antibody showed that instead of

lipoylating apoproteins, instead $LPLA_{T2}$-Strep was itself strongly lipoylated (data not shown).

6.3.3.2Effect of LA analogues on parasite growth

In the presence of 30-50 ng ml^{-1} of a LA analogue, selenolipoic acid (SeLA), *E. coli* growth is completely inhibited (Reed *et al.*, 1994). The isolation of a spontaneous mutant that was resistant to SeLA at concentrations up to 5 μg ml^{-1} (Reed *et al.*, 1994) resulted in the identification of the *lplA* gene, and a single point mutation in the LplA protein was attributed to the SeLA resistance phenotype (Morris *et al.*, 1995). Interestingly, a second mutant was also identified that was resistant to SeLA and which could proliferate in the absence of exogenous LA, even after ablation of the *lipA* gene (Morris *et al.*, 1995). Later studies revealed that this mutant had a duplication of the segment of the *E. coli* chromosome containing the *lipA* and *lipB* genes (Jordan & Cronan, 2002). Therefore, resistance to SeLA in *E. coli* can be attained by relying upon LA biosynthesis, or by reducing LplA activity. Since these initial studies, experiments have been carried out using a similar analogue, 8-BOA, with the aim of discerning the relative importance of LplA in LA metabolism in *T. gondii* (Crawford *et al.*, 2006) and *P. falciparum* (Allary *et al.*, 2007). Another analogue that has been used to study the role of LplA is OA (Allary *et al.*, 2007; Crawford *et al.*, 2006; Jordan & Cronan, 2003). Free OA is not used by *E. coli* LipB as substrate for acylation of apo-E2 and apo-H-protein (Jordan & Cronan, 2003), since ligation of OA to lipoyl-domains necessitates the thioester octanoylation of LipB by OA-ACP (Jordan & Cronan, 2003; Zhao *et al.*, 2005). However, *E. coli* LplA can octanoylate lipoyl-domains using either free OA (in the presence of ATP) or OA-ACP, albeit with a significantly reduced efficiency compared to the classical lipoylation reaction involving LA and ATP (Jordan & Cronan, 2003). Therefore, LplA is the only enzyme capable of catalyzing the octanoylation or 8'-bromooctanoylation of E2 and H-protein apoproteins (Allary *et al.*, 2007; Crawford *et al.*, 2006; Jordan & Cronan, 2003). As such, any effects resulting from OA or 8-BOA treatment (at physiologically-relevant concentrations) are attributable to the low efficiency action of LplA.

8-BOA and OA were used in this thesis in order to understand the importance of LPLA in the survival of promastigotes. In accordance with previous studies, the parasites used in this experiment were grown in lipid-depleted medium in order to minimise competition of LA found in FBS, with exogenously added 8-BOA as a

LPLA substrate. Specifically, promastigotes were grown for two sub-passages in lipid-depleted medium before starting the alamar blue assay (see Section 3.6), and any attempt to culture promastigotes for subsequent sub-passages in lipid-depleted medium resulted in a reduction of cell growth and by sub-passage five in lipid-depleted medium, growth was no longer apparent (data not shown). My results indicate that the IC_{50} for 8-BOA is five to tenfold higher in *L. major* promastigotes than it is in *T. gondii* and *P. falciparum*, and the IC_{50} of OA is similar to that in *T. gondii*. Interestingly, whereas supplementation of 1-2 μM LA in media containing the IC_{50} dose of 8-BOA or OA completely restored growth in *T. gondii* and *P. falciparum*, addition of 1 μM LA to the media did not positively affect *L. major* promastigotes.

A possible explanation for the inter-species differences observed could be due to the differential compartmentalisation of α-KADH complexes and GCC in *T. gondii* (and *P. falciparum*) and *L. major*. Biosynthesis enzymes LipA and LipB lipoylate the sole PDH located in the apicoplast (Gunther *et al.*, 2005; Gunther *et al.*, 2007; Gunther *et al.*, 2009b; Wrenger & Muller, 2004). Salvage enzyme LplA is mitochondrial and is required for lipoylation of E2k, E2b or H-protein (Allary *et al.*, 2007; Crawford *et al.*, 2006; Gunther *et al.*, 2009b). Therefore, treatment with 8-BOA and OA specifically impedes LplA-dependent α-KGDH, BCKDH and GCC complexes within the mitochondrion of *T. gondii* and *P. falciparum*. Lipoylation of the apicoplast PDH probably remains unaffected, since LipB does not use 8-BOA as a substrate (Crawford *et al.*, 2006). In *L. major*, it is highly likely that all three α-KADHs and the GCC are located in the same organelle (see Sections 6.2.1 and 5.4.1). Therefore, if one hypothesised that *L. major* promastigotes grown in lipid-depleted medium rely solely upon biosynthesis of LA, one would expect they be resistant to higher concentrations of LA analogues. Indeed, western blotting with α-LA antibody indicates that lipoylation patterns are not altered in the presence of 10 μM 8-BOA. Unlike in *T. gondii* and *P. falciparum*, 1 μM LA was not sufficient to rescue the growth defects caused by 8-BOA or OA, indicating that the inhibitory effects on *L. major* at high concentrations of 8-BOA and OA are probably attributable to interactions of this compound with other pathways operating in *L. major* and therefore cannot be alleviated by addition of LA. An interesting follow-up experiment would be to determine the expression levels of LA biosynthesis and LA salvage enzymes when promastigotes are incubated with different concentrations of LA analogues. If the above theory is correct, LA biosynthesis enzymes would be

up-regulated in lipid-depleted medium and LPLA down-regulated, and addition of OA or 8-BOA should not affect LIPB or LIPA expression levels under these conditions.

6.3.3.3Gene replacement in *L. major*

The results obtained from LA analogue experiments indicate that in *L. major* promastigotes, LA biosynthesis and salvage pathways might be redundant. In order to test this possibility further, a gene replacement strategy was taken, with the hypothesis that it would be possible to replace genes encoding either *LIPA* or *LPLA*. The results were very surprising, since every attempt to replace either gene in *L. major* promastigotes resulted in the duplication of the target gene by genome amplification. *Leishmania* are renowned for their genomic plasticity (Cruz *et al.*, 1993), and it is known that although *Leishmania* are mostly diploid, some chromosomes are aneuploid, such as chromosome 1 in *L. major* (Martinez-Calvillo *et al.*, 2005; Sunkin *et al.*, 2000). Also, it has been established in attempted knockout studies of *L. major* dihydrofolate reductase-thymidylate synthase (*DHFR-TS*) (Cruz *et al.*, 1993), a hypothetical gene (*LmjF01.0750*) (Martinez-Calvillo *et al.*, 2005), *L. mexicana* cdc2-related kinase 1 (*CRK1*) (Mottram *et al.*, 1996) and *L. tarentolae* J-binding protein 1 (*JBP1*) (Genest *et al.*, 2005) that targeting of essential genes in *Leishmania* frequently results in aneuploidy for the chromosome in question, tetraploidy, or the creation of ectopic DNA fragments called amplicons (Feng *et al.*, 2009b; Genest *et al.*, 2005).

The efficiency of correct targeting of the *LIPA* and *LPLA* genes was high, which rules out the possibility of either gene occupying a non-recombinogenic region of chromosomal DNA. Indeed, after integration of a stable re-expressing copy of *LPLA* into the *18S SSU RNA* locus it was possible to obtain a *LPLA* null mutant, which suggests that the endogenous *LPLA* gene is essential in promastigotes. However, despite the expression of an ectopic copy of *LPLA*, the growth of the *LPLA* null mutant is significantly slower than that of wild-type at all time points on the growth curve. Unfortunately, the immunogenic response raised against recombinant LPLA protein in rabbit was poor, and as such the α-*Lmj*LPLA polyclonal antibody generated was not useful in western blot analyses, despite multiple attempts to purify and concentrate the antibody (data not shown). It is therefore not possible to make any conclusions with regards to the expression level of LPLA under normal conditions or in the null mutant containing an ectopic

copy of *LPLA*. In terms of *LIPA*, it was not possible to generate a null mutant, even in the presence of an integrated ectopic copy of the gene. It is likely that *LIPA* is essential, however it is not possible to make as firm a conclusion as it is for *LPLA*.

The obvious question is, why would both LIPA and LPLA proteins be essential for survival of *L. major* promastigotes? In *E. coli*, LA biosynthesis and salvage pathways are redundant (Morris *et al.*, 1995), and based on the results gained from the effect of LA analogue experiments in this thesis, I expected the same to be true in *L. major*. The first possible explanation is that in *L. major* promastigotes, LIPA and LPLA are important but not essential, and that the inability to successfully replace both alleles could be due to limitation(s) in the methods used to achieve this. For example, promastigotes may be able to survive without salvage of LA, yet after replacement of the second *LPLA* allele, the parasites may take some time to up-regulate the biosynthesis pathway. The procedure I followed involved incubating transfected parasites for 24 h before cloning. Upon cloning it is clear that competition with other parasites is not an issue, however it could be that during the 24 h period preceding the initiation of cloning, the *LPLA* null mutants are out-competed/over-grown by parasites with duplicated genomes that do not need to switch metabolic pathways to survive. I would argue against this because promastigotes do not grow exponentially during the 24 h post transfection (data not shown). In terms of the attempted creation of a *LIPA* null mutant, 10 μM LA was included in each step of selection of the second-round allele replacement, yet it was still not possible to replace the remaining allele without the parasites duplicating their genome. This again provides support for the notion that both *LIPA* and *LPLA* are essential, since as explained above, if LPLA is important when LA is not limiting, replacement of the *LIPA* gene should be feasible.

A second possibility to explain the essentiality of both *LIPA* and *LPLA* is that only one of the two pathways used to acquire LA is active at any one time, and that the expression of the genes involved in these pathways is regulated by LA availability in the medium. It seems peculiar that if an organism has the genetic capacity to encode a compensatory enzyme/set of enzymes to complement a growth defect, it would not do so. For example, an ACP knock-down mutant in human HEK293 T cells resulted in decreased cell growth and eventually death by 72 h in culture, and the principal cause was due to an almost complete loss of protein lipoylation (Feng *et al.*, 2009a). This study highlighted the importance of LA biosynthesis in lipoylation of α-KADHs and the GCC (Feng *et al.*, 2009a). Also, it has previously

been shown in mouse that deletion of lipA has an embryonic-lethal phenotype (Yi & Maeda, 2005; Yi *et al.*, 2009), which further emphases the importance of LipA in mammals. Nevertheless, in the HEK293 T cell ACP knock-down line, addition of 2 μM exogenous LA in the growth medium did not rescue growth or result in increased lipoylation of α-KADHs and H-protein (Feng *et al.*, 2009a). This result argues against the possibility that LA availability in the medium was a limiting factor affecting the capacity of LA salvage to compensate for lack of LA biosynthesis. Also, in this thesis 10 μM LA was included in the growth medium at all steps during the selection of a *LIPA* null mutant, and yet LA salvage could still not compensate for a loss of LIPA protein. This fact thus opposes the theory that the LA salvage pathway is 'switched off' in LA-poor conditions.

A third possibility to explain the essentiality of both *LIPA* and *LPLA* is that the pathways for LA biosynthesis and salvage are indeed redundant in terms of lipoylating α-KADHs and the GCC, yet the enzymes involved could have essential roles other than catalysing lipoylation reactions. For example, in *E. coli*, LipB has a surprising additional role as a negative regulator of *dam* gene expression. The *dam* gene encodes a DNA methyltransferase which transfers methyl groups from SAM to adenine residues in the sequence 5'-GATC-3' in double stranded DNA, and has a large impact on the chromosome replication, gene expression and mismatch repair (Barras & Marinus, 1989). Also, in *P. falciparum*, an unexpected result was obtained whereby *LipB* null mutants exhibited a faster growth phenotype that was mainly due to accelerated progression through the intraerythrocytic cell cycle (Gunther *et al.*, 2007). Again, it could be that LipB has a role other than octanoylation of the PDH in *P. falciparum* (although it cannot be ruled out that in the *P. falciparum LipB* null line, increased growth rate is due to an increase in myristate production through type II FASI) (Gunther *et al.*, 2007).

A body of data is also accumulating that implicates LA biosynthesis in mitochondrial RNA processing (Hiltunen *et al.*, 2009). In 1993, a screen for mutants with altered mitochondrial tRNA precursor/product ratios resulted in the identification of the *lip5* mutant (Sulo & Martin, 1993). The *lip5* mutant had a non-functional *lipA* gene and as a result completely lacked lipoylated α-KADHs and GCC. Other phenotypes included accumulation of mitochondrial tRNA precursors, destabilisation of the mitochondrial genome and poor cell growth, which was not rescued by the addition of 2 μM LA (Sulo & Martin, 1993). In *S. cerevisiae*,

mutants in type II FAS genes lack lipoylated α-KADHs and GCC (Schonauer *et al.*, 2008). It had been hypothesised that the main role of type II FAS is the provision of octanoyl-ACP as a substrate for LA biosynthesis (Brody *et al.*, 1997; Hiltunen *et al.*, 2009; Wada *et al.*, 1997), however this particular study went a step further to show that FASII mutants had similar RNA processing phenotypes to that of the *lip5* mutant (Schonauer *et al.*, 2008). The authors illustrated that mutants of the *pdh, α-kgdh* and *gcc* genes did not cause RNA processing phenotypes, and it was proposed that biosynthesis of LA is not just involved in lipoylating α-KADHs and the GCC, but is also somehow implicated in RNA processing (Schonauer *et al.*, 2008). Interestingly, a similar phenotype of LA depletion and changes in mitochondrial ultrastructure was observed in a *T. brucei* PCF ACP knock-down line (Stephens *et al.*, 2007). A follow-up study subsequently showed that RNAi of ACP resulted in malformation of the mitochondrial membrane, and the authors concluded that the principle cause of the RNAi phenotype was due to a change in phospholipid composition of the mitochondrial membrane (Guler *et al.*, 2008). It would be interesting to determine whether RNA processing is also affected in this line, and whether LA is also the cause of this phenotype. Unfortunately, it was outside the scope of this thesis to investigate the exciting possibility that enzymes involved in LA metabolism play key roles in cellular functions other than lipoylation of α-KADHs and the GCC.

A fourth and final possibility to explain the essentiality of both *LIPA* and *LPLA* in *L. major* is that the LA biosynthesis and salvage pathways lipoylate their apoproteins in a substrate-specific manner. Given that all four apoproteins (H-protein, E2k, E2p and E2b) are lipoylated (and assumedly active) during logarithmic growth phase in *L. major* (see Figure 3.1), there is a possibility that three of them are essential to parasite survival (excluding the H-protein, which has been shown to be dispensable in *L. major* promastigotes (Scott *et al.*, 2008)). If LA biosynthesis and salvage have differential substrate specificities for the α-KADHs, one would then expect that these genes be essential to promastigote survival. An interesting observation in *E. coli* is that the *lipB* null mutant KER184 can grow on any medium that bypasses the need for the α-KGDH (for example, on succinate-containing medium), however acetate-containing minimal medium is not sufficient to sustain growth. This result infers that in the absence of LipB, LplA is not able to sufficiently lipoylate the α-KGDH to permit cell growth, yet can lipoylate the PDH enough to allow colony formation (Reed & Cronan, 1993). Another interesting study showed

that in *L. monocytogenes,* which encodes two LplA enzymes (LplA1 and LplA2) but lacks LA biosynthesis enzymes, LplA1 is essential for intracellular survival of the bacterium, and that LplA1 and LplA2 are not redundant (O'Riordan *et al.*, 2003). This interesting phenomenon was shown to be due to the fact that LplA1 can use lipoyl-peptides and ATP as substrates for lipoylation, whereas LplA2 requires free LA and ATP (Keeney *et al.*, 2007).

6.3.3.4Ectopic expression of lipoylating enzymes

Since it was not possible to study the functions of LIPA or LPLA by gene replacement, an alternative approach was taken. *LIPB-His*, *LPLA-His* and *LPLA^{H118A}-His* constructs were expressed in promastigotes and the effect of protein expression was assessed in terms of growth and lipoylation patterns at different stages. Unfortunately, it was not possible to express the *LIPA-His* construct in promastigotes, even when the G418 selection pressure was lowered to 10 μg ml^{-1} from the standard 50 μg ml^{-1} (data not shown). One possible explanation is that high ectopic expression of *LIPA* is not tolerated in promastigotes.

My results indicate that there is a dramatic overexpression of LIPB-His relative to endogenous LIPB in the WT[*LIPB-His*] line, since it was not possible to detect endogenous LIPB protein in wild-type promastigotes, whereas in the mutant line LIPB was readily detected by the α-His antibody (see Figure 5.5). Expression of LPLA-His and LPLA^{H118A}-His is easily detectable in the WT[*LPLA-His*] and WT[*LPLA^{H118A}-His*] lines using an α-His antibody, respectively (see Figure 4.6). The endogenous expression levels of LPLA remain unknown because the antibody raised against the recombinant protein proved to be cross-reactive with numerous other parasite proteins and no protein of the expected size of *Lmj*LPLA was detected in the blots. Growth rates were significantly slower in all three mutant lines compared to that observed in the empty vector control line, and this slow-growth phenotype was observed when G418-resistant lines were grown in the presence of either 10 μg ml^{-1} or 50 μg ml^{-1} selection pressure (see Figure 4.5 and Figure 5.7). The reason why G418-resistant lines were cultured in the presence of two different drug concentrations was because it had been shown before in *L. major* that overexpression levels of an ATPase protein VPS4, were correlated with the concentration of G418 included in the growth medium (Besteiro *et al.*, 2006). In

the experiments carried out in this thesis, expression levels of LIPB-His, LPLA-His and LPLA^{H118A}-His did not differ when the mutant lines were grown in the presence of either 10 μg ml^{-1} or 50 μg ml^{-1} G418 drug selection pressure (see Figure 4.6 and Figure 5.8).

It was hypothesised that the cause of reduced growth rate in the three mutant lines would be due to alterations in lipoylation patterns, which could impact on cell metabolism and growth. However, the lipoylation patterns in the WT[*LIPB-His*] and WT[*LPLA-His*] lines were comparable to those observed in the empty vector control lines (see Figure 4.7 and Figure 5.9). This potentially indicates that overexpression of LIPB-His and LPLA-His interferes with pathways other than lipoylation, although investigation of these options was not within the scope of this thesis. Interestingly, lipoylation of the H-protein was higher in the WT[*LPLA^{H118A}-His*] line at all time points investigated apart from in purified metacyclic promastigotes. Also, in WT[*LIPB-His*] and WT[*LPLA^{H118A}-His*] lines, but not in the WT[*LPLA^{-}His*] line, E2p and E2b were lipoylated in metacyclic promastigotes (see Figure 4.7 and Figure 5.9). I consistently found in multiple independent experiments carried out with the wild-type line, that metacyclic promastigotes had lipoylated E2k, but I had never observed lipoylation of E2p or E2b (for a representative result see Figure 3.1), and as such this was a potentially interesting phenotype. Based on sequence alignments with LplA and LT enzymes in other organisms whose three-dimensional structures have been determined, the His^{118} residue in *L. major* is likely to be one of the key residues involved in forming the hydrophobic crevice to which the lipoyl-AMP intermediate binds (Fujiwara *et al.*, 2005; Kim do *et al.*, 2005; McManus *et al.*, 2005). As explained in Section 4.4, the justification for creating the LPLA^{H118A}-His mutant was based on the published observation that an *E. coli* BPL mutant, *Ec*BPL^{R118G}, biotinylates substrates other than BCCP (Choi-Rhee *et al.*, 2004; Cronan, 2005). Given that the *Ec*BPL^{R118G} protein also bears a mutation within the ATP-binding domain, I hypothesised that the LPLA^{H118A}-His mutant protein would lipoylate substrates other than the endogenous substrates of wild-type LPLA. Therefore, one possibility is that 'promiscuous lipoylation' is the cause of the observed lipoylation of E2p and E2b in metacyclic promastigotes in the WT[*LPLA^{H118A}-His*] line. In addition, the fact that expression of LPLA-His does not result in lipoylation of E2p and E2b in metacyclic promastigotes suggests that LPLA may not normally lipoylate E2p and E2b. The fact that LIPB-His overexpression results in lipoylation of E2p and E2b in

metacyclic promastigotes is also interesting, and suggests that LIPB may have a substrate-specificity for E2p an E2b, and in wild-type metacyclic promastigotes, perhaps LIPB is not expressed and as such does not lipoylate E2p and E2b.

6.4 Conclusions and future directions

The overall aim of this thesis was to initiate an understanding of LA metabolism in *L. major*. I illustrated that *L. major* likely possesses genes that encode all subunits of the α-KGDH, PDH, BCKDH and GCC. All of the complexes are likely to be mitochondrial, although it is noteworthy that the subcellular localisations of the E3 and E3BP were not entirely predictable. As such, it would be interesting to determine the sub-cellular localisation of E3 (and E3BP), and perhaps the protein has dual localisation, as has been reported in *T. brucei* BSF (Danson *et al.*, 1987).

L. major was shown to possess a functional biosynthesis pathway consisting of LIPB and LIPA enzymes. Bacterial complementation studies did not yield a satisfactory answer regarding the functionality of LPLA; future studies would focus on optimising the expression and purification of LPLA to a high purity, such that it could be biochemically characterised. All three enzymes were shown to localise to the mitochondrion, which itself reinforces the likelihood that the α-KADH and GCC complexes are also localised to this organelle. In *E. coli*, LA salvage and biosynthesis pathways are redundant, and this was hypothesised to be the case in *L. major* also, because similarly to *E. coli*, both LA metabolism pathways are localised to the same organelle. Unexpectedly, LIPA and LPLA appear to have an essential function in *L. major*, since it was not possible to ablate either gene. Evidence is accumulating in other organisms that LipA and LipB may have roles other than lipoylation of α-KADHs and the GCC. Future work will be aimed at unravelling the seemingly complex role(s) that LipA and LipB play in the intersection between energy metabolism, type II FAS and mitochondrial RNA processing (Hiltunen *et al.*, 2009). Another hypothesis to explain the essentiality of LIPA and LPLA in *L. major* promastigotes is that LA salvage and biosynthesis enzymes may have differential substrate specificities. Preliminary results from overexpression of different enzymes involved in LA metabolism indicate that this could indeed be the case, and it would be of interest to build upon these results by determining whether LIPB and LPLA exhibit substrate-specificities for octanoylation/lipoylation *in vitro*. The potential for substrate-specificity between the

two different LA metabolic pathways is something that has not been addressed in the literature. However, one must pose the question, what is the purpose of having a LA salvage pathway when the LA biosynthesis pathway enables the cell to be completely self-sufficient?

7 Appendix

```
Rn_E1k          -----MFHLRTCAAKLRPLTASQTVKTFSQNKPAAIRTFQQIRCYSAPVAAEP-FLSGTS 54
Mm_E1k          -----MFHLRTCAAKLRPLTASQTVKTFSQNKPAAIRTFQQIRCYSAPVAAEP-FLSGTS 54
Hs_E1k          -----MFHLRTCAAKLRPLTASQTVKTFSQNRPAAARTFQQIRCYSAPVAAEP-FLSGTS 54
Rn_E1kL         -----MSQLRLLLFRLG----PQARKLLATR---DIAAFGGRRRSSGPPTTIPRSRGGVS 48
Mm_E1kL         -----MSQLRLLPFRLG----PRATKLLATR---AIPVFSGCRRSSGPPTTIPRSRSGVS 48
Hs_E1kL         -----MSQLRLLPSRLG----VQAARLLAAH---DVPVFGWRSRSSGPPATFPSSKGGGG 48
Rn_DHTKD1       ----------------------------------------------MASATVAAAGRA 12
Mm_DHTKD1       ----------------------------------------------MASAATVAAAGRA 13
Hs_DHTKD1       ----------------------------------------------MASATAAAARRG 12
Lm_E1kB         MMRR--ALSGVVAVR-----ASAMRSYTDARTIRKPN------------PYDQLVNAEN 40
Lm_E1kA         MMRRLVPVRGVVSCGSAVAPTSAFPCASHAALIMGRRRAAEAVPERQLLFDNDSFLSGSS 60


Rn_E1k          SNYVEEMYCAWLENPKSVHKSWDIFFRNTNAGAPPGTAYQSPLSLSRSSLATMAHAQSLV 114
Mm_E1k          SNYVEEMYCAWLENPKSVHKSWDIFFRNTNAGAPPGTAYQSPLSLSRSSLATMAHAQSLV 114
Hs_E1k          SNYVEEMYCAWLENPKSVHKSWDIFFRNTNAGAPPGTAYQSPLPLSRGSLAAVAHAQSLV 114
Rn_E1kL         PSYVEEMYFAWLENPQSVHKSWDNFFQRATKEASVGPAQPQPP-------AVIQESRASV 101
Mm_E1kL         SSYVEEMYFAWLENPQSVHKSWDSFFQRASKEASVGPAQPQLP-------AVLQESRTSV 101
Hs_E1kL         SSYMEEMYFAWLENPQSVHKSWDSFFREASEEAFSGSAQPRPP-------SVVHEGRSAV 101
Rn_DHTKD1       LRRAVPLLRRSYQTERGVYGYRP---RKAGSGEPRGDR-----------------ARPSV 52
Mm_DHTKD1       LRRAVLLLRRGYQTERGVYGYRP---RKAKSGEPRGDR-----------------ARPSV 53
Hs_DHTKD1       LGRALPLLWRGYQTERGVYGYRP---RKPESREPQGAL-----------------ERPPV 52
Lm_E1kB         QHYVEDLMRQYEADSALVDPSWVPVLEAIRSGSDDSPVVAT--------------FSRPT 86
Lm_E1kA         AMYMDGLYQQWKKDPASVDASWAELFSRSDLGNYNHALLDTPI----------CVLPAKS 110
                      :          *

Rn_E1k          EAQPNVDKLVEDHLAVQSLIRAYQIRGHHVAQLDPLGILDADLDSSVPAD------IISS 168
Mm_E1k          EAQPNVDKLVEDHLAVQSLIRAYQIRGHHVAQLDPLGILDADLDSSVPAD------IISS 168
Hs_E1k          EAQPNVDKLVEDHLAVQSLIRAYQIRGHHVAQLDPLGILDADLDSSVPAD------IISS 168
Rn_E1kL         SSCTKTSKLVEDHLAVQSLIRAYQIRGHHVAQLDPLGILDADLDSFVPSD------LITT 155
Mm_E1kL         SSCTKTSKLVEDHLAVQSLIRAYQIRGHHVAQLDPLGILDADLDSFVPSD------LITT 155
Hs_E1kL         SSRTKTSKLVEDHLAVQSLIRAYQIRGHHVAQLDPLGILDADLDSFVPSD------LITT 155
Rn_DHTKD1       DHG-----------LARLVTVYCEHGHKAAQINPLFPGQALLDTVPEIQ------ALVQ 94
Mm_DHTKD1       DHG-----------LARLVTVYCEHGHKAAQINPLFPGQALLDTVPEIQ------ALVR 95
Hs_DHTKD1       DHG-----------LARLVTVYCEHGHKAAKINPLFTGQALLENVPEIQ------ALVQ 94
Lm_E1kB         DAKSLSEKQRHDNMRLSWMIREYERFGHHMANVDPLSGYHADNCILG-SR------TLAP 139
Lm_E1kA         SDEAVVKQSLADCGRLIRMIHTFEDRGHLMAQTDPLNYVDTDVTERTPSRRYKEMVRLDL 170
                .             :  ::  :   **  *: :**   .:              :

Rn_E1k          TDKLGFYGLHESDLDKVFHLPTTTFIGGQEPALPLREIIRRLEMAYCQHIGVEFMFINDL 228
Mm_E1k          TDKLGFYGLHESDLDKVFHLPTTTFIGGQEPALPLREIIRRLEMAYCQHIGVEFMFINDL 228
Hs_E1k          TDKLGFYGLDESDLDKVFHLPTTTFIGGQESALPLREIIRRLEMAYCQHIGVEFMFINDL 228
Rn_E1kL         IDKLAFYDLQEADLDKEFRLPTTTFIGGSENTLSLREIIRRLESTYCQHIGLEFMFINDV 215
Mm_E1kL         IDKLAFYDLQEADLDKEFRLPTTTFIGGPENTLSLREIIRRLESTYCQHIGLEFMFINDV 215
Hs_E1kL         IDKLAFYDLQEADLDKEFQLPTTTFIGGSENTLSLREIIRRLENTYCQHIGLEFMFINDV 215
Rn_DHTKD1       TLQGPFT-------------TTGLLNMGKEEASLEEVLAYLNHIYCGPISIETAQLQSQ 140
Mm_DHTKD1       TLQGPFT-------------TTGLLNLGKEAASLEEVLAYLNHIYCGPISIETAQLQSQ 141
Hs_DHTKD1       TLQGPFH-------------TAGLLNMGKEEASLEEVLVYLNQIYCGQISIETSQLQSQ 140
Lm_E1kB         EEFGFTKDDLTHVFNVTFGASHEATFVSGGTAMTLQQIIDQLRRLYCGPIGFEFMSSGFF 199
Lm_E1kA         AYFGFSDKDLDRVVRVGFQNQMGGIYDTSSPQLTIRQLHELLTERYCGRIGFELVHLTDG 230
                                           .:.::   *   **  *..*

Rn_E1k          EQCQWIRQKFETPGIMQFTN-----EEKRTLLARLVRSTRFEEFLQRKWSSEKRFGLEGC 283
Mm_E1k          EQCQWIRQKFETPGIMQFTN-----EEKRTLLARLVRSTRFEEFLQRKWSSEKRFGLEGC 283
Hs_E1k          EQCQWIRQKFETPGIMQFTN-----EEKRTLLARLVRSTRFEEFLQRKWSSEKRFGLEGC 283
Rn_E1kL         EQCQWIRQKFETPGVMKFSI-----EEKRTLLARLVRSMRFEDFLARKWSSEKRFGLEGC 270
Mm_E1kL         EQCQWIRQKFETPGVMQFSV-----EEKRTLLARLVRSMRFEDFLARKWSSEKRFGLEGC 270
Hs_E1kL         EQCQWIRQKFETPGVMQFSS-----EEKRTLLARLVRSMRFEDFLARKWSSEKRFGLEGC 270
Rn_DHTKD1       EEKDWFARRFEELKKETFTT-----EERKHLSKLLLESQEFDHFLATKFATVKRYGGEGA 195
Mm_DHTKD1       EERDWFARRFEELKKETFTT-----EERKYLSKLLLESQEFDHFLATKFATVKRYGGEGA 196
Hs_DHTKD1       DEKDWFAKRFEELQKETFTT-----EERKHLSKLMLESQEFDHFLATKFSTVKRYGGEGA 195
Lm_E1kB         ELRNWFR--QEVMDSLQSLP----TEERRLYYNDVVKACGFEKFLQLKYATKHRFGLDGG 253
Lm_E1kA         DAKRFVRSQIELKDGCSALHRPMSREERLRIWDTVASAVFFEDFFKRKYSTQKRFGCDGA 290
                :   :.    *             **:      :  :  *:.*:  *::: :*:* :*

Rn_E1k          EVLIPALKTIIDMSSANGVDYVIMGMPHRGRLNVLANVIRKELEQIFCQFDSKLEAAD-- 341
Mm_E1k          EVLIPALKTIIDMSSANGVDYVIMGMPHRGRLNVLANVIRKELEQIFCQFDSKLEAAD-- 341
Hs_E1k          EVLIPALKTIIDKSSENGVDYVIMGMPHRGRLNVLANVIRKELEQIFCQFDSKLEAAD-- 341
Rn_E1kL         EVMIPALKTIIDKSSEMGVENVILGMPHRGRLNVLANVIRKDLEQIFCQFDPKLEAAD-- 328
```

```
Mm_E1kL          EVMIPALKTIIDKSSEMGIENVILGMPHRGRLNVLANVIRKDLEQIFCQFDPKLEAAD-- 328
Hs_E1kL          EVMIPALKTIIDKSSEMGIENVILGMPHRGRLNVLANVIRKDLEQIFCRFDPKLEAAD-- 328
Rn_DHTKD1        ESMMGFFHELLKLSAYGGITDIIIGMPHRGRLNLLTGLLQLPPELMFRKMRGLSEFPENV 255
Mm_DHTKD1        ESMMGFFHELLKLSAYGGITDIIIGMPHRGRLNLLTGLLQLPPELMFRKMRGLSEFPENV 256
Hs_DHTKD1        ESMMGFFHELLKMSAYSGITDVIIGMPHRGRLNLLTGLLQFPPELMFRKMRGLSEFPENF 255
Lm_E1kB          EALIPALKAAILTSSDLGVQSAIIGMPHRGRLNVLANVLRKSLRAILNEFEGR--VAIED 311
Lm_E1kA          ESMVAGLRALLEKSSELGVQAINLGMPHRGRLNVLCHVIGKPFEVILKEFVGVTGQELHP 350
                 * ::  ::  :  *:  *:    :**.******:*  ::    . :: .:

Rn_E1k           EGSGDMKYHLGMYHRRINRVTDRNITLSLVANPSHLEAADPVVMGKTKAEQFYCGDTE-- 399
Mm_E1k           EGSGDMKYHLGMYHRRINRVTDRNITLSLVANPSHLEAADPVVMGKTKAEQFYCGDTE-- 399
Hs_E1k           EGSGDMKYHLGMYHRRINRVTDRNITLSLVANPSHLEAADPVVMGKTKAEQFYCGDTE-- 399
Rn_E1kL          EGSGDMKYHLGMYHERINRVTNRNITLSLVANPSHLEAVDPVVQGKTKAEQFYRGDAQ-- 386
Mm_E1kL          EGSGDMKYHLGMYHERINRVTNRNITLSLVANPSHLEAVDPVVQGKTKAEQFYRGDAQ-- 386
Hs_E1kL          EGSGDMKYHLGMYHERINRVTNRNITLSLVANPSHLEAVDPVVQGKTKAEQFYRGDAQ-- 386
Rn_DHTKD1        AAIGDVLSHLTSSVD-LDFGAHRPLHVTMLPNPSHLEAINPVAVGKTRGRQQSQEDGDYS 314
Mm_DHTKD1        ATIGDVLSHLTSSVD-LDFGAHQPLHVTMLPNPSHLEAVNPVAVGKTRGRQQSREDGDYS 315
Hs_DHTKD1        SATGDVLSHLTSSVD-LYFGAHHPLHVTMLPNPSHLEAVNPVAVGKTRGRQQSRQDGDYS 314
Lm_E1kB          AHLTGDVEYHLGKRKHVSLPNNKSIELDLLPNPSHLEAVNPLVLGKARARQTYTNDVE-- 369
Lm_E1kA          FQIQSDVKYHLGYRGQLKLNSGKVMETEMLFNPSHLEAVNPFVQGYTRAMQVSLGEKG-- 408
                     .   :       :     : :   :: ******* :*.. * ::. *    :

Rn_E1k           -------GKKVMSILLHGDAAFAGQGIVYETFHLSDLPSYTTHGTVHVVVNNQIGFTTDP 452
Mm_E1k           -------GKKVMSILLHGDAAFAGQGIVYETFHLSDLPSYTTHGTVHVVVNNQIGFTTDP 452
Hs_E1k           -------GKKVMSILLHGDAAFAGQGIVYETFHLSDLPSYTTHGTVHVVVNNQIGFTTDP 452
Rn_E1kL          -------GRKVMSILVHGDAAFAGQGVVYETFHLSDLPSYTTNGTVHVVVNNQIGFTTDP 439
Mm_E1kL          -------GRKVMSILVHGDAAFAGQGVVYETFHLSDLPSYTTNGTVHVVVNNQIGFTTDP 439
Hs_E1kL          -------GKKVMSILVHGDAAFAGQGVVYETFHLSDLPSYTTNGTVHVVVNNQIGFTTDP 439
Rn_DHTKD1        PNGSAQPGDKVICLQVHGDASFCGQGIVLETFTLSNLPHFRIGGSIHLIVNNQLGYTTPA 374
Mm_DHTKD1        PNGSAQPGDKVICLQVHGDASFCGQGIVLETFTLSNLPHFRIGGSIHLIVNNQLGYTTPA 375
Hs_DHTKD1        PDNSAQPGDRVICLQVHGDASFCGQGIVPETFTLSNLPHFRIGGSVHLIVNNQLGYTTPA 374
Lm_E1kB          -------CTAVLPILIHGDAAFAGQGSCYETMGFCELENFHVGGTLHLVINNQIGFTTNP 422
Lm_E1kA          -------REKVLPIEIHGDAAFAGQGVAFETMCISEVGEQDTGGTVHVVCNNQIGFTTDP 461
                           *: : :****:*.***   **: :.::      *::*:: ***:*:** .

Rn_E1k           RMARSSPYPTDVARVVNAPIFHVNSDDPEAVMYVCKVAAEWRNTFHKDVVVDLVCYRRNG 512
Mm_E1k           RMARSSPYPTDVARVVNAPIFHVNSDDPEAVMYVCKVAAEWRNTFHKDVVVDLVCYRRNG 512
Hs_E1k           RMARSSPYPTDVARVVNAPIFHVNSDDPEAVMYVCKVAAEWRSTFHKDVVVDLVCYRRNG 512
Rn_E1kL          RMARSSPYPTDVARVVNAPIFHVNADDPEAVIYVCSVAAEWRNTFNKDVVVDLVCYRRRG 499
Mm_E1kL          RMARSSPYPTDVARVVNAPIFHVNADDPEAVIYVCSVAAEWRNTFNKDVVVDLVCYRRRG 499
Hs_E1kL          RMARSSPYPTDVARVVNAPIFHVNADDPEAVIYVCSVAAEWRNTFNKDVVVDLVCYRRRG 499
Rn_DHTKD1        ERGRSSLYSSDIGKLVGCAIIHVNGDSPEEVVRATRLAFEYQRQFRKDVIIDLLCYRQWG 434
Mm_DHTKD1        ERGRSSLYSSDIGKLVGCAIIHVNGDSPEEVVRATRLAFEYQRQFRKDVIVDLLCYRQWG 435
Hs_DHTKD1        ERGRSSLYCSDIGKLVGCAIIHVNGDSPEEVVRATRLAFEYQRQFRKDVIIDLLCYRQWG 434
Lm_E1kB          KDSRASAYCTDLSKVNNAPVMHVNGDDVDACVKAAKIAARFRQQFHHDIIIDLVCYRRYG 482
Lm_E1kA          KSSRSSAYCSDLGRVYNCPILHVNGDYPEEVIRVFEFAAEYRARFHKSVVIDLVCYRRFG 521
                 . .*:* * :*:.:: ...::***.*  :  : .  .* .::  *.:.:::**:***: *

Rn_E1k           HNEMDEPMFTQPLMYKQIRKQKPVLQKYAELLVSQGVVNQPEYEEEISKYDKICEEAFTR 572
Mm_E1k           HNEMDEPMFTQPLMYKQIRKQKPVLQKYAELLVSQGVVNQPEYEEEISKYDKICEEAFTR 572
Hs_E1k           HNEMDEPMFTQPLMYKQIRKQKPVLQKYAELLVSQGVVNQPEYEEEISKYDKICEEAFAR 572
Rn_E1kL          HNEMDEPMFTQPLMYKQIHKQVPVLKKYADKLIAEGTVTLQEFEEEIAKYDRICEEAYGR 559
Mm_E1kL          HNEMDEPMFTQPLMYKQIHKQVPVLKKYADKLIAEGTVTLQEFEEEIAKYDRICEEAYGR 559
Hs_E1kL          HNEMDEPMFTQPLMYKQIHRQVPVLKKYADKLIAEGTVTLQEFEEEIAKYDRICEEAYGR 559
Rn_DHTKD1        HNELDEPFFTNPVMYKIIRARKSIPDTYAEHLIASGLMTQEEVSDIKASYYAKLNGHLAN 494
Mm_DHTKD1        HNELDEPFFTNPVMYKIIRARKSIPDTYAEHLIASGLMTQEEVSDIKTSYYTKLNDHLAN 495
Hs_DHTKD1        HNELDEPFFTNPIMYKIIRARKSIPDTYAEHLIAGGLMTQEEVSEIKSSYYAKLNDHLNN 494
Lm_E1kB          HNELDEPFFTNPQLYHQIRQHPSVVDIYTKTLIKDGVLTAEEAKAKDKDWEGVLRQAYDR 542
Lm_E1kA          HNENDDPSITQPLMYERVRAMPDVFRRYTDALITQGILTPQQSTQKAID-EKARYGSYQE 580
                 *** * *  *:* :*. ::    :   *:. *:  * :.  :      .         .

Rn_E1k           SKDEK---------ILHIKHWLDSPWPGFFTLDGQPRSMTCPSTGLEEDILTHIGNVASS 623
Mm_E1k           SKDEK---------ILHIKHWLDSPWPGFFTLDGQPRSMTCPSTGLEEDVLFHIGKVASS 623
Hs_E1k           SKDEK---------ILHIKHWLDSPWPGFFTLDGQPRSMSCPSTGLTEDILTHIGNVASS 623
Rn_E1kL          SKDKK---------ILHIKHWLDSPWPGFFNVDGEPKSMTYPTTGIPEDTLSHIGNVASS 610
Mm_E1kL          SKDKK---------ILHIKHWLDSPWPGFFNVDGEPKSMTCPTTGIPEEMLTHIGSVASS 610
Hs_E1kL          SKDKK---------ILHIKHWLDSPWPGFFNVDGEPKSMTCPATGIPEDMLTHIGSVASS 610
Rn_DHTKD1        VAHYS---------PPAPH--LQARWQGLVQP---AACVTTWDTGVPLELLRFVGVKSVE 540
Mm_DHTKD1        VAHYS---------PPATN--LQARWQGLVQP---EACVTTWDTGVPLELLRFIGVKSVE 541
Hs_DHTKD1        MAHYR---------PPALN--LQAHWQGLAQP---EAQITTWSTGVPLDLLRFVGMKSVE 540
Lm_E1kB          MNSAQNFVKVMPVFDPESENTSADLSSAKIAATRVPPPVSAVDTGVETQTLRAAGLRLAS 602
Lm_E1kA          AAAQVNYAEYLKKSIPD-KWKCMKYSDELGNVTQHP-------TAITQETVDKVLKALKT 632
                                 .                        *.:  : :
```

```
Rn_E1k          VPVENFTIHGGLSRILKTRRELVTNRT-VDWALAEYMAFGSLLKEGIHVRLSGQDVERGT 682
Mm_E1k          VPVENFTIHGGLSRILKTRRELVTNRT-VDWALAEYMAFGSLLKEGIHVRLSGQDVERGT 682
Hs_E1k          VPVENFTIHGGLSRILKTRGEMVKNRT-VDWALAEYMAFGSLLKEGIHIRLSGQDVERGT 682
Rn_E1kL         VPLEDFKIHTGLSRILRGRADMTKKRT-VDWALAEYMAFGSLLKEGIHVRLSGQDVERGT 669
Mm_E1kL         VPLEDFKIHTGLSRILRGRADMTKKRT-VDWALAEYMAFGSLLKEGIHVRLSGQDVERGT 669
Hs_E1kL         VPLEDFKIHTGLSRILRGRADMTKNRT-VDWALAEYMAFGSLLKEGIHVRLSGQDVERGT 669
Rn_DHTKD1       VPEELQLHSHLLKMYVQSRMEKVKNGTNLDWATAETLALGSLLAQGFNVRLSGQDVERGT 600
Mm_DHTKD1       VPEELQVHSHLLKMYVQSRMEKVKNGSGLDWATAETLALGSLLAQGFNVRLSGQDVERGT 601
Hs_DHTKD1       VPRELQMHSHLLKTHVQSRMEKMMDGIKLDWATAEALALGSLLAQGFNVRLSGQDVERGT 600
Lm_E1kB         IPKEMQKPHPVVERTYAARKKGTEQGDAIEWCQAELMALATLSMQGVPIRLTGEDVERGT 662
Lm_E1kA         YPEGFQL-HPKLKAVLDRRNETIETGEGIEWGTAEALAFGSLLLEGHQVRVTGEDVERGT 691
                 *         :.     * .       ::*  ** :*:.:*  :*  :*::*:** ***

Rn_E1k          FSHRHHVLHDQNVDKRTCIPMNHLWPNQAPY-TVCNSSLSEYGVLGFELGFAMASPNALV 741
Mm_E1k          FSHRHHVLHDQNVDKRTCIPMNHLWPNQAPY-TVCNSSLSEYGVLGFELGFAMASPNALV 741
Hs_E1k          FSHRHHVLHDQNVDKRTCIPMNHLWPNQAPY-TVCNSSLSEYGVLGFELGFAMASPNALV 741
Rn_E1kL         FSHRHHVLHDQDVDRRTCVPMNHLWPDQAPY-TVCNSSLSEYGVLGFELGYAMASPNALV 728
Mm_E1kL         FSHRHHVLHDQEVDRRTCVPMNHLWPDQAPY-TVCNSSLSEYGVLGFELGYAMASPNALV 728
Hs_E1kL         FSHRHHVLHDQEVDRRTCVPMNHLWPDQAPY-TVCNSSLSEYGVLGFELGYAMASPNALV 728
Rn_DHTKD1       FSQRHAMVVCQNTDD-VYIPLNHMDPNQKGFLEVSNSPLSEEAVLGFEYGMSIESPKLLP 659
Mm_DHTKD1       FSQRHAMVVCQDTDD-AYIPLNHMDPNQKGFLEVSNSPLSEEAVLGFEYGMSIESPTLLP 660
Hs_DHTKD1       FSQRHAIVVCQETDD-TYIPLNHMDPNQKGFLEVSNSPLSEEAVLGFEYGMSIESPKLLP 659
Lm_E1kB         FTQRHAGITDMKTNL-KYFPVKMLSPS-QALITISNSSLSELGVCGFEMGYNMENTRSIT 720
Lm_E1kA         FAQRHAVIHDQSQER-TYVPLAHISDT-QGRMIINNSPLSEYGMLGYAAGYSLYDPTSLV 749
                *::**  :   . :    .*:  :         : **.*** .: *:  *  : ..  :

Rn_E1k          LWEAQFGDFNNMAQCIIDQFICPGQAKWVRQNGIVLLLPHGMEGMGPEHSSARPERFLQM 801
Mm_E1k          LWEAQFGDFNNMAQCIIDQFICPGQAKWVRQNGIVLLLPHGMEGMGPEHSSARPERFLQM 801
Hs_E1k          LWEAQFGDFHNTAQCIIDQFICPGQAKWVRQNGIVLLLPHGMEGMGPEHSSARPERFLQM 801
Rn_E1kL         LWEAQFGDFHNTAQCIIDQFISTGQAKWVRHNGIVLLLPHGMEGMGPEHSSARPERFLQM 788
Mm_E1kL         LWEAQFGDFHNTAQCIIDQFISTGQAKWVRHNGIVLLLPHGMEGMGPEHSSARPERFLQM 788
Hs_E1kL         LWEAQFGDFHNTAQCIIDQFISTGQAKWVRHNGIVLLLPHGMEGMGPEHSSARPERFLQM 788
Rn_DHTKD1       LWEAQFGDFFNGAQIIFDTFISGGEAKWLLQSGLVILLPHGYDGAGPDHSSCRIERFLQM 719
Mm_DHTKD1       LWEAQFGDFFNGAQIIFDTFISGGEAKWLLQSGLVILLPHGYDGAGPEHSSCRIERFLQM 720
Hs_DHTKD1       LWEAQFGDFFNGAQIIFDTFISGGEAKWLLQSGIVILLPHGYDGAGPDHSSCRIERFLQM 719
Lm_E1kB         IWEAQFGDFANGAQVIFDQFLSCCEEKWNEHSSLVLSLPHGYSGAGPEHSSARVERFLQL 780
Lm_E1kA         IWEAQYGDFANGATIVFDQFLSAGESKWNQQQSCIVTLPHGYDGKGAEHSSGRLERFLQM 809
                :****:*** * *  ::* *:.  : **  :.. :: **** .* *.:*** * *****:

Rn_E1k          CNDDPDVLP---NLQEENFDISQLYDCNWIVVNCSTPGNFFHVLRRQILLPFRKP----- 853
Mm_E1k          CNDDPDVLP---DLQEENFDINQLYDCNWIVVNCSTPGNFFHVLRRQILLPFRKP----- 853
Hs_E1k          CNDDPDVLP---DLKEANFDINQLYDCNWVVVNCSTPGNFFHVLRRQILLPFRKP----- 853
Rn_E1kL         SNDDSDAYP---VFTED-FEVSQLYDCNWIVVNCSTPASYFHVLRRQVLLPFRKPGWMWG 844
Mm_E1kL         SNDDSDAYP---VFTED-FEVSQLYDCNWIVVNCSTPASYFHVLRRQILLPFRKP----- 839
Hs_E1kL         SNDDSDAYP---AFTKD-FEVSQLYDCNWIVVNCSTPANYFHVLRRQILLPFRKP----- 839
Rn_DHTKD1       CDSAEEGVD---SDTVN----------MFVVHPTTPAQYFHLLRRQMMRNFRKP----- 760
Mm_DHTKD1       CDSAEEGVD---SDTVN----------MFVVHPTTPAQYFHLLRRQMIRNFRKP----- 761
Hs_DHTKD1       CDSAEEGVD---GDTVN----------MFVVHPTTPAQYFHLLRRQMVRNFRKP----- 760
Lm_E1kB         SDDSDRVPSDFRHFPNDQALEIRIRRHNWQVTYPSTPANYFHLLRRQGLREFAKP----- 835
Lm_E1kA         SSEDVTTP--------AYSKEERAHRINWEITYPSTPAQYFHLLRRHQKRNFRKA----- 856
                ...                          :.  :**..:**:***:    * *.

Rn_E1k          --------------LIVFTPKSLLRHPEARTSFDEMLPGTHFQRVIPEDGPAAQNPDKVK 899
Mm_E1k          --------------LIVFTPKSLLRHPEARTSFDEMLPGTHFQRVIPENGPAAQDPHKVK 899
Hs_E1k          --------------LIIFTPKSLLRHPEARSSFDEMLPGTHFQRVIPEDGPAAQNPENVK 899
Rn_E1kL         PIDGAPGGWLFAFQLIVFTPKSLLRHPDAKSSFDQMVSGTSFQRMIPEDGPAAQSPERVE 904
Mm_E1kL         --------------LIVFTPKSLLRHPDAKSSFDQMVSGTSFQRLIPEDGPAAHSPEQVQ 885
Hs_E1kL         --------------LIIFTPKSLLRHPEAKSSFDQMVSGTSFQRVIPEDGAAARAPEQVQ 885
Rn_DHTKD1       --------------LIVASPKMLLRYPVAVSTLEEMAPGTAFKPVIGD---SSVDPKNVK 803
Mm_DHTKD1       --------------LIVASPKMLLRYPAAVSTLEEMAPGTAFKPVIGD---SSVDPKNVK 804
Hs_DHTKD1       --------------LIVASPKMLLRLPAAVSTLQEMAPGTTFNPVIGD---SSVDPKKVK 803
Lm_E1kB         --------------LINLFSKARLRAPN-LSKLSDMTQGTSFKAVIDT---ARAKDTVAR 877
Lm_E1kA         --------------LVIFFSKKYLRAPN-VSTLEELTSG-EFQPVIPD---LSVPASQAR 897
                              *:   .*  ** *   :.:.::  *  *: :*           ..

Rn_E1k          RLLFCTGKVYYDLTRER----KARDMAE-EVAITRIEQLSPFPFDLLLKEAQKYPN---- 950
Mm_E1k          RLLFCTGKVYYDLTRER----KARNMEE-EVAITRIEQLSPFPFDLLLKEAQKYPN---- 950
Hs_E1k          RLLFCTGKVYYDLTRER----KARDMVG-QVAITRIEQLSPFPFDLLLKEVQKYPN---- 950
Rn_E1kL         RLIFCTGKVYYDLVKER----SSQGLEK-QVAITRLEQISPFPFDLIMREAEKYSG---- 955
Mm_E1kL         RLIFCTGKVYYDLVKER----SSQGLEQ-QVAITRLEQISPFPFDLIMREAEKYSG---- 936
Hs_E1kL         RLIFCTGKVYYDLVKER----SSQDLEE-KVAITRLEQISPFPFDLIKQEAEKYPG---- 936
Rn_DHTKD1       TLIFCSGKHFYALLKQR----ESLGAKKRDFAIIRLEELCPFPLDSLQQEMGKYKHV--- 856
Mm_DHTKD1       TLIFCSGKHFYALLKQR----ESLGTKKHDFAIIRLEELCPFPLDALQQEMSKYKHV--- 857
Hs_DHTKD1       TLVFCSGKHFYSLVKQR----ESLGAKKHDFAIIRVEELCPFPLDSLQQEMSKYKHV--- 856
```

```
Lm_E1kB          KVVFCSGQIESIVNDAKTAMQKETPGVHDDVVLVTVEQLAPFPWEEVADVMEKYAQRNPN 937
Lm_E1kA          RLVMCTGQIYHYLNKYR-----ETKGVKD-VALVRVEELSPFPVAEVQQLLAEYEK---- 947
                  :::*:*:    :   :             ..:  :*::.***   :     :*

Rn_E1k           AELAWCQEEHKNQGYYDYVKPRLRTTIDRAK----PVWYAGRDPAAAPATGNKKTHLTEL 1006
Mm_E1k           AELAWCQEEHKNQGYYDYVKPRLRTTIDRAK----PVWYAGRDPAAAPATGNKKTHLTEL 1006
Hs_E1k           AELAWCQEEHKNQGYYDYVKPRLRTTISRAK----PVWYAGRDPAAAPATGNKKTHLTEL 1006
Rn_E1kL          AELVWCQEEHKNMGYYDYISPRFMTLLGHSR----PIWYVGREPAAAPATGNKNTHLVSL 1011
Mm_E1kL          AELVWCQEEHKNMGYYDYISPRFMTLLGHSR----PIWYVGRDPAAAPATGNKNAHLVSL 992
Hs_E1kL          AELAWCQEEHKNMGYYDYISPRFMTILRRAR----PIWYVGRDPAAAPATGNRNTHLVSL 992
Rn_DHTKD1        QDIIWSQEEPQNMGPWSFVYPRFEKQLACK------LRLVSRPPLPAPAVGIGTVHQQQH 910
Mm_DHTKD1        RDVIWSQEEPQNMGPWSFVSPRFEKQLACR------LRLVSRPPLPAPAVGIGTVHQQQH 911
Hs_DHTKD1        KDHIWSQEEPQNMGPWSFVSPRFEKQLACK------LRLVGRPPLPVPAVGIGTVHLHQH 910
Lm_E1kB          TEFVWLQEEPRNMGMWTHMRPRMNSLMRHLGLKQTRVNVVSRPSAASPSTGYGSVHVAEE 997
Lm_E1kA          AELMWAQEEPKNMGSWAHVEPRIEDYTK----GERELRYAGRSITAAPSTGYKSKHEKEQ 1003
                  :  * *** :* * : .: **:            :  ..*   . *:.*  . *  .

Rn_E1k           QRFLDTAFDLDAFKKFS-                                          1023
Mm_E1k           QRFLDTAFDLDAFKKFS-                                          1023
Hs_E1k           QRLLDTAFDLDVFKNFS-                                          1023
Rn_E1kL          RKFLDTAFNLKAFEGKTF                                          1029
Mm_E1kL          RRFLDTAFNLKAFEGKTF                                          1010
Hs_E1kL          KKFLDTAFNLQAFEGKTF                                          1010
Rn_DHTKD1        EAILFKTFTS--------                                          920
Mm_DHTKD1        EDILSKTFTQ--------                                          921
Hs_DHTKD1        EDILAKTFA---------                                          919
Lm_E1kB          KKLIAETLA---------                                          1006
Lm_E1kA          EIICEMVFH---------                                          1012
                 . :   .:
```

Figure 7.1 Alignment of E1k, E1k-like and DHTKD1 protein sequences
ClustalW alignment of *L. major* E1k-A and E1k-B (*Lm*E1k-A and *Lm*E1k-B) (see Table 3.1) with: *R. norvegicus* E1k (*Rn*E1k) (accession number NP_001017461); *M. musculus* E1k (*Mm*E1k) (accession number CAI24405); *H. sapiens* E1k (*Hs*E1k) (accession number NP_002532); *R. norvegicus* E1k-like (*Rn*E1kL) (accession number NP_001099532); *M. musculus* E1k-like (*Mm*E1kL) (accession number AAI57971); *H. sapiens* E1k-like (*Hs*E1kL) (accession number NP_060715); *R. norvegicus* DHTKD1 (*Rn*DHTKD1) (accession number NP_001099532); *M. musculus* DHTKD1 (*Mm*DHTKD1) (accession number AAI57971); *H. sapiens* DHTKD1 (*Hs* DHTKD1) (accession number NP_061176). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues indicate conserved amino acids within different motifs that fulfil various roles within E1k enzymes. Blue residues and green residues are those that are conserved in E1k/E1kL proteins but not in DHTKD1 or *Lm*E1kA/*Lm*E1kB, respectively. Refer also to Table 7.1.

	Signature specific motif				
Function[a]	**E1k consensus**[a]	**Hs E1k**[a]	**Hs DHTKD1**[a]	**Lm E1kA**[b]	**Lm E1kB**[b]
Lipoyl domain complementarity	KRF(G/s)(L/i)EG	KRFGLEG	KRYGGEG	KRFGCDG	HRFGLDG
2-Oxo substrate side chain binding pocket ThDP interaction	GM(A/p)HRGRLN	GMPHRGRLN	GMPHRGRLN	GMAHRGRLN	GMPHRGRLN
2-Oxo substrate side chain binding pocket	GD(V/m)KYH(L/m,q)G	GDVKYHLG	GDVLSHLT	SDVKYHLG	GDVEYHLG
Stabilization of ThDP binding Shielding S of ES from the solvent	NPSHLE	NPSHLE	NPSHLE	NPSHLE	NPSHLE
Binding of ThDP pyrophosphate group	G(Q/e)G	GQG	GQG	GQG	GQG
Binding of ThDP pyrophosphate group and thiazole ring Hydrogen bond network	GFTT	GFTT	GYTT	GFTT	GFTT
Interaction with ThDP through loop 1 Catalysis of decarboxylation and reductive acylation Lipoyl domain interaction	GHNExD(E/q)PxxTQ	GHNEMDEPMFTQ	GHNELDEPFYTN	GHNENDDPSITQ	GHNETDLPDFTQ
Lipoyl domain complementarity	Wxx(A/g)E	WALAE	WATAE	WGTAE	WCQAE
Lipoyl domain complementarity	G(Q/e)DxxRGTF	GQDVERGTF	GQDVGRGTF	GEDVERGTF	GEDVERGTF
ThDP binding	WEAQ(F/y)GDF	WEAQFGDF	WEAQFGDF	WEAQYGDF	WEAQFGDF
Catalysis Shielding leaving carboxylate from solvent Lipoyl domain interaction	LPHG(Y/m)(E/d)GxGPEHSS	LPHGMEGMGPEHSS	LPHGYDGAGPDHSS	LPHGYDGKGAEHSS	LPHGYSGAGPEHSS
Lipoyl domain complementarity	LLR	LLR	LLR	YLR	RLR
Ca2+ binding	DxDxDx	DADLDS	QALLEN	DTDVTE	HADNCI
	ExDxDx	ESDLDK	-	FSDKDL	?
	NDDxDx	NDDPDV	MCDSAE?	SEDVTT	SDDSDR

Table 7.1 Summary of conserved motifs in *Lm*E1k-A and *Lm*E1k-B
This table summarises which motifs contribute to specific activities within E1k enzymes (Bunik & Degtyarev, 2008)[a]. Based upon ClustalW alignment (see Figure 7.1), conservation of sequences comprising these motifs was determined for *Lm*E1k-A and *Lm*E1k-B[b]. Blue residues and green residues are those that are conserved in E1k/E1kL motifs in other species but not in *Hs*DHTKD1 or *Lm*E1kA/*Lm*E1kB, respectively.

```
Hs_E1p-α          MRKMLAAVSRVLSGASQKPASRVLVASRNFANDATFEIKKCDLHRLEEGPPVTTVLTRED 60
Rn_E1p-α          MRKMLAAVSHVLAGAAQKPASRVLVASRNFANDATFEIKKCDLHRLEEGPPVTTVLTRED 60
Lm_E1p-α          ---MFKCATRCLLDTKTVP----LKPQR------PFKLHTAGRTDMAP-LPTQAVYDAEQ 46
                     *: ..:: * .:   *    * ..*      .*::::..   :    *. :*   *:

Hs_E1p-α          GLKYYRMMQTVRRMELKADQLYKQKIIRGFCHLCDGQEACCVGLEAGINPTDHLITAYRA 120
Rn_E1p-α          GLKYYRMMQTVRRMELKADQLYKQKIIRGFCHLCDGQEACCVGLEAGINPTDHLITAYRA 120
Lm_E1p-α          LKQSLALMFRIRRMESLCDQSYKLKKIRGFCHLYIGQEAIPAGMENVLTFEDPIITGYRD 106
                    :   :*  :**** .** ** * *******  ****  .*:*  :.  * :**.**

Hs_E1p-α          HGFTFTRGLSVREILAELTGRKGGCAKGKGGSMHMYAKN--FYGGNGIVGAQVPLGAGIA 178
Rn_E1p-α          HGFTFTRGLPVRAILAELTGRRGGCAKGKGGSMHMYAKN--FYGGNGIVGAQVPLGAGIA 178
Lm_E1p-α          HGWYISRGGKPEDVFAEMFGRQGGCSKGKGGSMHMYRVDNGFYGGNGIVGAQVSIGAGLA 166
                  **: ::**   . ::**: **:***:********** :  ***********.:***:*

Hs_E1p-α          LACKYNGKD---EVCLTLYGDGAANQGQIFEAYNMAALWKLPCIFICENNRYGMGTSVER 235
Rn_E1p-α          LACKYNGKD---EVCLTLYGDGAANQGQIFEAYNMAALWKLPCIFICENNRYGMGTSVER 235
Lm_E1p-α          WRFAMENRDSPKHVAVTFYGDGAANQGQIYESMNIAALQRLPVIFAVENNHFGMGTSAAR 226
                       :.:*   .*.:*:**********:*: *:*** :** **  ***::*****. *

Hs_E1p-α          AAASTDYYKRGDFIPGLRVDGMDILCVREATRFAAAYCRSGKGPILMELQTYRYHGHSMS 295
Rn_E1p-α          AAASTDYYKRGDFIPGLRVDGMDILCVREATKFAAAYCRSGKGPILMELQTYRYHGHSMS 295
Lm_E1p-α          GSYQAEFYRRGDYIPGIKVDGMDVLAVQEGTRYARDYCMSGKGPIVMELDCYRYMGHSMS 286
                  .: .:::*:***:***::*****:*.*:*.*::*  ** ******:***: *** *****

Hs_E1p-α          DPGVSYRTREEIQEVRSKSDPIMLLKDRMVNSNLASVEELKEIDVEVRKEIEDAAQFATA 355
Rn_E1p-α          DPGVSYRTREEIQEVRSKSDPIMLLKDRMVNSNLASVEELKEIDVEVRKEIEDAAQFATA 355
Lm_E1p-α          DPDNQYRTKSDIQHVKQERDCIRKMREFMATEGIMTEDEMSKMEKDVKKEVDQDLQKAQK 346
                  **. .***:.:**.*:.: * *  ::: *....: : :*:.::: :*:**:::  * *

Hs_E1p-α          DPEPPLEELGYHIYSSDPPFEVRGANQWIKFKSVS                          390
Rn_E1p-α          DPEPPLEELGYHIYSSDPPFEVRGANQWIKFKSVS                          390
Lm_E1p-α          QPMTKLDELFTDIYVG-EQYEHRTCQGTVYHKP--                          378
                  :* . *:**  .** .   :* * .:  : .*.
```

Figure 7.2 Alignment of E1p-α protein sequences
ClustalW alignment of *L. major* E1p-α (*Lm*E1p-α) (see Table 3.1) with homologues in *R. norvegicus* (*Rn*E1p-α) (accession number CAA78146) and *H. sapiens* (*Hs*E1p-α) (accession number NP_000275). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues indicate conserved amino acids within different motifs that fulfil various roles within E1p-α enzymes. Green residues are those that are conserved in E1p-α proteins found in other species but not in *Lm*E1p-α. Refer also to Table 7.2.

```
Hs_E1p-β        MAAVSGLVRRPLREVSGLLKRRFHWTAPAALQVTVRDAINQGMDEELERDEKVFLLGEEV 60
Rn_E1p-β        MAAVAGLVRGPLRQASGLLKRRFHRSAPAAVQLTVREAINQGMDEELERDEKVFLLGEEV 60
Lm_E1p-β        ---MRRFASRALFSASAAMAARCATTN-----MTVRDAIHSALDEELAREEKVFVIGEEV 52
                   :  :.  .* ..*. :  *   :      :***:**:..:**** *:****::****

Hs_E1p-β        AQYDGAYKVSRGLWKKYGDKRIIDTPISEMGFAGIAVGAAMAGLRPICEFMTFNFSMQAI 120
Rn_E1p-β        AQYDGAYKVSRGLWKKYGDKRIIDTPISEMGFAGIAVGAAMAGLRPICEFMTFNFSMQAI 120
Lm_E1p-β        AQYQGAYKVTKGLMDKYGKDRIIDMPITEHGFAGMAVGAALSGLRPVCEFMTFNFAMQAI 112
                ***:*****::** .***..**** **:* ****:*****::****:********:****

Hs_E1p-β        DQVINSAAKTYYMSGGLQPVPIVFRGPNGASAGVAAQHSQCFAAWYGHCPGLKVVSPWNS 180
Rn_E1p-β        DQVINSAAKTYYMSAGLQPVPIVFRGPNGASAGVAAQHSQCFAAWYGHCPGLKVVSPWNS 180
Lm_E1p-β        DQLVNSAGKSLYMSGGQMKCPIVFRGPNGASAGVGAQHSQCFGPWYASVPGLKVIAPYNC 172
                **::***.*: ***.*    **************.*******..**.  *****::*:*.

Hs_E1p-β        EDAKGLIKSAIRDNNPVVVLENELMYGVPFEFLPEAQSKDFLIPIGKAKIERQGTHITVV 240
Rn_E1p-β        EDAKGLIKSAIRDDNPVVMLENELMYGVAFELPTEAQSKDFLIPIGKAKIERQGTHITVV 240
Lm_E1p-β        EDARGMIKAAIRDDNAVVVLEHELLYSESFPVTDEAADKNFVIPFGKAKIEREGKDITLI 232
                ***:*:**:****:*.**:**:**:*. .* .  ** .*:*:**:*******:*..**::

Hs_E1p-β        SHSRPVGHCLEAAAVLSKEGVECEVINMRTIRPMDMETIEASVMKTNHLVTVEGGWPQFG 300
Rn_E1p-β        AHSRPVGHCLEAAAVLSKEGIECEVINLRTIRPMDIEAIEASVMKTNHLVTVEGGWPQFG 300
Lm_E1p-β        GFSRGVDLCLKAAEKLAAEGVQAEVINLRSLRPLDRHTILSSIKKTHRAVTVDESFPVCN 292
                ..** *. **:**  *: **::.****:*::**:* .:* :*: **:: ***: .:*  .

Hs_E1p-β        VGAEICARIMEGPAFNFLDAPAVRVTGADVPMPYAKILEDNSIPQVKDIIFAIKKTLNI 359
Rn_E1p-β        VGAEICARIMEGPAFNFLDAPAVRVTGADVPMPYAKILEDNSIPQVKDIIFAIKKTLNI 359
Lm_E1p-β        IGAEICACVMESDTFDYLDAPIERVSCADCPTPYSKDIEMASQPQVADVMAAAKRVLS- 350
                :****** :**. :*::****  **: ** * **:* :*  * *** *:: * *:.*.
```

Figure 7.3 Alignment of E1p-β protein sequences
ClustalW alignment of *L. major* E1p-β (*Lm*E1p-β) (see Table 3.1) with homologues in *R. norvegicus* *Rn*E1p-β) (accession number NP_001007621) and *H. sapiens* (*Hs*E1p-β) (accession number NP_000916). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues indicate conserved amino acids within different motifs that fulfil various roles within E1p-α enzymes. Green residues are those that are conserved in E1p-β proteins found in other species but not in *Lm*E1p-β. Refer also to Table 7.2.

	Signature specific motif				
Function[a]	**E1p consensus**[a]	**Hs E1p-alpha**[a]	**Lm E1p-alpha**[b]	**Hs E1p-beta**[a]	**Lm E1p-beta**[b]
Lipoyl domain complementarity	IRGFCHL	IRGFCHL	IRGFCHL	-	-
2-Oxo substrate side chain binding pocket ThDP interaction	IT(A/s)YR	ITAYR	IT**G**YR	-	-
2-Oxo substrate side chain binding pocket	GKGGSMH	GKGGSMH	GKGGSMH	-	-
Stabilization of ThDP binding Shielding S of ES from the solvent	FYGG(N/h)GIVGAQ	FYGGNGIVGAQ	FYGGNGIVGAQ	-	-
Binding of ThDP pyrophosphate group	NQGQ	NQGQ	NQGQ	-	-
Binding of ThDP pyrophosphate group and thiazole ring Hydrogen bond network	GMGT	GMGT	GMGT	-	-
Interaction with ThDP through loop 1 Catalysis of decarboxylation and reductive acylation Lipoyl domain interaction	TYRY(H/g)GHSMSDPG	TYRYHGHSMSDPG	**C**YRY**M**GHSMSDP**D**	-	-
Lipoyl domain complementarity	TVR(D/e)A(L/i)N	-	-	TVRDAIN	TVRDAI**H**
Lipoyl domain complementarity	GEEVxQYxGAYK	-	-	GEEVAQYDGAYK	GEEVAQYQGAYK
ThDP binding	EFM(T/s)FNFSMQAID	-	-	EFMTFNFSMQAID	EFMTFNF**A**MQAID
Catalysis Shielding leaving carboxylate from solvent Lipoyl domain interaction	VFRGPNGAxxGVxAQHSQ	-	-	VFRGPNGASAGVAAQHSQ	VFRGPNGASAGVGAQHSQ
Lipoyl domain complementarity	(L/m)YG	-	-	MYG	LY**S**

Table 7.2 Summary of conserved motifs in *Lm*E1p-α and *Lm*E1p-β
This table summarises which motifs contribute to specific activities within E1p enzymes (Bunik & Degtyarev, 2008)[a]. Based upon ClustalW alignments (see Figure 7.2 and Figure 7.3), conservation of sequences comprising these motifs was determined for *Lm*E1p-αand *Lm*E1p-β[b]. Green residues are those that are conserved in E1p-α/E1p-β motifs in other species but not in *Lm*E1p-αand *Lm*E1p-β.

```
Hs_E1b-α          ------------MAVAIAAARVWRLNRGLSQAALLLLRQPG---ARGLARSHPPRQQQQ- 44
Rn_E1b-α          ------------MAVAMSAAKIWRPSRGLRQAALLLLGRPG---ARGLARFHPSRQQQQQ 45
Lm_E1b-α          MFHISRAVRCNALATAIAGRTSLSDAIRQVQKVWNLDFKDGPVITSTLAFNDEPDPAAPI 60
                              :*.*::.           * .  *  : *   :  **  . .

Hs_E1b-α          FSSLDDK-PQFPGASAEFIDKLEFIQPNVISGIPIYRVMDRQGQIINPSEDPH--LPKEK 101
Rn_E1b-α          FPSLDDK-PQFPGASAEFVDKLEFIQPNVISGIPIYRVMDRQGQIINPSEDPH--LPQEE 102
Lm_E1b-α          FHVLDLQGRVFVEDKADSRAAAATPSTVNEAGKMESQSADVGEVFRYHAEDEMSVITREV 120
                  *  ** :   *   .*:        ..   :*    :  *    :   :**    :.:*

Hs_E1b-α          VLKLYKSMTLLNTMDRILYESQRQGRISFYMTNYGEEGTHVGSAAALDNTDLVFGQYREA 161
Rn_E1b-α          VLKLYRSMTLLNTMDRILYESQRQGRISFYMTNYGEEGTHVGSAAALERTDLVFGQYREA 162
Lm_E1b-α          AQGMMSAMLTHNTMDKIMLEAQRQGRISFYMTMFGEEAAVIGAAAGLASNDELFAQYREA 180
                  .  :  :*   ****:*: *:*********** :***.: :*:**.*  .* :*.*****

Hs_E1b-α          GVLMYRDYPLELFMAQCYGNISDLGKGRQMPVHYGCKERHFVTISSPLATQIPQAVGAAY 221
Rn_E1b-α          GVLMYRDYPLELFMAQCYGNVSDPGKGRQMPVHYGCKERHFVTISSPLATQIPQAVGAAY 222
Lm_E1b-α          GILTYRGYTIPEFIAQCMGNCECDAKGRQMPIHYGSKRLHAQMVSSPLATQIPHGAGAGY 240
                  *:* **.*.:  *:*** ** .  .******:***.*. *   :*********:..**.*

Hs_E1b-α          AAKRAN----------------ANRVVICYFGEGAASEGDAHAGFNFAATLECPIIFFC 264
Rn_E1b-α          AAKRAN----------------ANQIVICYFGEGAASEGDAHAGFNFAATLECPIIFFC 265
Lm_E1b-α          AFRLENQALERRLPAGTLLSTIPEARICATFFGEGAASEGDFHAGLNFASTVGSHTLFFV 300
                  * :  *                  ::   :********** ***:***:*: .  :**

Hs_E1b-α          RNNGYAISTPTSEQYRGDGIAARGPGYGIMSIRVDGNDVFAVYNATKEARRRAVAENQPF 324
Rn_E1b-α          RNNGYAISTPTSEQYRGDGIAARGPGYGIMSIRVDGNDVFAVYNATKEARRRAVAENQPF 325
Lm_E1b-α          RNNGYAISTPTHSQYMGDGILSRAVGYGIPAARVDGLDALAVYHTVRKAREMILNSHRPV 360
                  ********** .** **** :*. **** : **** *.:***::.::**.  : .::*.

Hs_E1b-α          LIEAMTYRIGHHSTSDDSSAYRSVDEVNYWDKQDHPISRLRHYLLSQGWWDEEQEKAWRK 384
Rn_E1b-α          LIEAMTYRIGHHSTSDDSSAYRSVDEVNYWDKQDHPISRLRQYLLNQGWWDEEQEKAWRK 385
Lm_E1b-α          LVEALTYRLSHHSTSDDSTAYRSRDEIEHFAETFSPIERFEHFVTARGWWTPEQSREVVE 420
                  *:**:***:.*******:**** **:::: :   **.*:.:::  :***  **.:   :

Hs_E1b-α          QSRRKVMEAFEQAERKPKPNPNLLFSDVYQEMPAQLRKQQESLARHLQTYGEHYPLDHFD 444
Rn_E1b-α          QSRKKVMEAFEQAERKLKPNPSLLFSDVYQEMPAQLRRQQESLARHLQTYGEHYPLDHFD 445
Lm_E1b-α          RTRSEVLSELRRQEKLPAWPVSTLCDDVFEHLTPELERQRTQLVEHYQAHRSIYDQEKL- 479
                  ::* :*:. :.: *:      . * .**::.:..:*.:*: .*..* *:: . *  :::

Hs_E1b-α          K                                                        445
Rn_E1b-α          K                                                        446
Lm_E1b-α          -
```

Figure 7.4 Alignment of E1b-α protein sequences
ClustalW alignment of *L. major* E1b-α (*Lm*E1b-α) (see Table 3.1) with homologues in *R. norvegicus* (*Rn*E1b-α) (accession number NP_036914) and *H. sapiens* (*Hs*E1b-α) (accession number NP_000700). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues indicate conserved amino acids within different motifs that fulfil various roles within E1b-α enzymes. Green residues are those that are conserved in E1b-α proteins found in other species but not in *Lm*E1b-α. Refer also to Table 7.3.

Figure 7.5 Alignment of E1b-β protein sequences
ClustalW alignment of *L. major* E1b-β (*Lm*E1b-β) (see Table 3.1) with homologues in *R. norvegicus* (*Rn*E1b-β) (accession number NP_062140) and *H. sapiens* (*Hs*E1b-β) (accession number NP_000047). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues indicate conserved amino acids within different motifs that fulfil various roles within E1b-β enzymes. Green residues are those that are conserved in E1b-β proteins found in other species but not in *Lm*E1b-β. Refer also to Table 7.3.

		Signature specific motif			
Function[a]	**E1b consensus**[a]	**Hs E1b-alpha**[a]	**Lm E1b-alpha**[b]	**Hs E1b-beta**[a]	**Lm E1b-beta**[b]
Lipoyl domain complementarity	-	-	-	-	-
2-Oxo substrate side chain binding pocket ThDP interaction	YR(D/e,q)	YRE	YRE	-	-
2-Oxo substrate side chain binding pocket	(K/s)G(G)RQ(M/l)P	KGRQMP	KGRQMP	-	-
Stabilization of ThDP binding Shielding S of ES from the solvent	-	-	-	-	-
Binding of ThDP pyrophosphate group	(N/t)FA	NFA	NFA	-	-
Binding of ThDP pyrophosphate group and thiazole ring Hydrogen bond network	AIS	AIS	AIS	-	-
Interaction with ThDP through loop 1 Catalysis of decarboxylation and reductive acylation Lipoyl domain interaction	(P/h)H(S/t)xx(G)DD(P/d,s)	HHSTSDDS	HHSTSDDS	-	-
Lipoyl domain complementarity	(Q/d)(A/s)xxxA	-	-	QSVTSA	QAVNS**G**
Lipoyl domain complementarity	G(E/q)DVxxxGGVF	-	-	GEDVAFGGVF	GEDVAFGGVF
ThDP binding	E(I/m)QF	-	-	EIQF	E**V**QF
Lipoyl domain complementarity	(L/a)Y(R/n)	-	-	LYR	LYR

Table 7.3 Summary of conserved motifs in *Lm*E1b-α and *Lm*E1b-β
This table summarises which motifs contribute to specific activities within E1b enzymes (Bunik & Degtyarev, 2008)[a]. Based upon ClustalW alignments (see Figure 7.4 and Figure 7.5), conservation of sequences comprising these motifs was determined for *Lm*E1b-αand *Lm*E1b-β[b]. Green residues are those that are conserved in E1b-α/E1b-β motifs in other species but not in *Lm*E1b-αand *Lm*E1b-β.

```
Ec_P-protein      -------------------------------------------------MTQTLSQLE-NS 11
Hs_P-protein      MQSCARAWGLRLGRGVGGGRRLAGGSGPCWAPRSRDSSSGGGDSAAAGASRLLERLLPRH 60
Lm_P-protein      --VSALPSGFTP--------------PPFLSPACARMLRRLLRVHGVPAPAGLARYT-ST 43
                                                                 .  * :

Ec_P-protein      GAFIERHIGPDAAQQQEMLNAVGAQSLNALTGQIVPKDIQLATPPQVGAPATEYAALAEL 71
Hs_P-protein      DDFARRHIGPGDKDQREMLQTLGLASIDELIEKTVPANIRLKRPLKMEDPVCENEILATL 120
Lm_P-protein      DAYLNRHIGPTRKETAEMLKTVGKESLADLMTTVLPSDI-LRTPLNNFKCLSETAALSYL 102
                  . : .*****   :  ***:::*  *:  *    :* :* *  * :      *   *: *

Ec_P-protein      KAIASRNKRFTSYIGMGYTAVQLPPVILRNMLENPGWYTAYTPYQPEVSQGRLEALLNFQ 131
Hs_P-protein      HAISSKNQIWRSYIGMGYYNCSVPQTILRNLLENSGWITQYTPYQPEVSQGRLESLLNYQ 180
Lm_P-protein      KSLGAQNKVLKSMIGQGYYECIVPSAIMRNVLENPMWYTPYTPFQSEIAQGRLESLLNFQ 162
                  :::.::*:   * ** **    :* .*:**:***. * * ***:*.*::*****:***:*

Ec_P-protein      QVTLDLTGLDMASASLLDEATAAAEAMAMAKRVSKLKNANRFFVASDVHPQTLDVVRTRA 191
Hs_P-protein      TMVCDITGLDMANASLLDEGTAAAEALQLCYRHNKRR---KFLVDPRCHPQTIAVVQTRA 237
Lm_P-protein      TMVTDLTKMDISNASLLDQATAAGECLYLALNQHRHKR-RKFFVSRDVFLSSIEMIRTRA 221
                   :. *:* :*::.*****:.***.*.: :. .  : :   :*:*    . .:: :::***

Ec_P-protein      ETFG-FEVIVDDAQKVLDHQDVFGVLLQQVGTTGEIHDYTALISELKSRKIVVSVAADIM 250
Hs_P-protein      KYTGVLTELKLPCEMDFSGKDVSGVLFQYPDTEGKVEDFTELVERAHQSGSLACCATDLL 297
Lm_P-protein      HPLGAQVIVGDVQSLDLDDAELSGIFVQTPDAKGELHDFTTIFARAKANGVVCCAGVDLM 281
                  .  *    :    .  :.  :: *::.*  .: *::.*:* :. . :    : . ..*::

Ec_P-protein      ALVLLTAPGKQGADIVFGSAQRFGVPMGYGGPHAAFFAAKDEYKRSMPGRIIGVSKDAAG 310
Hs_P-protein      ALCILRPPGEFGVDIALGSSQRFGVPLGYGGPHAAFFAVRESLVRMMPGRMVGVTRDATG 357
Lm_P-protein      ASCLVKPAGEMGADVVVGCAQRFGTPLGYGGPHAAFMATTDNLKRLSPGRIVGISKDNAG 341
                  *  :: ..*: *.*:..*.:****.*:*********:*. :.  *  ***::*:::* :*

Ec_P-protein      NTALRMAMQTREQHIRREKANSNICTSQVLLANIASLYAVYHGPVGLKRIANRIHRLTDI 370
Hs_P-protein      KEVYRLALQTREQHIRRDKATSNICTAQALLANMAAMFAIYHGSHGLEHIARRVHNATLI 417
Lm_P-protein      DPAIRVALQTREQHIKRERATSNICTAQALLANMNAFYAIYHGPEGLKQLAREIHQKAKL 401
                  . . *:*:*******:*::*.*****:*.****: :::*:***. **:::*..:*. : :

Ec_P-protein      LAAGLQQKGLKLRHAHYFDTLCVEV----ADKAGVLTRAEAAEINLRSD-ILNAVGITLD 425
Hs_P-protein      LSEGLKRAGHQLQHDLFFDTLKIQCG---CSVKEVLGRAAQRQINFRLF-EDGTLGISLD 473
Lm_P-protein      FAVGMESLGFSPVNTTYFDTLSFSMEAAPMTAADYAQRCVERGINLFVDGSTNQVSVSLD 461
                  :: *::  * .  :  :**** ..             *.    **:      . :.::**

Ec_P-protein      ETTTRENVMQLFNVLLGDNHGLDIDTLDKDVAHDSRSIQPAMLR-DDEILTHPVFNRYHS 484
Hs_P-protein      ETVNEKDLDDLLWIFGCESS---AELVAESMGEECRGIPGSVFKRTSPFLTHQVFNSYHS 530
Lm_P-protein      EATTEQHIAALLQAAGMPTP--KIEALTR-VADTICVIPEALLR-KSKFLQSTVFNSHKS 517
                  *:...:.:  *:      .     : : . :..    *  ::::  . :*   *** ::*

Ec_P-protein      ETEMMRYMHSLERKDLALNQAMIPLGSCTMKLNAAAEMIPITWPEFAELHPFCPPEQAEG 544
Hs_P-protein      ETNIVRYMKKLENKDISLVHSMIPLGSCTMKLNSSSELAPITWKEFANIHPFVPLDQAQG 590
Lm_P-protein      ETELMRYAQHLQRKDYGLTHGMIPLGSCTMKLNSAAAMRALSWPEYTALHPYAPEDQARG 577
                  **:::** : *:.** .* :.************::: : .::* *:: :**: * :**.*

Ec_P-protein      YQQMIAQLADWLVKLTGYDAVCMQPNSGAQGEYAGLLAIRHYHESRNEGHRDICLIPASA 604
Hs_P-protein      YQQLFRELEKDLCELTGYDQVCFQPNSGAQGEYAGLATIRAYLNQKGEGHRTVCLIPKSA 650
Lm_P-protein      YHTLLADLKQKLCDITGMAACSIQPNSGAQGEYAGLRIIRAYHESRGEAHRDVCFIPISA 637
                  *: :: :* . * .:**    .:*************  ** * :.:.*.** :*:** **

Ec_P-protein      HGTNPASAHMAGMQVVVVACDKNGNIDLTDLRAKAEQAGDNLSCIMVTYPSTHGVYEETI 664
Hs_P-protein      HGTNPASAHMAGMKIQPVEVDKYGNIDAVHLKAMVDKHKENLAAIMITYPSTNGVFEENI 710
Lm_P-protein      HGTNPASAVLAGLKVVTVKCLDDGSVDMVDLETKCVKHARDLACLMITYPSTYGLYDQNI 697
                  ******** :**:::  *   . *.:* ..*.:   :   :*:.:*:***** *::::.*

Ec_P-protein      REVCEVVHQFGGQVYLDGANMNAQVGITSPGFIGADVSHLNLHKTFCIPHGGGGPGMGPI 724
Hs_P-protein      SDVCDLIHQHGGQVYLDGANMNAQVGICRPGDFGSDVSHLNLHKTFCIPHGGGGPGMGPI 770
Lm_P-protein      RKITSMVHEHGGQCYIDGANLNALVGYTGPGFIGGDVCHINMHKTFSIPHGGGGPGLGPI 757
                   .: .::*:.*** *:****:** **   ** :*.**.*:*:****.*********:***

Ec_P-protein      GVKAHLAPFVPGHSVVQIEG--MLTRQGAVSAAPFGSASILPISWMYIRMMGAEGLKKAS 782
Hs_P-protein      GVKKHLAPFLPNHPVISLKRNEDACPVGTVSAAPWGSSSILPISWAYIKMMGGKGLKQAT 830
Lm_P-protein      TVRPHLAPFLPNSTYGPAVG--GSQAFGQVSQAGNGSASIATISYAFMMMLGSHGLKTCT 815
                   *: *****:*. .            * ** *  **:** .**: :: *:*..*** .:

Ec_P-protein      QVAILNANYIASRLQDAFPVLYTGRDGRVAHECILDIRPLKEETGISELDIAKRLIDYGF 842
Hs_P-protein      ETAILNANYMAKRLETHYRILFRGARGYVGHEFILDTRPFKKSANIEAVDVAKRLQDYGF 890
Lm_P-protein      EYAVLNANYLKKRLEEHYTICFLDHSQFCAHEFILDIRPFKKTAHIDAEDVAKRLIDYGF 875
                  : *:*****: .**:  : : : .     .** *** **:*: : *.  *:**** ****
```

```
Ec_P-protein      HAPTMSFPVAGTLMVEPTESESKVELDRFIDAMLAIRAEIDQVKAGVWPLEDNPLVNAPH 902
Hs_P-protein      HAPTMSWPVAGTLMVEPTESEDKAELDRFCDAMISIRQEIADIEEGRIDPRVNPLKMSPH 950
Lm_P-protein      HAPTLAFPVEGTLMIEPTESESKRELDRLADALISIRREIAAVERGDQPKDNNVLTNAPH 935
                  ****:::** ****:******.* ****: **:::** **  :: *      * *  :**

Ec_P-protein      -IQSELVAEWAHPYSREVAVFPAGVA---DKYWPTVKRLDDVYGDRNLFCSCVPISEYQ- 957
Hs_P-protein      SLTCVTSSHWDRPYSREVAAFPLPFVKPENKFWPTIARIDDIYGDQHLVCTCPPMEVYES 1010
Lm_P-protein      TAKCVTADEWNRPYSRQLAAYPTRHQYR-EKFWPSVGRVDNTYGDRNLMCSCAPLEFY-- 992
                     .    .* :****::*.:*       :*:**:: *:*: ***::*.*:* *:. *

Ec_P-protein      ----------
Hs_P-protein      PFSEQKRASS                                                  1020
Lm_P-protein      ----------
```

Figure 7.6 Alignment of P-protein sequences
ClustalW alignment of *L. major* P-protein (*Lm*P-protein) (see Table 3.1) with homologues in *E. coli* (*Ec*P-protein) (accession number AAC75941) and *H. sapiens* (*Hs*P-protein) (accession number NP_000161). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.).Red residues indicate conserved amino acids within different motifs that fulfil various roles within P-proteins. Green residues are those amino acids that are not conserved in *L. major*, but which are strictly conserved in other species.

```
Hs_E2k          LSRSHCVPGVQHSLSAFQKGNCPLGRHSLPGVSLCQGPGYPNSRKVVINNSVFSVRFFRT 60
Lm_E2k          ----------------FRR----VSTRVLP--TACSAAHNLN-----------LRFC-- 24
Ec_E2k          ------------------------SS------------------------------- 2
                                        .

Hs_E2k          TTVCKYDLVTVKTPAFAEPVTEGDVR-WEKAVGDTVAEDEVVCEIETDKTLVQVPSPANG 119
Lm_E2k          --------LSINVPTIAESISTGKVVNWTKKVGDAVAEDEVICQIESDKLNVDVRAPANG 76
Ec_E2k          --------VDILVPDLPESVADATVATWHKKPGDAVVRDEVLVEIETDKVVLEVPASADG 54
                        : : .* :.*.:: . *  * *  **:*..***: :**:**  ::* :.*:*

Hs_E2k          MIEALFVPDGGKVEGGTPLFTLRKTGAAPAKAKPAEAPAAAAPKAE----------PIA 168
Lm_E2k          VITKINFEEGADVEVGAQLSTMKEGPAPAAAAPKAAEVKLDAPKAE----------PPK 125
Ec_E2k          ILDAVLEDEGTTVTSRQILGRLREGNSAGKETSAKSEEKASTPAQRQQASLEEQNNDALS 114
                ::  :   :*  *     *  :::  :.   :         :*  .           .

Hs_E2k          AAVPPRAAPIPTQMPPVPSP--------------SQPPSSK-PVSAVKPTAVPPLAEPG 212
Lm_E2k          AAAPAASAPA---APAAPAA--------------AAKPAMH-TIAGADPRTKS------ 160
Ec_E2k          PAIRRLLAEHNLDASAIKGTGVGGRLTREDVEKHLAKAPAKESAPAAAAPAAQPALAAR- 173
                .*     *      ..  ..               :  *: . . :.. * : .

Hs_E2k          AGKGLHSEHREKMNRMRQCIAQRLKEAQNTVPMLTIFNEIDVSNIQKMRARHKEAFLKKH 272
Lm_E2k          ----------VRISSMRRRIADRLKASQNTCAMLTTFNEIDMTPLFQLRDKYKDEFHKRH 210
Ec_E2k          ------SEKRVPMTRLRKRVAERLLEAKNSTAMLTTFNEVNMKPIMDLRKQYGEAFEKRH 227
                            :. :*: :*:**  ::*: .*** ***:::. : .:* :: : * *:*

Hs_E2k          NLKLGFMSASVKASAFALQEQPVVNAVIDDITKEVVYRDYIDISVAVATPQGLVVPVIRN 332
Lm_E2k          DVKLGLMSPFVKASAIALKDVPIVNASFG--KDTIDYHEFVDIAIAVATPRGLVVPVIRD 268
Ec_E2k          GIRLGFMSFYVKAVVEALKRYPEVNASID--GDDVVYHNYFDVSMAVSTPRGLVTPVLRD 285
                .::**:**  *** . **:  * *** :.   . : *:::.*:::**:**:***.**:*:

Hs_E2k          VEAMNYADIEQTITELGEKARKNEFAIEDMDGGTFTISNGGVFGSLFEHPLS-TPLSAIL 391
Lm_E2k          VQNMNLANIETAIADYAARARINKLTMAEMTGGTFTISNGGVFGSWMGTPIINPPHSAIL 328
Ec_E2k          VDTLGMADIEKKIKELAVKGRDGKLTVEDLTGGNFTITNGGVFGSLMSTPIINPPQSAIL 345
                *: :. *:**  * : . :.* .:::: :: **.***:******* :  *:  .* ****

Hs_E2k          GMHGIFDKPVAIGGKVEVRPMMYVALTYDHRLIDGREAVTFLRKIKAAVEDPRVLLFDL- 450
Lm_E2k          GMHAIKKKPWVVGNEIKIRDIMAVALTYDHRLIDGSDAVTFLVKVKNLIEDPARMVLDLS 388
Ec_E2k          GMHAIKDRPMAVNGQVEILPMMYLALSYDHRLIDGRESVGFLVTIKELLEDPTRLLLDV- 404
                ***.* .:* .:..::::  :* :**:******** ::* ** .:*  :***  :::*:
```

Figure 7.7 Alignment of E2k protein sequences
ClustalW alignment of *L. major* E2k (*Lm*E2k) (see Table 3.1) with homologues in *E. coli* (*Ec*E2k) (accession number AAA23898) and *H. sapiens* (*Hs*E2k) (accession number NP_000161). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.).Red residues indicate conserved amino acids within different motifs that fulfil various roles within E1k enzymes. Green residues are those that are conserved in E1k proteins found in other species but not in *Lm*E1k.

```
Hs_E2p          MWRVCARRAQNVAPWAGLEARWTALQEVPGTPRVTSRSGPAPARRNSVTTGYGGVRALCG 60
Hs_E3BP         ----------MAASWR----------LGCDPRLLRYLVGFPGRRS-----VGLVKGALG 34
Lm_E2p          ------------------------------------------------------MLRCRAV 7
Ec_E2p          -MAIEIKVPDIGADEVEITEILVKVGDKVEAEQSLITVEGDKASMEVPSPQAGIVKEIKV 59
Lm_E2pL         ------------------------------------------------------MMRRTLL 7
                                                                       ::

Hs_E2p          WTPSSGATPRNRLLLQLLGSPGRRYYSLPPHQKVPLPSLSPTMQAGT------------- 107
Hs_E3BP         WSVSRGANWRWFHSTQWLRGD---------PIKILMPSLSPTMEEGN------------- 72
Lm_E2p          SKLATLAALRFLTITP-----------------IPMPALSPTMEKGK------------- 37
Ec_E2p          SVGDKTQTGALIMIFDSADGAADAAPAQAEEKKEAAPAAAPAAAAAKDVNVPDIGSDEVE 119
Lm_E2pL         WLVNFEP-------------------------VFMPALSPSMETGT------------- 28
                                               *: :*:   ..

Hs_E2p          IARWEKKEGDKINEGDLIAEVETDKATVGFE-SLEECYMAKILVAEGTRDVPIGAIICIT 166
Hs_E3BP         IVKWLKKEGEAVSAGDALCEIETDKAVVTLD-ASDDGILAKIVVEEGSKNIRLGSLIGLI 131
Lm_E2p          ITEWCKQPGDFIRPGDTFCNIETDKAVVSYDNATEEGFFARVITSPG-EETVVGQTVCLI 96
Ec_E2p          VTEILVKVGDKVEAEQSLITVEGDKASMEVP-APFAGTVKEIKVNVGDKVSTGSLIMVFE 178
Lm_E2pL         VVEWKKKIGELVKESDVFCTIQTDKAVVDYTNTFESGYLAKIYCGNGQSAPVAKTIAVMV 88
                :..   : *: :   : :  :: *** :    :     . .:    *           :

Hs_E2p          VGKPEDIEAFKNYTLDSSAAPTPQAAPAPTPAATAS--PPTPSAQAPGSSYPPHMQVLLP 224
Hs_E3BP         VEEGE----------DWKHVEIPKDVGPPPPVSKPS--EPRPSPE-PQISIPVKKEHIPG 178
Lm_E2p          VDEKEGVHS------DEVKNWKPEAEEAPAAAAEEA--PAAPAATTPVAAAPVAAS---- 144
Ec_E2p          VAGEAGAAAP-----AAKQEAAPAAAPAPAAGVKEVNVPDIGGDEVEVTEVMVKVGDKVA 233
Lm_E2pL         SDAADVSKAD----EYTPEGEVPAAEAEAPTAAAVAAAPAAGGASSKAPEGVTCEPVFMP 144
                                    *     ...          .

Hs_E2p          ALSPTMTMGTVQRWEKKVGEKLSEGDLLAEIETDKATIGFEVQEEGYLAKILVPEGTRDV 284
Hs_E3BP         TLRFRLSPAARNILEKHS-------------------------------LDASQGTATG 206
Lm_E2p          GDRVKASPYARKMAAEKN---------------------------------VSLRGIKGT 171
Ec_E2p          AEQSLITVEGDKASMEVPAPFAG----------------------VVKELKVNVGDKVK 270
Lm_E2pL         ALSPSMETGTVVEWKKKIG---------------------------------------- 163
                              :

Hs_E2p          PLGTPLCIIVEKEADISAFADYRPTEVTDLKPQVPPPTPPPVAAVPPTPQPLAPTPSAPC 344
Hs_E3BP         PRG-----IFTKEDALKLVQLKQTGKITESRPTPAP------TATPTAPSPLQAT----- 250
Lm_E2p          GGG------VGRITSKDVAAAVASGTASSAAEVAAP-------AKTAATAALAAP----- 213
Ec_E2p          TGSLIMIFEVEGAAPAAAPAKQEAAAPAPAAKAEAP------AAAPAAKAEGKSEFAEND 324
Lm_E2pL         --------ELVKESDVFCTIQTDKAVVDYTNTFESG----------------------- 191
                        .                       .

Hs_E2p          PATPAGPKGRVFVSPLAKKLAVEKGIDLTQVKGTGPDGRITKKDIDSFVPSKVAPAPAAV 404
Hs_E3BP         ---------------------------------SGPS-----------------YPRPV 259
Lm_E2p          ------------------------------------------------------------
Ec_E2p          AYVHATPLIRRLAREFG----------VNLAKVKGTGRKGRILREDVQAYVKEAIKRAEA 374
Lm_E2pL         ------------------------------------------------------------


Hs_E2p          VPPTGPGMAP----------VPTGVFTDIPISNIRRVIAQRLMQSK-QTIPHYYLSIDVN 453
Hs_E3BP         IPPVSTPGQP----------NAVGTFTEIPASNIRRVIAKRLTESK-STVPHAYATADCD 308
Lm_E2p          AKPAAAKGTP----------PANPNFTDIPVTTMRSVIAKRLHQSKNLEIPHYYLFDDCR 263
Ec_E2p          APAATGGGIPGMLPWPKVDFSKFGEIEEVELGRIQKISGANLSRNW-VMIPHVTHFDKTD 433
Lm_E2pL         --------------------------------YLAKIYCGNGQSAPVAKTIAVMVSDAAD 219
                                               :  :   .

Hs_E2p          MGEVLLVRKELN---KILEGRSKISVNDFIIKASALACLKVPEANSSWM--DTVIRQNHV 508
Hs_E3BP         LGAVLKVRQDL------VKDDIKVSVNDFIIKAAAVTLKQMPDVNVSWD--GEGPKQLPF 360
Lm_E2p          VDNMLALIKQLN---AKGNGEYKITVNDYIVKAVARANTLVPEVNSSWQ--GDFIRQYAT 318
Ec_E2p          ITELEAFRKQQNEEAAKRKLDVKITPVVFIMKAVAAALEQMPRFNSSLSEDGQRLTLKKY 493
Lm_E2pL         VEKVANYYPEDA-----VGGPPASAADPSAAAAAAASARPAPSAASAKH-------YGGS 267
                :  :    :            :      * * :    *    :

Hs_E2p          VDVSVAVSTPAGLITPIVFNAHIKGVETIANDVVSLATKAREGKLQPHEFQG--GTFTIS 566
Hs_E3BP         IDISVAVATDKGLLTPIIKDAAAKGIQEIADSVKALSKKARDGKLLPEEYQG--GSFSIS 418
Lm_E2p          VDVSVAVATPTGLITPIIRNAQAKGLVEISKETKALAKKARDGTLQPSEFQG--GTCSVS 376
Ec_E2p          INIGVAVDTPNGLVVPVFKDVNKKGIIELSRELMTISKKARDGKLTAGEMQG--GCFTIS 551
Lm_E2pL         LDAAVAASGPS---------VARIAAGLETSTLAGIAPSGKGGRFLKSDFSGQPGFDYND 318
                :: .**.             .   .    :     :: ..: * :   : .*  *    .
```

```
Hs_E2p          NLGMFGIKNFSAIINPPQACILAIGASEDKLVPADNEKG---FDVAS----MMSVTLSCD 619
Hs_E3BP         NLGMFGIDEFTAVINPPQACILAVGRFRPVLKLTEDEEGNAKLQQRQ----LITVTMSSD 474
Lm_E2p          NLGATGIPGFTAIINPPQAMILAVGSAKPRAEIVKSEETG-EFEMTGRVENVVSFSASFD 435
Ec_E2p          SIGGLGTTHFAPIVNAPEVAILGVSKSAMEPVWNGKEFVP---------RLMLPISLSFD 602
Lm_E2pL         TTPARAMQQKAAPAAAADEASKTAAKSAAPAAVSGDIYN-----------VVLKPGPVY 366
                .    .    :.   ..:      .          .               ::  . .

Hs_E2p          HRVVDGAVGAQWLAEFRKYLEKPITMLL                                647
Hs_E3BP         SRVVDDELATRFLKSFKANLENPIRLA-                                501
Lm_E2p          HRIVDGALGAKWFQHFHDAMENPLSLLL                                463
Ec_E2p          HRVIDGADGARFITIINNTLSDIRRLVM                                630
Lm_E2pL         KSVSDTALLKKLMHTMHVPKPKLKKAAE                                394
                    : *     : :  :.    .
```

Figure 7.8 Alignment of E2p and E3BP protein sequences
ClustalW alignment of *L. major* E2p (*Lm*E2p) and *L. major* E2p-like (*Lm*E2pL) (see Table 3.1) with: *E. coli* E2p (*Ec*E2p) (accession number AAC73226); *H. sapiens* E2p (*Hs*E2p) (accession number NP_001922); *H. sapiens* E3BP (*Hs*E3BP) (accession number NP_001128496). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.).Red residues indicate conserved motifs within E2p and E3BP enzymes. Green residues are those that are conserved in E1p proteins but not in E3BP or *Lm*E2pL.

```
Hs_E2b            --TPAVRRLAMENNIKLSEVVGSGKD-GRILKEDILNYLEKQTGA 42
Rn_E2b            --TPAVRRLAMENNIKLSEVVGSGKD-GRILKEDILNFLEKQTGA 42
Bs_E2b            --MPSVRKYAREKGVDIRLVQGTGKN-GRVLKEDIDAFLAGGAKP 42
Ec_E2p            --TPLIRRLAREFGVNLAKVKGTGRK-GRILREDVQAYVKEAIKR 42
Hs_E2p            FVSPLAKKLAVEKGIDLTQVKGTGPD-GRITKKDIDSFVP--SKV 42
Mm_E2p            FVSPLAKKLAAEKGIDLTQVKGTGPE-GRIIKKDIDSFVP--SKA 42
Sc_E2p            --SPLAKTIALEKGISLKDVHGTGPR-GRITKADIESYLEKSSKQ 42
Lm_E2p            KASPYARKMAAEKNVSLRGIKGTGGGVGRITSKDVAAAVASGT-- 43
Ec_E2k            --SPAIRRLLAEHNLDASAIKGTGVG-GRLTREDVEKHLAKAPAK 42
Hs_E3BP           RLSPAARNILEKHSLDASQGTATGPR-GIFTKEDALKLVQLKQ-- 42
Lm_E2b            -EECELTRLMEVRG-SLKDVVKERSK-GKAKLSFMPFFLKAASIA 42
Lm_E2k            --LGLMSPFVKASAIALKDVPIVNASFGKDTID-YHEFVDIAIAV 42
Lm_E2pL           --PSMETGTVVEWKKKIGELVKESDVFC-TIQTDKAVVDYTNTFE 42
```

Figure 7.9 Alignment of the E1/E3 binding sequences found within E2p and E3BP proteins, respectively
ClustalW alignment of the E1/E3 binding domains of *Lm*E2p, *Lm*E2pL, *Hs*E2p, *Hs*E3BP, *Ec*E2p (see Figure 7.8), *Lm*E2b, *Hs*E2b, *Rn*E2b (see Figure 7.10), *Lm*E2k, *Ec*E2k (see Figure 7.7) and: *B. stearothermophilus* E2p (*Bs*E2p) (accession number CAA37630); *M. musculus* E2p (*Mm*E2p) (accession number NP_663589); *S. cerevisiae* E2p (*Sc*E2p) (accession number NP_014328). Red residues are those in *H. sapiens* E3 that interact with E3BP. Replacement of Pro with a less constraining residue (blue residue), along with a positively charged residue instead of Ile (blue residue), is thought to permit E1 binding as well as E3 binding (Ciszak *et al.*, 2006). Green residues represent those from *L. major* E2 proteins that do not conform to the consensus, but which could perform similar functions. For example, the Val in *Lm*E2p does not align well with Pro and Lys residues due to insertion of two Gly residues, but could be homologous to the Val observed in *Ec*E2k. *Lm*E2b and *Lm*E2k potentially fulfil E1 and E3 binding since they possess both Ser and Lys residues, instead of Lys and Arg.

```
Hs_E2b          MAAVRMLRTWSRNAGKLICVRYFQTCGNVHVLKPNYVCFFGYPSFKYSHPHHFLKTTAAL 60
Rn_E2b          MAAARVLGTWSRNAVRLTCVRYFQTYNNVHILKPQHVCSVGYPLLKSSQPRHSLRTAAVL 60
Lm_E2b          MRSARCMVRR--------CIGAAAGAATVAHAS-------GWPYTASR--RHLFATTCAP 43
                * :.* :           *:       .*   .       *:*       :* : *:..

Hs_E2b          RGQVVQFKLSDIGEGIREVTVKEWYVKEGDTVSQFDSICEVQSDKASVTITSRYDGVIKK 120
Rn_E2b          QGQVVQFKLSDIGEGIREVTIKEWYVKEGDTVSQFDSICEVQSDKASVTITSRYDGVIKR 120
Lm_E2b          LGRCIPYRLADIGEGITEVQVLGVCVKAGDTINEFDPICEVQSDKATVDITSRYTGVVKA 103
                 *: : ::*:****** ** :    ** ***:.:**.*********:* ***** **:*

Hs_E2b          LYYNLDDIAYVGKPLVDIETEALKDSEEDVVET-----PAVSHD---EHTHQEIK----- 167
Rn_E2b          LYYNLDDIAYVGKPLIDIETEALKDSEEDVVET-----PAVAHD---EHTHQEIK----- 167
Lm_E2b          VYLQPGATAKVGSVMLDIVPEGADDAPEAASPSRSAPPPSSAPDSAPQATYSASKPSSDA 163
                :* : .  * **. ::** .*. .*: * .  :     *: : *   : *:.  *

Hs_E2b          -GRKTLATPAVRRLAMENNIKLSEVVGSGKDGRILKEDILNYLEKQTGAILPPSPKVEIM 226
Rn_E2b          -GQKTLATPAVRRLAMENNIKLSEVVGSGKDGRILKEDILNFLEKQTGAILPPSPKSEIT 226
Lm_E2b          SAGKVLATPATRYLAREHKLDLAHVPATGKGGRVTKEDVLQFMDAGMSAAAAPSPPSTAS 223
                 . *.*****.* ** *:::.*:.* .:**.**: ***:*::::   .*  .***

Hs_E2b          PPPPKPKDMTVPILVSKPPVFTGKDKTEPIKGFQKAMVKTMSAALKIPHFGYCDEIDLTE 286
Rn_E2b          PPPPQPRDRPFPTPVSKPPVFLGKDRTEPVTGFQKAMVKTMSAALKIPHFGYCDEVDLTE 286
Lm_E2b          SAATAPPG----TVVSGLQTEAG-DTVMPITGVRRGMVKTMSQAASIPTFTFSEECELTR 278
                .... * .      **   .  * * . *:.*.::.***** * .** * :.:* :**.

Hs_E2b          LVKLREELKPIAFARG---IKLSFMPFFLKAASLGLLQFPILNASVDENCQNITYKASHN 343
Rn_E2b          LVKLREELKPVALARG---IKLSFMPFFLKAASLGLLQFPILNASVDENCQSITYKASHN 343
Lm_E2b          LMEVRGSLKDVVKERSKGKAKLSFMPFFLKAASIALQHHPDINAHCPVDCSALVRKAAHN 338
                *:::* .** :.  *.    *************:.* :.* :**    :*. :. **:**

Hs_E2b          IGIAMDTEQGLIVPNVKNVQICSIFDIATELNRLQKLGSVSQLSTTDLTGGTFTLSNIGS 403
Rn_E2b          IGIAMDTERGLIVPNVKNVQVRSVFEIAMELNRLQKLGSLGQLSTTDLTGGTFTLSNIGS 403
Lm_E2b          IGFAMDTPNGLIVPVVKHVERKSILDIANDMQVLIERGKSNKLTTQDMTGGTFTLSNIGV 398
                **:**** .***** **:*:  *:::** ::: * : *. .:*:* *:***********

Hs_E2b          IGGTFAKPVIMPPEVAIGALGSIKAIPRFNQKGEVYKAQIMNVSWSADHRVIDGATMSRF 463
Rn_E2b          IGGTYAKPVILPPEVAIGALGAIKALPRFDQKGDVYKAQIMNVSWSADHRVIDGATMSRF 463
Lm_E2b          IGATVTTPVLLPPQVAIGAIGRLQKLPRFDANGSLYAANLICVSFTADHRVIDGASMVRF 458
                **.* :.**::**:*****:* :: :***: :*.:* *::: **::*********:* **

Hs_E2b          SNLWKSYLENPAFMLLDLK                                           482
Rn_E2b          SNLWKSYLENPAFMLLDLK                                           482
Lm_E2b          ANTYKQLLEHPENMLVDLR                                           477
                :* :*. **:*  **:**:
```

Figure 7.10 Alignment of E2b protein sequences
ClustalW alignment of *L. major* E2b (*Lm*E2b) (see Table 3.1) with homologues in *H. sapiens* (*Hs*E2b) (accession number NP_001909) and *R. norvegicus* (*Rn*E2b) (accession number NP_445764). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues indicate conserved amino acids within different motifs that fulfil various roles within E2b enzymes.

```
Hs_H-protein      MALRVVRSVRALLCTLRAVPLPAAPCPPRPWQLGVGAVRTLRTGPALLSVRKFTEKHEWV 60
Rn_H-protein      MSLRVVRSVRAVALQPALSPWP--PCPPRPWRRALAAVRSLRTGSALLSVRKFTEKHEWV 58
Lm_H-protein      -----MR--RAFASVP-----------------VAAAAYLRC----YATKHFTDSHEWV 31
                       :*  **.                     :.*.  **      :.::**:.****

Hs_H-protein      TTENGIGTVGISNFAQEALGDVVYCSLPEVGTKLNKQDEFGALESVKAASELYSPLSGEV 120
Rn_H-protein      TAKDGIGTVGISNFAQEALGDVVYCSLPEVGTKLKKQEEFGGLESVKAASELYSPLSGEV 118
Lm_H-protein      MQCEDEITIGISSYAQENLGDVVYVSLPQVGDTVKEKDVIGEVESVKATSNVYSPVDGTV 91
                     :.  *:***.:*** ****** ***:** .::::: :* :*****:*::***:.* *

Hs_H-protein      TEINEALAENPGLVNKSCYEDGWLIKMTLSNPSELDELMSEEAYEKYIKSIEE 173
Rn_H-protein      TEVNEALPENPGLVNKSMYEDGWLIKMTLSDPSELDELMSEEAYEKYVKSN-- 169
Lm_H-protein      SAVNENLKDEPGLVNQSPEEKGWLIKVKCSEIP--KGLMDEAAYKKFLE---- 138
                  : :** * ::*****:*  *.*****:. *: .  . **.* **:*:::
```

Figure 7.11 Alignment of H-protein sequences
ClustalW alignment of *L. major* H-protein (*Lm*H-protein) (see Table 3.1) with homologues in *H. sapiens* (*Hs*H-protein) (accession number NP_004474) and *R. norvegicus* (*Rn*H-protein) (accession number CAB56621). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.).Red residues indicate conserved amino acids within different motifs that fulfil various roles within H-protein. Green residues are those that are conserved in H-protein motifs in other species but not in *Lm*H-protein.

```
Lm_T-protein-A      ----------------------------MSASLKKTALHLFHLAQQAKMDAFAGYHMPIS 32
Lm_T-protein-B      ----------------------------MSASLKKTALHLFHLAQQAKMDAFAGYHMPIS 32
Hs_T-protein        MQRAVSVVARLGFRLQAFPPALCRPLSCAQEVLRRTPLYDFHLAHGGKMVAFAGWSLPVQ 60
Ec_T-protein        ------------------------------MAQQTPLYEQHTLCGARMVDFHGWMMPLH 29
                                                   ::*.*:  *    .:*  * *: :*:

Lm_T-protein-A      YGRLGVLKEHLYTREVAGIFDVSHVGQYEVRGADRERFLEHVTPVDLQRIR-AGHGALTM 91
Lm_T-protein-B      YGRLGVLKEHLYTREVAGIFDVSHVGQYEVRGADRERFLEHVTPVDLQRIR-AGHGALTM 91
Hs_T-protein        Y-RDSHTDSHLHTRQHCSLFDVSHMLQTKILGSDRVKLMESLVVGDIAELR-PNQGTLSL 118
Ec_T-protein        Y--GSQIDEHHAVRTDAGMFDVSHMTIVDLRGSRTREFLRYLLANDVAKLTKSGKALYSG 87
                    *   .  ..*  .*  ..:*****:   .: *:   .::. :   *: .:  ..:.  :

Lm_T-protein-A      LTNAQGGIKDDCIVTKMAD-HLFLVLNAGCKEKDVAHMESVLRESAMKGADVQLVPLDRS 150
Lm_T-protein-B      LTNAQGGIKDDCIVTKMAD-HLFLVLNAGCKEKDVAHMESVLRESAMKGADVQLVPLDRS 150
Hs_T-protein        FTNEAGGILDDLIVTNTSEGHLYVVSNAGCWEKDLALMQDKVRELQNQGRDVGLEVLDNA 178
Ec_T-protein        MLNASGGVIDDLIVYYFTEDFFRLVVNSATREKDLSWITQHAEPFGIEIT----VRDDLS 143
                    : *  **: ** **   :: .: :* *:.  ***:: : .  .    :         * :

Lm_T-protein-A      LIALQGPQAAAILSE-FMDDVPGMGFMQCRQRVNIKGMEVQVTRCGYTGEDGFELSVSNT 209
Lm_T-protein-B      LIALQGPQAAAILSE-FMDDVPGMGFMQCRQRVNIKGMEVQVTRCGYTGEDGFELSVSNT 209
Hs_T-protein        LLALQGPTAAQVLQAGVADDLRKLPFMTSAVMEVFGVSGCRVTRCGYTGEDGVEISVPVA 238
Ec_T-protein        MIAVQGPNAQAKAAT-LFNDAQRQAVEGMKPFFGVQAGDLFIATTGYTGEAGYEIALPNE 202
                    ::*:*** *       . :*     .        .      ::  ***** * *:::.

Lm_T-protein-A      DIVALVELLMSR-KAEMIGLGARDSLRLEAGLNLYGHELTEDINPVAARFMWVISKRRMA 268
Lm_T-protein-B      DIVALVELLMSR-KAEMIGLGARDSLRLEAGLNLYGHELTEDINPVAARFMWVISKRRMA 268
Hs_T-protein        GAVHLATAILKNPEVKLAGLAARDSLRLEAGLCLYGNDIDEHTTPVEGSLSWTLGKRRRA 298
Ec_T-protein        KAADFWRALVEA-GVKPCGLGARDTLRLEAGMNLYGQEMDETISPLAANMGWTIAWEP-A 260
                      . :   ::.   .:  **.***:******: ***::: *  .*: . : *.:. .  *

Lm_T-protein-A      EGGFIGYEPIKYLRDNASKGAVPRLRVGLVSTGPVAREKT---VIEVGGKPVGEVTSGCP 325
Lm_T-protein-B      EGGFIGYEPIKYLRDNASKGAVPRLRVGLVSTGPVAREKT---VIEVGGKPVGEVTSGCP 325
Hs_T-protein        AMDFPGAKVIVPQ----LKGRVQRRRVGLMCEGAPMRAHSP--ILNMEGTKIGTVTSGCP 352
Ec_T-protein        DRDFIGREALEVQR---EHGTEKLVGLVMTEKGVLRNELPVRFTDAQGNQHEGIITSGTF 317
                      .* * : :        :*      : :   *   .  .        .   * :***

Lm_T-protein-A      SPCLKKNIAIGYLDRELAKDGVKVDLVVRGRRVAAVVVTPPFVPARYYRKPPVKGTEERR 385
Lm_T-protein-B      SPCLKKNIAIGYLDRELAKDGVKVDLVVRGRRVAAVVVTPPFVPARYYRKPK-------- 377
Hs_T-protein        SPSLKKNVAMGYVPCEYSRPGTMLLVEVRRKQQMAVVSKMPFVPTNYYTLK--------- 403
Ec_T-protein        SPTLGYSIALARVPEGIG---ETAIVQIRNREMPVKVTKPVFVRNGKAVA---------- 364
                    ** *  .:*:. :    .       : :* :.  . * .  **

Lm_T-protein-A      TPTIKNIAS                                                   394
Lm_T-protein-B      ---------
Hs_T-protein        ---------
Ec_T-protein        ---------
```

Figure 7.12 Alignment of T-protein sequences
ClustalW alignment of *L. major* T-protein-A and -B (*Lm*T-protein-A and -B) (see Table 3.1) with homologues in *H. sapiens* (*Hs*T-protein) (accession number NP_000472) and *E. coli* (*Ec*T-protein) (accession number AAC75943). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.).Red residues indicate conserved motifs within T-proteins. Green residues are those that are conserved in bacterial T-proteins but not in *Hs*T-protein or *Lm*T-protein-A/B.

```
Lm_LipDH        VTTLSGPRSNTHTFPHFTHT----QAHTSMFRRNIAHLAS-YDVTVIGGGPGGYVAAIKA 55
Hs_LipDH        MQSWSRVYCSLAKRGHFNRISHGLQGLSAVPLRTYADQPIDADVTVIGSGPGGYVAAIKA 60
Ec_LipDH        ----------------------------------MSTEIKTQVVVLGAGPAGYSAAFRC 25
                                                   :*.*:*.**.** **::.

Lm_LipDH        AQLGLKTACIEKRGALGGTCLNVGCIPSKALLHATHLYHDAH-ANFAQYGLRGGENVTMD 114
Hs_LipDH        AQLGFKTVCIEKNETLGGTCLNVGCIPSKALLNNSHYYHMAHGKDFASRGIEMSE-VRLN 119
Ec_LipDH        ADLGLETVIVERYNTLGGVCLNVGCIPSKALLHVAKVIEEAK--ALAEHGIVFGE-PKTD 82
                 *:**::*. :*:  :***.*************: ::  . *:   :*. *:  .*    :

Lm_LipDH        VSAMQAQKGKGVKALTGGVEYLFKKNKVTYYKGEGSFVNPNTIKVKGLDGKEETLESKKT 174
Hs_LipDH        LDKMMEQKSTAVKALTGGIAHLFKQNKVVHVNGYGKITGKNQVTATKADGGTQVIDTKNI 179
Ec_LipDH        IDKIRTWKEKVINQLTGGLAGMAKGRKVKVVNGLGKFTGANTLEVEGENGKT-VINFDNA 141
                 :. :   * . :: ****:  : * .**   :* *.:.. * : .   :*   .:: .:

Lm_LipDH        IVATGSEPTELPFLPFDEKVVMSSTGALDLDHVPKKMIVVGGGVIGLELGSVWARLGAEV 234
Hs_LipDH        LIATGSEVTPFPGITIDEDTIVSSTGALSLKKVPEKMVVIGAGVIGVELGSVWQRLGADV 239
Ec_LipDH        IIAAGSRPIQLPFIPHEDPRIWDSTDALELKEVPERLLVMGGGIIGLEMGTVYHALGSQI 201
                 ::*:**.   :* :. ::  : .**.**.*..**::::*:*.*:**:*:*:*:  **:::

Lm_LipDH        TVVEFASRCAAT-TDADVSKALTDALVKHEKMKIMTNTKVVSGTNNGSS-VTIEVEDKDG 292
Hs_LipDH        TAVEFLGHVGGVGIDMEISKNFQRILQKQG-FKFKLNTKVTGATKKSDGKIDVSIEAASG 298
Ec_LipDH        DVVEMFDQVIPA-ADKDIVKVFTKRISKKF--NLMLETKVTAVEAKEDG-IYVTMEGKKA 257
                  .**: .:   .  * :: * :   : *:   ::  :***..   : .. : : :*  ..

Lm_LipDH        -KHQTLEADALLCSVGRRPHTTGLNAEAINLQME-RGFICINDHFETNVPNVYAIGDVVN 350
Hs_LipDH        GKAEVITCDVLLVCIGRRPFTKNLGLEELGIELDPRGRIPVNTRFQTKIPNIYAIGDVVA 358
Ec_LipDH        -PAEPQRYDAVLVAIGRVPNGKNLDAGKAGVEVDDRGFIRVDKQLRTNVPHIFAIGDIVG 316
                    :    *.:* .:** *  ..*.    .:::: ** * :: ::.*::*:::****:*

Lm_LipDH        KGPMLAHKAEEEGVACAEILAGKPGHVNYSVIPGVIYTNPEVAQVGETEEQVKKRGIDYK 410
Hs_LipDH        -GPMLAHKAEDEGIICVEGMAGGAVHIDYNCVPSVIYTHPEVAWVGKSEEQLKEEGIEYK 417
Ec_LipDH        -QPMLAHKGVHEGHVAAEVIAGKKHYFDPKVIPSIAYTEPEVAWVGLTEKEAKEKGISYE 375
                   ******. .**  ..* :**   :.: . :*.: **.**** ** :*:: *:.**.*:

Lm_LipDH        VGKFPFSANSRAKAVGTEDGFVKVVTDKKTDRILGVQIVCTAAGEMIAEPTLAMEYGASS 470
Hs_LipDH        VGKFPFAANSRAKTNADTDGMVKILGQKSTDRVLGAHILGPGAGEMVNEAALALEYGASC 477
Ec_LipDH        TATFPWAASGRAIASDCADGMTKLIFDKESHRVIGGAIVGTNGGELLGEIGLAIEMGCDA 435
                 ...**::*..** :    **:.*:: :*.:.*::*  *: . .**:: *  **:* *...

Lm_LipDH        EDLGRTCHAHPTMSEAVKEACMAC-FAQTINF---------                     501
Hs_LipDH        EDIARVCHAHPTLSEAFREANLAASFGKSINF---------                     509
Ec_LipDH        EDIALTIHAHPTLHESVGLA--AEVFEGSITDLPNPKAKKK                     474
                 **:. . *****: *:.  *  *  *  :*.
```

Figure 7.13 Alignment of LipDH protein sequences
ClustalW alignment of *L. major* LipDH (*Lm*LipDH) (see Table 3.1) with homologues in *H. sapiens* (*Hs*LipDH) (accession number NP_000099) and *E. coli* (*Ec*LipDH) (accession number AAC73227). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues indicate conserved motifs within LipDH proteins.

```
Hs_PDK-1        -ELFKNAMRATMEHHANR--------GVYPPIQVHVTLGN--EDLTVKMSDRGGGVPLRK 49
Hs_PDK-2        -ELFKNAMRATVESHESS--------LILPPIKVMVALGE--EDLSIKMSDRGGGVPLRK 49
Hs_PDK-3        -ELFKNSMRATVELYEDRK-------EGYPAVKTLVTLGK--EDLSIKISDLGGGVPLRK 50
Hs_PDK-4        -ELFKNAMRATVEHQENQ--------PSLTPIEVIVVLGK--EDLTIKISDRGGGVPLRI 49
Hs_BCKDK        PELLKNAMRATMESHLDTP-------YNVPDVVITIANND--VDLIIRISDRGGGIAHKD 51
LmjF20.0280     LELMKNAFRATVDSHMKRNDVGMVTCADMPPVRVLINLQEGTEHACICISDEGMGMTDEA 60
                 **:**::***::   .            . :   :   .   .  : :** * *:. .

Hs_PDK-1        IDRLFNYMYSTAPR----PRV----------ETSRAVPLAGFGYGLPISRLYAQYFQGDL 95
Hs_PDK-2        IERLFSYMYSTAPT----PQP----------GTG-GTPLAGFGYGLPISRLYAKYFQGDL 94
Hs_PDK-3        IDRLFNYMYSTAPR----PSL----------EPTRAAPLAGFGYGLPISRLYARYFQGDL 96
Hs_PDK-4        IDRLFSYTYSTAPT----PVM----------DNSRNAPLAGFGYGLPISRLYAKYFQGDL 95
Hs_BCKDK        LDRVMDYHFTTAEASTQDPRISPLFGHLDMHSGAQSGPMHGFGFGLPTSRAYAEYLGGSL 111
LmjF20.0280     LTMAMAYSYTSVSK----PALQLGESG-ERCASTAPSPLAGYGYGLPMSRVYAQSLGGDL 115
                :   : * :::.      *                  *: *:*:*** ** **. : *.*

Hs_PDK-1        KLYSLEGYGTDAVIYIKALSTDSIERLPVYNKAAWKHYNTNHEADDWCVPSREPKDMTTF 155
Hs_PDK-2        QLFSMEGFGTDAVIYLKALSTDSVERLPVYNKSAWRHYQTIQEAGDWCVPSTEPKNTSTY 154
Hs_PDK-3        KLYSMEGVGTDAVIYLKALSSESFERLPVFNKSAWRHYKTTPEADDWSNPSSEPRDASKY 156
Hs_PDK-4        NLYSLSGYGTDAIIYLKALSSESIEKLPVFNKSAFKHYQMSSEADDWCIPSREPKNLAKE 155
Hs_BCKDK        QLQSLQGIGTDVYLRLRHIDG---------REESFRI---------------------- 139
LmjF20.0280     LLQTMEGYGTRAYYYIKIADAQP------LCDEETK---------------------- 145
                 * ::.* ** .   ::  .           .   :

Hs_PDK-1        RSA-                                                       158
Hs_PDK-2        RVS-                                                       157
Hs_PDK-3        KAKQ                                                       160
Hs_PDK-4        VAM-                                                       158
Hs_BCKDK        ----
LmjF20.0280     ----
```

Figure 7.14 Alignment of the ATPase/kinase domains of known PDH kinase proteins with a potential *L. major* α-KADH kinase, *LmjF*20.0280
ClustalW alignment of *LmjF*20.0280 (see Table 3.1) with: *H. sapiens* PDK-1 (*Hs*PDK-1) (accession number NP_002601); *H. sapiens* PDK-2 (*Hs*PDK-2) (accession number NP_002602); *H. sapiens* PDK-3 (*Hs*PDK-3) (accession number NP_005382); *H. sapiens* PDK-4 (*Hs*PDK-4) (accession number NP_002603); *H. sapiens* BCKDK (*Hs*BCKDK) (accession number NP_005872). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues indicate conserved motifs within PDK and BCKDK proteins. Orange residues are those required for binding of PDK proteins to the E2p lipoyl-domain 2 (L2) of the PDH complex.

```
Hs_PDK-1        -ELFKNAMRATMEHHANR-GVYPPIQVHVTLG--NEDLTVKMSDRGGGVPLRKIDRLFN- 55
Hs_PDK-2        -ELFKNAMRATVESHESS-LILPPIKVMVALG--EEDLSIKMSDRGGGVPLRKIERLFS- 55
Hs_PDK-3        -ELFKNSMRATVELYEDRKEGYPAVKTLVTLG--KEDLSIKISDLGGGVPLRKIDRLFN- 56
Hs_PDK-4        -ELFKNAMRATVEHQENQ-PSLTPIEVIVVLG--KEDLTIKISDRGGGVPLRIIDRLFS- 55
Hs_BCKDK        PELLKNAMRATMESHLDTPYNVPDVVITIANN--DVDLIIRISDRGGGIAHKDLDRVMD- 57
LmjF24.0010     CAMLEDAVSANVDRQERTGKECTKIEVTLAQWPTNKRFVLRISDTAGGMTLRQASMQLSC 60
                  :::::: *.::         . :   :.    .  : :::** .**:. :  .  :.

Hs_PDK-1        -YMYSTAPR----PRV----------ETSRAVPLAGFGYGLPISRLYAQYFQGDLKLYSL 100
Hs_PDK-2        -YMYSTAPT----PQP----------GTG-GTPLAGFGYGLPISRLYAKYFQGDLQLFSM 99
Hs_PDK-3        -YMYSTAPR----PSL----------EPTRAAPLAGFGYGLPISRLYARYFQGDLKLYSM 101
Hs_PDK-4        -YTYSTAPT----PVM----------DNSRNAPLAGFGYGLPISRLYAKYFQGDLNLYSL 100
Hs_BCKDK        -YHFTTAEASTQDPRISPLFGHLDMHSGAQSGPMHGFGFGLPTSRAYAEYLGGSLQLQSL 116
LmjF24.0010     WSLYRNIQG----------------HNQDTISTWTSSPIRLPYAYNAARVIGGNITLASI 104
                   : .                          .  .    ** :   *. : *.: * *:

Hs_PDK-1        EGYGTDAVIYIKALSTDSIERLPVYNKAAWKHYNTNHEADDWCVPSREPKDMTTFRSA- 158
Hs_PDK-2        EGFGTDAVIYLKALSTDSVERLPVYNKSAWRHYQTIQEAGDWCVPSTEPKNTSTYRVS- 157
Hs_PDK-3        EGVGTDAVIYLKALSSESFERLPVFNKSAWRHYKTTPEADDWSNPSSEPRDASKYKAKQ 160
Hs_PDK-4        SGYGTDAIIYLKALSSESIEKLPVFNKSAFKHYQMSSEADDWCIPSREPKNLAKEVAM- 158
Hs_BCKDK        QGIGTDVYLRLRHIDG---------REESFRI------------------------- 139
LmjF24.0010     EGYGTDRQLYLPSTGLAGVSL------------------------------------ 125
                .* ***  : :   .
```

Figure 7.15 Alignment of the ATPase/kinase domains of known PDH kinase proteins with a potential *L. major* α-KADH kinase, *LmjF*24.0010
ClustalW alignment of *LmjF*24.0010 (see Table 3.1) with: *H. sapiens* PDK-1 (*Hs*PDK-1) (accession number NP_002601); *H. sapiens* PDK-2 (*Hs*PDK-2) (accession number NP_002602); *H. sapiens* PDK-3 (*Hs*PDK-3) (accession number NP_005382); *H. sapiens* PDK-4 (*Hs*PDK-4) (accession number NP_002603); *H. sapiens* BCKDK (*Hs*BCKDK) (accession number NP_005872). The alignment indicates identical residues (*), conserved residues (:) and homologous residues (.). Red residues indicate conserved motifs within PDK and BCKDK proteins. Green residues are those that are strictly conserved in PDK and BCKDK proteins but not in *LmjF*24.0010. Orange residues are those required for binding of PDK proteins to the E2p lipoyl-domain 2 (L2) of the PDH complex.

References

Aevarsson, A., Seger, K., Turley, S., Sokatch, J. R. & Hol, W. G. (1999). Crystal structure of 2-oxoisovalerate and dehydrogenase and the architecture of 2-oxo acid dehydrogenase multienzyme complexes. *Nat Struct Biol* **6**, 785-792.

Aevarsson, A., Chuang, J. L., Wynn, R. M., Turley, S., Chuang, D. T. & Hol, W. G. (2000). Crystal structure of human branched-chain alpha-ketoacid dehydrogenase and the molecular basis of multienzyme complex deficiency in maple syrup urine disease. *Structure* **8**, 277-291.

Agabian, N. (1990). Trans splicing of nuclear pre-mRNAs. *Cell* **61**, 1157-1160.

Allary, M., Lu, J. Z., Zhu, L. & Prigge, S. T. (2007). Scavenging of the cofactor lipoate is essential for the survival of the malaria parasite *Plasmodium falciparum*. *Molecular Microbiology* **63**, 1331-1344.

Allen, A. G., Perham, R. N., Allison, N., Miles, J. S. & Guest, J. R. (1989). Reductive acetylation of tandemly repeated lipoyl domains in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli* is random order. *J Mol Biol* **208**, 623-633.

Anderson, M. D., Che, P., Song, J., Nikolau, B. J. & Wurtele, E. S. (1998). 3-Methylcrotonyl-coenzyme A carboxylase is a component of the mitochondrial leucine catabolic pathway in plants. *Plant Physiol* **118**, 1127-1138.

Antoine, J. C., Prina, E., Courret, N. & Lang, T. (2004). *Leishmania spp.*: on the interactions they establish with antigen-presenting cells of their mammalian hosts. *Adv Parasitol* **58**, 1-68.

Appling, D. R. (1991). Compartmentation of folate-mediated one-carbon metabolism in eukaryotes. *FASEB J* **5**, 2645-2651.

Atmaca, G. (2004). Antioxidant effects of sulfur-containing amino acids. *Yonsei Med J* **45**, 776-788.

Autio, K. J., Guler, J. L., Kastaniotis, A. J., Englund, P. T. & Hiltunen, J. K. (2008a). The 3-hydroxyacyl-ACP dehydratase of mitochondrial fatty acid synthesis in *Trypanosoma brucei*. *FEBS Lett* **582**, 729-733.

Autio, K. J., Kastaniotis, A. J., Pospiech, H., Miinalainen, I. J., Schonauer, M. S., Dieckmann, C. L. & Hiltunen, J. K. (2008b). An ancient genetic link between vertebrate mitochondrial fatty acid synthesis and RNA processing. *Faseb J* **22**, 569-578.

Bagautdinov, B., Kuroishi, C., Sugahara, M. & Kunishima, N. (2005). Crystal structures of biotin protein ligase from *Pyrococcus horikoshii* OT3 and its complexes: structural basis of biotin activation. *J Mol Biol* **353**, 322-333.

Bakker, B. M., Mensonides, F. I. C., Teusink, B., van Hoek, P., Michels, P. A. M. & Westerhoff, H. V. (2000). From the Cover: Compartmentation protects trypanosomes from the dangerous design of glycolysis. *PNAS* **97**, 2087-2092.

Bao, H., Kasten, S. A., Yan, X. & Roche, T. E. (2004). Pyruvate dehydrogenase kinase isoform 2 activity limited and further inhibited by slowing down the rate of dissociation of ADP. *Biochemistry* **43**, 13432-13441.

Barford, D., Das, A. K. & Egloff, M. P. (1998). The structure and mechanism of protein phosphatases: insights into catalysis and regulation. *Annu Rev Biophys Biomol Struct* **27**, 133-164.

Barras, F. & Marinus, M. G. (1989). The great GATC: DNA methylation in *E. coli*. *Trends Genet* **5**, 139-143.

Barros, V. C., Oliveira, J. S., Melo, M. N. & Gontijo, N. F. (2006). *Leishmania amazonensis*: chemotaxic and osmotaxic responses in promastigotes and their probable role in development in the phlebotomine gut. *Exp Parasitol* **112**, 152-157.

Bastin, P., Pullen, T. J., Moreira-Leite, F. F. & Gull, K. (2000). Inside and outside of the trypanosome flagellum:a multifunctional organelle. *Microbes Infect* **2**, 1865-1874.

Beckett, D. & Matthews, B. W. (1997). *Escherichia coli* repressor of biotin biosynthesis. *Methods Enzymol* **279**, 362-376.

Beitner, H. (2003). Randomized, placebo-controlled, double blind study on the clinical efficacy of a cream containing 5% alpha-lipoic acid related to photoageing of facial skin. *Br J Dermatol* **149**, 841-849.

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

Besteiro, S., Williams, R. A., Morrison, L. S., Coombs, G. H. & Mottram, J. C. (2006). Endosome sorting and autophagy are essential for differentiation and virulence of *Leishmania major*. *J Biol Chem* **281**, 11384-11396.

Besteiro, S., Williams, R. A., Coombs, G. H. & Mottram, J. C. (2007). Protein turnover and differentiation in Leishmania. *Int J Parasitol* **37**, 1063-1075.

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

Bienen, E. J., Hammadi, E. & Hill, G. C. (1981). *Trypanosoma brucei*: biochemical and morphological changes during in vitro transformation of bloodstream- to procyclic-trypomastigotes. *Exp Parasitol* **51**, 408-417.

Bilwes, A. M., Alex, L. A., Crane, B. R. & Simon, M. I. (1999). Structure of CheA, a signal-transducing histidine kinase. *Cell* **96**, 131-141.

Bogdan, C. & Rollinghoff, M. (1999). How do protozoan parasites survive inside macrophages? *Parasitol Today* **15**, 22-28.

Booker, S. J., Cicchillo, R. M. & Grove, T. L. (2007). Self-sacrifice in radical S-adenosylmethionine proteins. *Curr Opin Chem Biol* **11**, 543-552.

Bork, P., Brown, N. P., Hegyi, H. & Schultz, J. (1996). The protein phosphatase 2C (PP2C) superfamily: detection of bacterial homologues. *Protein Sci* **5**, 1421-1425.

Bourguignon, J., Merand, V., Rawsthorne, S., Forest, E. & Douce, R. (1996). Glycine decarboxylase and pyruvate dehydrogenase complexes share the same dihydrolipoamide dehydrogenase in pea leaf mitochondria: evidence from mass spectrometry and primary-structure analysis. *Biochem J* **313 (Pt 1)**, 229-234.

Bouvier, L. A., Miranda, M. R., Canepa, G. E., Alves, M. J. & Pereira, C. A. (2006). An expanded adenylate kinase gene family in the protozoan parasite *Trypanosoma cruzi*. *Biochim Biophys Acta* **1760**, 913-921.

Bowker-Kinley, M. & Popov, K. M. (1999). Evidence that pyruvate dehydrogenase kinase belongs to the ATPase/kinase superfamily. *Biochem J* **344 Pt 1**, 47-53.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248-254.

Bringaud, F., Riviere, L. & Coustou, V. (2006). Energy metabolism of trypanosomatids: adaptation to available carbon sources. *Mol Biochem Parasitol* **149**, 1-9.

Brody, S., Oh, C., Hoja, U. & Schweizer, E. (1997). Mitochondrial acyl carrier protein is involved in lipoic acid synthesis in *Saccharomyces cerevisiae*. *FEBS Lett* **408**, 217-220.

Bunik, V., Kaehne, T., Degtyarev, D., Shcherbakova, T. & Reiser, G. (2008). Novel isoenzyme of 2-oxoglutarate dehydrogenase is identified in brain, but not in heart. *Febs J*.

Bunik, V. I. & Degtyarev, D. (2008). Structure-function relationships in the 2-oxo acid dehydrogenase family: substrate-specific signatures and functional predictions for the 2-oxoglutarate dehydrogenase-like proteins. *Proteins* **71**, 874-890.

Burchmore, R. J., Rodriguez-Contreras, D., McBride, K., Merkel, P., Barrett, M. P., Modi, G., Sacks, D. & Landfear, S. M. (2003). Genetic characterization of glucose transporter function in *Leishmania mexicana*. *Proc Natl Acad Sci U S A* **100**, 3901-3906.

Bursell, E. (1960). Free amino-acids of the tsetse fly (*Glossina*). *Nature* **187**, 778.

Bursell, E. (1963). Aspects of metabolism of amino acids in the tsetse fly, *Glossina* (Diptera). *Journal of Insect Physiology* **9**, 439-452.

Bursell, E. (1967). The conversion of glutamate to alanine in the tsetse fly (*Glossina morsitans*). *Comp Biochem Physiol* **23**, 825-829.

Burton, G. W. & Ingold, K. U. (1981). Autoxidation of Biological Molecules. 1. The Antioxidant Activity of Vitamin E and Related Chain-Breaking Phenolic Antioxidants in Vitro. *J Am Chem Soc* **103**, 6472-6477.

Cazzulo, J. J., Franke de Cazzulo, B. M., Engel, J. C. & Cannata, J. J. (1985). End products and enzyme levels of aerobic glucose fermentation in trypanosomatids. *Mol Biochem Parasitol* **16**, 329-343.

Chappuis, F., Sundar, S., Hailu, A., Ghalib, H., Rijal, S., Peeling, R. W., Alvar, J. and Boelaert, M. (2007). Visceral leishmaniasis: what are the needs for diagnosis, treatment and control? *Nat Rev Microbiol* **5**, 873-882.

Charlab, R., Tesh, R. B., Rowton, E. D. & Ribeiro, J. M. (1995). *Leishmania amazonensis*: sensitivity of different promastigote morphotypes to salivary gland homogenates of the sand fly *Lutzomyia longipalpis*. *Exp Parasitol* **80**, 167-175.

Chen, X. J. (1997). Cloning and characterization of the lipoyl-protein ligase gene LIPB from the yeast *Kluyveromyces lactis*: synergistic respiratory deficiency due to mutations in LIPB and mitochondrial F1-ATPase subunits. *Mol Gen Genet* **255**, 341-349.

Chenik, M., Lakhal, S., Ben Khalef, N., Zribi, L., Louzir, H. & Dellagi, K. (2006). Approaches for the identification of potential excreted/secreted proteins of *Leishmania major* parasites. *Parasitology*, 1-17.

Choi-Rhee, E., Schulman, H. & Cronan, J. E. (2004). Promiscuous protein biotinylation by *Escherichia coli* biotin protein ligase. *Protein Sci* **13**, 3043-3050.

Cicchillo, R. M., Iwig, D. F., Jones, A. D., Nesbitt, N. M., Baleanu-Gogonea, C., Souder, M. G., Tu, L. & Booker, S. J. (2004a). Lipoyl synthase requires two equivalents of S-adenosyl-L-methionine to synthesize one equivalent of lipoic acid. *Biochemistry* **43**, 6378-6386.

Cicchillo, R. M., Lee, K. H., Baleanu-Gogonea, C., Nesbitt, N. M., Krebs, C. & Booker, S. J. (2004b). *Escherichia coli* lipoyl synthase binds two distinct [4Fe-4S] clusters per polypeptide. *Biochemistry* **43**, 11770-11781.

Cicchillo, R. M. & Booker, S. J. (2005). Mechanistic investigations of lipoic acid biosynthesis in *Escherichia coli*: both sulfur atoms in lipoic acid are contributed by the same lipoyl synthase polypeptide. *J Am Chem Soc* **127**, 2860-2861.

Cinti, D. L., Cook, L., Nagi, M. N. & Suneja, S. K. (1992). The fatty acid chain elongation system of mammalian endoplasmic reticulum. *Prog Lipid Res* **31**, 1-51.

Ciszak, E. M., Makal, A., Hong, Y. S., Vettaikkorumakankauv, A. K., Korotchkina, L. G. & Patel, M. S. (2006). How dihydrolipoamide dehydrogenase-binding protein binds dihydrolipoamide dehydrogenase in the human pyruvate dehydrogenase complex. *J Biol Chem* **281**, 648-655.

Claros, M. G. & Vincens, P. (1996). Computational method to predict mitochondrially imported proteins and their targeting sequences. *Eur J Biochem* **241**, 779-786.

Clayton, C., Hausler, T. & Blattner, J. (1995). Protein trafficking in kinetoplastid protozoa. *Microbiol Rev* **59**, 325-344.

Coombs, G. H., Tetley, L., Moss, V. A. & Vickerman, K. (1986). Three dimensional structure of the Leishmania amastigote as revealed by computer-aided reconstruction from serial sections. *Parasitology* **92 (Pt 1)**, 13-23.

Cossins, E. A. & Chen, L. (1997). Folates and one-carbon metabolism in plants and fungi. *Phytochemistry* **45**, 437.

Coustou, V., Besteiro, S., Riviere, L., Biran, M., Biteau, N., Franconi, J.-M., Boshart, M., Baltz, T. & Bringaud, F. (2005). A mitochondrial NADH-dependent fumarate reductase involved in the production of succinate excreted by procyclic *Trypanosoma brucei*. *J Biol Chem* **280**, 16559-16570.

Coustou, V., Biran, M., Breton, M., Guegan, F., Riviere, L., Plazolles, N., Nolan, D., Barrett, M. P., Franconi, J. M. & Bringaud, F. (2008). Glucose-induced remodeling of intermediary and energy metabolism in procyclic *Trypanosoma brucei*. *J Biol Chem* **283**, 16342-16354.

Crawford, M. J., Thomsen-Zieger, N., Ray, M., Schachtner, J., Roos, D. S. & Seeber, F. (2006). *Toxoplasma gondii* scavenges host-derived lipoic acid despite its *de novo* synthesis in the apicoplast. *Embo J* **25**, 3214-3222.

Croft, S. L. and Coombs, G. H. (2003). Leishmaniasis - current chemotherapy and recent advances in the search for novel drugs. *Trends Parasitol* **19**, 502-508.

Croft, S. L., Yardley, V. and Kendrick, H. (2002). Drug sensitivity of *Leishmania* species: some unresolved problems. *Trans R Soc Trop Med Hyg* **96**, *Supp 1*, S127-129.

Croft, S. L., Barrett, M. P. & Urbina, J. A. (2005). Chemotherapy of *trypanosomiases* and *leishmaniasis*. *Trends Parasitol* **21**, 508-512.

Cronan, J. E. (2005). Targeted and proximity-dependent promiscuous protein biotinylation by a mutant *Escherichia coli* biotin protein ligase. *J Nutr Biochem* **16**, 416-418.

Cronan, J. E., Zhao, X. & Jiang, Y. (2005). Function, attachment and synthesis of lipoic acid in *Escherichia coli*. *Adv Microb Physiol* **50**, 103-146.

Cronan, J. E., Jr. (1989). The *E. coli* bio operon: transcriptional repression by an essential protein modification enzyme. *Cell* **58**, 427-429.

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

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

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

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

Cynober, L. & Harris, R. A. (2006). Symposium on branched-chain amino acids: conference summary. *J Nutr* **136**, 333S-336S.

Damuni, Z., Merryfield, M. L., Humphreys, J. S. & Reed, L. J. (1984). Purification and properties of branched-chain alpha-keto acid dehydrogenase phosphatase from bovine kidney. *Proc Natl Acad Sci U S A* **81**, 4335-4338.

Damuni, Z. & Reed, L. J. (1987). Purification and properties of the catalytic subunit of the branched-chain alpha-keto acid dehydrogenase phosphatase from bovine kidney mitochondria. *J Biol Chem* **262**, 5129-5132.

Davies, C. R., Kaye, P. M., Croft, S. L. and Sundar, S. (2003). Leishmaniasis: new approaches to disease control. *BMJ* **326**, 377-382.

Danson, M. J., Conroy, K., McQuattie, A. & Stevenson, K. J. (1987). Dihydrolipoamide dehydrogenase from *Trypanosoma brucei*. Characterization and cellular location. *Biochem J* **243**, 661-665.

de Kok, A., Hengeveld, A. F., Martin, A. & Westphal, A. H. (1998). The pyruvate dehydrogenase multi-enzyme complex from Gram-negative bacteria. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1385**, 353.

De Marcucci, O. & Lindsay, J. G. (1985). Component X. An immunologically distinct polypeptide associated with mammalian pyruvate dehydrogenase multi-enzyme complex. *Eur J Biochem* **149**, 641-648.

Desjeux, P. (2004). Leishmaniasis: current situation and new perspectives. *Comp Immunol Microbiol Infect Dis* **27**, 305-318.

Devasagayam, T. P., Di Mascio, P., Kaiser, S. & Sies, H. (1991). Singlet oxygen induced single-strand breaks in plasmid pBR322 DNA: the enhancing effect of thiols. *Biochim Biophys Acta* **1088**, 409-412.

Donelson, J. E. (2003). Antigenic variation and the African trypanosome genome. *Acta Trop* **85**, 391-404.

Douce, R., Bourguignon, J., Neuburger, M. & Rebeille, F. (2001). The glycine decarboxylase system: a fascinating complex. *Trends in Plant Science* **6**, 167.

Douglas, P., Kriek, M., Bryant, P. & Roach, P. L. (2006). Lipoyl synthase inserts sulfur atoms into an octanoyl substrate in a stepwise manner. *Angew Chem Int Ed Engl* **45**, 5197-5199.

Dutta, R. & Inouye, M. (2000). GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem Sci* **25**, 24-28.

Else, A. J., Clarke, J. F., Willis, A., Jackman, S. A., Hough, D. W. & Danson, M. J. (1994). Dihydrolipoamide dehydrogenase in the *Trypanosoma* subgenus, *trypanozoon*. *Mol Biochem Parasitol* **64**, 233-239.

Emanuelsson, O., Nielsen, H., Brunak, S. & von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J Mol Biol* **300**, 1005-1016.

Estrada, D. E., Ewart, H. S., Tsakiridis, T., Volchuk, A., Ramlal, T., Tritschler, H. & Klip, A. (1996). Stimulation of glucose uptake by the natural coenzyme alpha-lipoic acid/thioctic acid: participation of elements of the insulin signaling pathway. *Diabetes* **45**, 1798-1804.

Ewald, R., Kolukisaoglu, U., Bauwe, U., Mikkat, S. & Bauwe, H. (2007). Mitochondrial protein lipoylation does not exclusively depend on the mtKAS pathway of *de novo* fatty acid synthesis in *Arabidopsis*. *Plant Physiol* **145**, 41-48.

Fairlamb, A. H. & Opperdoes, F. (1986). Carbohydrate metabolism in african trypanosomes, with special reference to the glycosome. In *Carbohydrate Metabolism in Cultured Cells*. Edited by M. J. Morgan: Plenum Publishing Corporation.

Feng, D., Witkowski, A. & Smith, S. (2009a). Downregulation of mitochondrial acyl carrier protein in mammalian cells compromises protein lipoylation and respiratory complex I and results in cell death. *J Biol Chem*.

Feng, X., Rodriguez-Contreras, D., Buffalo, C., Bouwer, H. G., Kruvand, E., Beverley, S. M. & Landfear, S. M. (2009b). Amplification of an alternate transporter gene suppresses the avirulent phenotype of glucose transporter null mutants in *Leishmania mexicana*. *Mol Microbiol* **71**, 369-381.

Flinn, H. M., Rangarajan, D. & Smith, D. F. (1994). Expression of a hydrophilic surface protein in infective stages of *Leishmania major*. *Mol Biochem Parasitol* **65**, 259-270.

Foucher, A. L., Papadopoulou, B. & Ouellette, M. (2006). Prefractionation by digitonin extraction increases representation of the cytosolic and intracellular proteome of *Leishmania infantum*. *J Proteome Res* **5**, 1741-1750.

Frank, R. A., Price, A. J., Northrop, F. D., Perham, R. N. & Luisi, B. F. (2007). Crystal structure of the E1 component of the *Escherichia coli* 2-oxoglutarate dehydrogenase multienzyme complex. *J Mol Biol* **368**, 639-651.

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

Frey, P. A., Hegeman, A. D. & Ruzicka, F. J. (2008). The Radical SAM Superfamily. *Crit Rev Biochem Mol Biol* **43**, 63-88.

Fries, M., Stott, K. M., Reynolds, S. & Perham, R. N. (2007). Distinct modes of recognition of the lipoyl domain as substrate by the E1 and E3 components of the pyruvate dehydrogenase multienzyme complex. *J Mol Biol* **366**, 132-139.

Frost, S. C. & Lane, M. D. (1985). Evidence for the involvement of vicinal sulfhydryl groups in insulin-activated hexose transport by 3T3-L1 adipocytes. *J Biol Chem* **260**, 2646-2652.

Fujiwara, K., Okamura-Ikeda, K. & Motokawa, Y. (1997a). Cloning and expression of a cDNA encoding bovine lipoyltransferase. *J Biol Chem* **272**, 31974-31978.

Fujiwara, K., Okamura-Ikeda, K., Packer, L. & Motokawa, Y. (1997b). Synthesis and characterization of selenolipoylated H-protein of the glycine cleavage system. *J Biol Chem* **272**, 19880-19883.

Fujiwara, K., Suzuki, M., Okumachi, Y., Okamura-Ikeda, K., Fujiwara, T., Takahashi, E. & Motokawa, Y. (1999). Molecular cloning, structural characterization and chromosomal localization of human lipoyltransferase gene. *Eur J Biochem* **260**, 761-767.

Fujiwara, K., Takeuchi, S., Okamura-Ikeda, K. & Motokawa, Y. (2001). Purification, characterization, and cDNA cloning of lipoate-activating enzyme from bovine liver. *J Biol Chem* **276**, 28819-28823.

Fujiwara, K., Toma, S., Okamura-Ikeda, K., Motokawa, Y., Nakagawa, A. & Taniguchi, H. (2005). Crystal structure of lipoate-protein ligase A from *Escherichia coli*. Determination of the lipoic acid-binding site. *J Biol Chem* **280**, 33645-33651.

Fujiwara, K., Hosaka, H., Matsuda, M., Okamura-Ikeda, K., Motokawa, Y., Suzuki, M., Nakagawa, A. & Taniguchi, H. (2007). Crystal structure of bovine lipoyltransferase in complex with lipoyl-AMP. *J Mol Biol* **371**, 222-234.

Gagnon, D., Foucher, A., Girard, I. & Ouellette, M. (2006). Stage specific gene expression and cellular localization of two isoforms of the serine hydroxymethyltransferase in the protozoan parasite *Leishmania*. *Mol Biochem Parasitol* **150**, 63-71.

Genest, P. A., ter Riet, B., Dumas, C., Papadopoulou, B., van Luenen, H. G. & Borst, P. (2005). Formation of linear inverted repeat amplicons following targeting of an essential gene in *Leishmania*. *Nucleic Acids Res* **33**, 1699-1709.

Gordon, S. (2003). Alternative activation of macrophages. *Nat Rev Immunol* **3**, 23-35.

Gossage, S. M., Rogers, M. E. & Bates, P. A. (2003). Two separate growth phases during the development of *Leishmania* in sand flies: implications for understanding the life cycle. *Int J Parasitol* **33**, 1027-1034.

Green, D. E., Morris, T. W., Green, J., Cronan, J. E., Jr. & Guest, J. R. (1995). Purification and properties of the lipoate protein ligase of *Escherichia coli*. *Biochem J* **309 (Pt 3)**, 853-862.

Green, J. D., Perham, R. N., Ullrich, S. J. & Appella, E. (1992). Conformational studies of the interdomain linker peptides in the dihydrolipoyl acetyltransferase component of the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J Biol Chem* **267**, 23484-23488.

Gueguen, V., Macherel, D., Jaquinod, M., Douce, R. & Bourguignon, J. (2000). Fatty acid and lipoic acid biosynthesis in higher plant mitochondria. *J Biol Chem* **275**, 5016-5025.

Guest, J. R., Lewis, H. M., Graham, L. D., Packman, L. C. & Perham, R. N. (1985). Genetic reconstruction and functional analysis of the repeating lipoyl domains in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *J Mol Biol* **185**, 743-754.

Guler, J. L., Kriegova, E., Smith, T. K., Lukes, J. & Englund, P. T. (2008). Mitochondrial fatty acid synthesis is required for normal mitochondrial morphology and function in *Trypanosoma brucei*. *Mol Microbiol* **67**, 1125-1142.

Gunther, S., McMillan, P. J., Wallace, L. J. & Muller, S. (2005). *Plasmodium falciparum* possesses organelle-specific alpha-keto acid dehydrogenase complexes and lipoylation pathways. *Biochem Soc Trans* **33**, 977-980.

Gunther, S., Wallace, L., Patzewitz, E. M., McMillan, P. J., Storm, J., Wrenger, C., Bissett, R., Smith, T. K. & Muller, S. (2007). Apicoplast lipoic acid protein ligase B is not essential for *Plasmodium falciparum*. *PLoS Pathog* **3**, e189.

Gunther, S., Matuschewski, K. & Muller, S. (2009a). Knockout studies reveal an important role of *Plasmodium* lipoic Acid protein ligase A1 for asexual blood stage parasite survival. *PLoS ONE* **4**, e5510.

Gunther, S., Storm, J. & Muller, S. (2009b). *Plasmodium falciparum*: organelle-specific acquisition of lipoic acid. *Int J Biochem Cell Biol* **41**, 748-752.

Haanstra, J. R., van Tuijl, A., Kessler, P., Reijnders, W., Michels, P. A., Westerhoff, H. V., Parsons, M. & Bakker, B. M. (2008). Compartmentation prevents a lethal turbo-explosion of glycolysis in trypanosomes. *Proc Natl Acad Sci U S A* **105**, 17718-17723.

Hackert, M. L., Oliver, R. M. & Reed, L. J. (1983a). A computer model analysis of the active-site coupling mechanism in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *Proc Natl Acad Sci U S A* **80**, 2907-2911.

Hackert, M. L., Oliver, R. M. & Reed, L. J. (1983b). Evidence for a multiple random coupling mechanism in the alpha-ketoglutarate dehydrogenase multienzyme complex of *Escherichia coli*: a computer model analysis. *Proc Natl Acad Sci U S A* **80**, 2226-2230.

Haenen, G. R. & Bast, A. (1991). Scavenging of hypochlorous acid by lipoic acid. *Biochem Pharmacol* **42**, 2244-2246.

Hager, K., Marahrens, A., Kenklies, M., Riederer, P. & Munch, G. (2001). Alpha-lipoic acid as a new treatment option for Azheimer type dementia. *Arch Gerontol Geriatr* **32**, 275-282.

Handman, E. & Bullen, D. V. (2002). Interaction of *Leishmania* with the host macrophage. *Trends Parasitol* **18**, 332-334.

Hart, D. T. & Coombs, G. H. (1982). *Leishmania mexicana*: energy metabolism of amastigotes and promastigotes. *Exp Parasitol* **54**, 397-409.

Hayakawa, K. & Oizumi, J. (1988). Human serum lipoamidase. *Enzyme* **40**, 30-36.

Hayakawa, K., Guo, L., Terentyeva, E. A., Li, X. K., Kimura, H., Hirano, M., Yoshikawa, K., Nagamine, T., Katsumata, N., Ogata, T. & Tanaka, T. (2006). Determination of specific activities and kinetic constants of biotinidase and lipoamidase in LEW rat and *Lactobacillus casei* (*Shirota*). *J Chromatogr B Analyt Technol Biomed Life Sci* **844**, 240-250.

Hayden, M. A., Huang, I., Bussiere, D. E. & Ashley, G. W. (1992). The biosynthesis of lipoic acid. Cloning of lip, a lipoate biosynthetic locus of *Escherichia coli*. *J Biol Chem* **267**, 9512-9515.

Hellemond, J. J., Bakker, B. M. & Tielens, A. G. (2005). Energy metabolism and its compartmentation in *Trypanosoma brucei*. *Adv Microb Physiol* **50**, 199-226.

Herwaldt, B. L. (1999). Leishmaniasis. *Lancet* **354**, 1191-1199.

Hiltunen, J. K., Okubo, F., Kursu, V. A., Autio, K. J. & Kastaniotis, A. J. (2005). Mitochondrial fatty acid synthesis and maintenance of respiratory competent mitochondria in yeast. *Biochem Soc Trans* **33**, 1162-1165.

Hiltunen, J. K., Schonauer, M. S., Autio, K. J., Mittelmeier, T. M., Kastaniotis, A. J. & Dieckmann, C. L. (2009). Mitochondrial Fatty Acid Synthesis Type II: More than Just Fatty Acids. *J Biol Chem* **284**, 9011-9015.

Huang, B., Gudi, R., Wu, P., Harris, R. A., Hamilton, J. & Popov, K. M. (1998). Isoenzymes of pyruvate dehydrogenase phosphatase. DNA-derived amino acid sequences, expression, and regulation. *J Biol Chem* **273**, 17680-17688.

Ilg, T., Handman, E. & Stierhof, Y. D. (1999). Proteophosphoglycans from *Leishmania* promastigotes and amastigotes. *Biochem Soc Trans* **27**, 518-525.

Ilg, T., Demar, M. & Harbecke, D. (2001). Phosphoglycan repeat-deficient *Leishmania mexicana* parasites remain infectious to macrophages and mice. *J Biol Chem* **276**, 4988-4997.

Islam, M. M., Wallin, R., Wynn, R. M., Conway, M., Fujii, H., Mobley, J. A., Chuang, D. T. & Hutson, S. M. (2007). A novel branched-chain amino acid metabolon. Protein-protein interactions in a supramolecular complex. *J Biol Chem* **282**, 11893-11903.

Ivens, A. C., Peacock, C. S., Worthey, E. A., Murphy, L., Aggarwal, G., Berriman, M., Sisk, E., Rajandream, M. A., Adlem, E., Aert, R., Anupama, A., Apostolou, Z., Attipoe, P., Bason, N., Bauser, C., Beck, A., Beverley, S. M., Bianchettin, G., Borzym, K., Bothe, G., Bruschi, C. V., Collins, M., Cadag, E., Ciarloni, L., Clayton, C., Coulson, R. M., Cronin, A., Cruz, A. K., Davies, R. M., De Gaudenzi, J., Dobson, D. E., Duesterhoeft, A., Fazelina, G., Fosker, N., Frasch, A. C., Fraser, A., Fuchs, M., Gabel, C., Goble, A., Goffeau, A., Harris, D., Hertz-Fowler, C., Hilbert, H., Horn, D., Huang, Y., Klages, S., Knights, A., Kube, M., Larke, N., Litvin, L., Lord, A., Louie, T., Marra, M., Masuy, D., Matthews, K., Michaeli, S., Mottram, J. C., Muller-Auer, S., Munden, H., Nelson, S., Norbertczak, H., Oliver, K., O'Neil, S., Pentony, M., Pohl, T. M., Price, C., Purnelle, B., Quail, M. A., Rabbinowitsch, E., Reinhardt, R., Rieger, M., Rinta, J., Robben, J., Robertson, L., Ruiz, J. C., Rutter, S., Saunders, D., Schafer, M., Schein, J., Schwartz, D. C., Seeger, K., Seyler, A., Sharp, S., Shin, H., Sivam, D., Squares, R., Squares, S., Tosato, V., Vogt, C., Volckaert, G., Wambutt, R., Warren, T., Wedler, H., Woodward, J., Zhou,

S., Zimmermann, W., Smith, D. F., Blackwell, J. M., Stuart, K. D., Barrell, B. & Myler, P. J. (2005). The genome of the kinetoplastid parasite, *Leishmania major*. *Science* **309**, 436-442.

Jacob, S., Ruus, P., Hermann, R., Tritschler, H. J., Maerker, E., Renn, W., Augustin, H. J., Dietze, G. J. & Rett, K. (1999). Oral administration of RAC-alpha-lipoic acid modulates insulin sensitivity in patients with type-2 diabetes mellitus: a placebo-controlled pilot trial. *Free Radic Biol Med* **27**, 309-314.

Jacobson, R. L. & Schlein, Y. (2001). *Phlebotomus papatasi* and *Leishmania major* parasites express alpha-amylase and alpha-glucosidase. *Acta Trop* **78**, 41-49.

Jacobson, R. L., Schlein, Y. & Eisenberger, C. L. (2001). The biological function of sand fly and *Leishmania* glycosidases. *Med Microbiol Immunol* **190**, 51-55.

Jarrett, J. T. (2005). The novel structure and chemistry of iron-sulfur clusters in the adenosylmethionine-dependent radical enzyme biotin synthase. *Arch Biochem Biophys* **433**, 312-321.

Jiang, Y. & Cronan, J. E. (2005). Expression cloning and demonstration of *Enterococcus faecalis* lipoamidase (pyruvate dehydrogenase inactivase) as a Ser-Ser-Lys triad amidohydrolase. *J Biol Chem* **280**, 2244-2256.

Jocelyn, P. C. (1967). The standard redox potential of cysteine-cystine from the thiol-disulphide exchange reaction with glutathione and lipoic acid. *Eur J Biochem* **2**, 327-331.

Jones, D. D. & Perham, R. N. (2008). The role of loop and beta-turn residues as structural and functional determinants for the lipoyl domain from the *Escherichia coli* 2-oxoglutarate dehydrogenase complex. *Biochem J* **409**, 357-366.

Jordan, S. W. & Cronan, J. E., Jr. (1997a). Biosynthesis of lipoic acid and posttranslational modification with lipoic acid in *Escherichia coli*. *Methods Enzymol* **279**, 176-183.

Jordan, S. W. & Cronan, J. E., Jr. (1997b). A new metabolic link. The acyl carrier protein of lipid synthesis donates lipoic acid to the pyruvate dehydrogenase complex in *Escherichia coli* and mitochondria. *J Biol Chem* **272**, 17903-17906.

Jordan, S. W. & Cronan, J. E., Jr. (2002). Chromosomal amplification of the *Escherichia coli* lipB region confers high-level resistance to selenolipoic acid. *J Bacteriol* **184**, 5495-5501.

Jordan, S. W. & Cronan, J. E., Jr. (2003). The *Escherichia coli lipB* gene encodes lipoyl (octanoyl)-acyl carrier protein:protein transferase. *J Bacteriol* **185**, 1582-1589.

Joshi, M., Jeoung, N. H., Popov, K. M. & Harris, R. A. (2007). Identification of a novel PP2C-type mitochondrial phosphatase. *Biochem Biophys Res Commun* **356**, 38-44.

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

Kagan, V. E., Shvedova, A., Serbinova, E., Khan, S., Swanson, C., Powell, R. & Packer, L. (1992). Dihydrolipoic acid--a universal antioxidant both in the membrane and in the aqueous phase. Reduction of peroxyl, ascorbyl and chromanoxyl radicals. *Biochem Pharmacol* **44**, 1637-1649.

Kamhawi, S. (2006). Phlebotomine sand flies and *Leishmania* parasites: friends or foes? *Trends Parasitol* **22**, 439-445.

Kang, S. G., Jeong, H. K., Lee, E. & Natarajan, S. (2007). Characterization of a lipoate-protein ligase A gene of rice (*Oryza sativa* L.). *Gene* **393**, 53-61.

Kato, M., Chuang, J. L., Tso, S. C., Wynn, R. M. & Chuang, D. T. (2005). Crystal structure of pyruvate dehydrogenase kinase 3 bound to lipoyl domain 2 of human pyruvate dehydrogenase complex. *Embo J* **24**, 1763-1774.

Kaur, K. J. & Ruben, L. (1994). Protein translation elongation factor-1 alpha from *Trypanosoma brucei* binds calmodulin. *J Biol Chem* **269**, 23045-23050.

Keeney, K. M., Stuckey, J. A. & O'Riordan, M. X. (2007). LplA1-dependent utilization of host lipoyl peptides enables *Listeria* cytosolic growth and virulence. *Mol Microbiol* **66**, 758-770.

Kim do, J., Kim, K. H., Lee, H. H., Lee, S. J., Ha, J. Y., Yoon, H. J. & Suh, S. W. (2005). Crystal structure of lipoate-protein ligase A bound with the activated intermediate: insights into interaction with lipoyl domains. *J Biol Chem* **280**, 38081-38089.

Kim do, J., Lee, S. J., Kim, H. S., Kim, K. H., Lee, H. H., Yoon, H. J. & Suh, S. W. (2008). Structural basis of octanoic acid recognition by lipoate-protein ligase B. *Proteins* **70**, 1620-1625.

Koretke, K. K., Lupas, A. N., Warren, P. V., Rosenberg, M. & Brown, J. R. (2000). Evolution of two-component signal transduction. *Mol Biol Evol* **17**, 1956-1970.

Krauth-Siegel, R. L. & Schoneck, R. (1995). Flavoprotein structure and mechanism. 5. Trypanothione reductase and lipoamide dehydrogenase as targets for a structure-based drug design. *Faseb J* **9**, 1138-1146.

Kume, A., Koyata, H., Sakakibara, T., Ishiguro, Y., Kure, S. & Hiraga, K. (1991). The glycine cleavage system. Molecular cloning of the chicken and human glycine decarboxylase cDNAs and some characteristics involved in the deduced protein structures. *J Biol Chem* **266**, 3323-3329.

Kwon, K. & Beckett, D. (2000). Function of a conserved sequence motif in biotin holoenzyme synthetases. *Protein Sci* **9**, 1530-1539.

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

Lamour, N., Riviere, L., Coustou, V., Coombs, G. H., Barrett, M. P. & Bringaud, F. (2005). Proline metabolism in procyclic *Trypanosoma brucei* is down-regulated in the presence of glucose. *J Biol Chem* **280**, 11902-11910.

Laufs, H., Muller, K., Fleischer, J., Reiling, N., Jahnke, N., Jensenius, J. C., Solbach, W. & Laskay, T. (2002). Intracellular survival of *Leishmania major* in neutrophil granulocytes after uptake in the absence of heat-labile serum factors. *Infect Immun* **70**, 826-835.

Lawlis, V. B. & Roche, T. E. (1981). Regulation of bovine kidney alpha-ketoglutarate dehydrogenase complex by calcium ion and adenine nucleotides. Effects on S0.5 for alpha-ketoglutarate. *Biochemistry* **20**, 2512-2518.

Lawson, J. E., Niu, X. D., Browning, K. S., Trong, H. L., Yan, J. & Reed, L. J. (1993). Molecular cloning and expression of the catalytic subunit of bovine pyruvate dehydrogenase phosphatase and sequence similarity with protein phosphatase 2C. *Biochemistry* **32**, 8987-8993.

Lawson, J. E., Park, S. H., Mattison, A. R., Yan, J. & Reed, L. J. (1997). Cloning, expression, and properties of the regulatory subunit of bovine pyruvate dehydrogenase phosphatase. *J Biol Chem* **272**, 31625-31629.

Lee, H. H., Kim, D. J., Ahn, H. J., Ha, J. Y. & Suh, S. W. (2004). Crystal structure of T-protein of the glycine cleavage system. Cofactor binding, insights into H-protein recognition, and molecular basis for understanding nonketotic hyperglycinemia. *J Biol Chem* **279**, 50514-50523.

Lee, S. H., Stephens, J. L., Paul, K. S. & Englund, P. T. (2006). Fatty acid synthesis by elongases in trypanosomes. *Cell* **126**, 691-699.

Lee, S. H., Stephens, J. L. & Englund, P. T. (2007). A fatty-acid synthesis mechanism specialized for parasitism. *Nat Rev Microbiol* **5**, 287-297.

Leonard, A. E., Pereira, S. L., Sprecher, H. & Huang, Y. S. (2004). Elongation of long-chain fatty acids. *Prog Lipid Res* **43**, 36-54.

Leroux, A., Fleming-Canepa, X., Aranda, A., Maugeri, D., Cazzulo, J. J., Sanchez, M. A. & Nowicki, C. (2006). Functional characterization and subcellular localization of the three malate dehydrogenase isozymes in *Leishmania spp*. *Mol Biochem Parasitol* **149**, 74-85.

Lessard, I. A. & Perham, R. N. (1994). Expression in *Escherichia coli* of genes encoding the E1 alpha and E1 beta subunits of the pyruvate dehydrogenase complex of *Bacillus stearothermophilus* and assembly of a functional E1 component (alpha 2 beta 2) in vitro. *J Biol Chem* **269**, 10378-10383.

Liu, B., Liu, Y., Motyka, S. A., Agbo, E. E. & Englund, P. T. (2005). Fellowship of the rings: the replication of kinetoplast DNA. *Trends Parasitol* **21**, 363-369.

Lotierzo, M., Tse Sum Bui, B., Florentin, D., Escalettes, F. & Marquet, A. (2005). Biotin synthase mechanism: an overview. *Biochem Soc Trans* **33**, 820-823.

Lutziger, I. & Oliver, D. J. (2000). Molecular evidence of a unique lipoamide dehydrogenase in plastids: analysis of plastidic lipoamide dehydrogenase from Arabidopsis thaliana. *FEBS Lett* **484**, 12-16.

Ma, Q., Zhao, X., Nasser Eddine, A., Geerlof, A., Li, X., Cronan, J. E., Kaufmann, S. H. & Wilmanns, M. (2006). The *Mycobacterium tuberculosis* LipB enzyme functions as a cysteine/lysine dyad acyltransferase. *Proc Natl Acad Sci U S A* **103**, 8662-8667.

Macherel, D., Bourguignon, J. & Douce, R. (1992). Cloning of the gene (gdcH) encoding H-protein, a component of the glycine decarboxylase complex of pea (*Pisum sativum L.*). *Biochem J* **286 (Pt 2)**, 627-630.

Machius, M., Chuang, J. L., Wynn, R. M., Tomchick, D. R. & Chuang, D. T. (2001). Structure of rat BCKD kinase: nucleotide-induced domain communication in a mitochondrial protein kinase. *Proc Natl Acad Sci U S A* **98**, 11218-11223.

Maeng, C. Y., Yazdi, M. A., Niu, X. D., Lee, H. Y. & Reed, L. J. (1994). Expression, purification, and characterization of the dihydrolipoamide dehydrogenase-binding protein of the pyruvate dehydrogenase complex from *Saccharomyces cerevisiae*. *Biochemistry* **33**, 13801-13807.

Magnuson, K., Jackowski, S., Rock, C. O. & Cronan, J. E., Jr. (1993). Regulation of fatty acid biosynthesis in *Escherichia coli*. *Microbiol Rev* **57**, 522-542.

Maitra, I., Serbinova, E., Trischler, H. & Packer, L. (1995). Alpha-lipoic acid prevents buthionine sulfoximine-induced cataract formation in newborn rats. *Free Radic Biol Med* **18**, 823-829.

Marche, S., Michels, P. A. & Opperdoes, F. R. (2000). Comparative study of *Leishmania mexicana* and *Trypanosoma brucei* NAD-dependent glycerol-3-phosphate dehydrogenase. *Mol Biochem Parasitol* **106**, 83-91.

Martinez-Calvillo, S., Stuart, K. & Myler, P. J. (2005). Ploidy changes associated with disruption of two adjacent genes on *Leishmania major* chromosome 1. *Int J Parasitol* **35**, 419-429.

Massey, L. K., Sokatch, J. R. & Conrad, R. S. (1976). Branched-chain amino acid catabolism in bacteria. *Bacteriol Rev* **40**, 42-54.

Mattevi, A., Schierbeek, A. J. & Hol, W. G. (1991). Refined crystal structure of lipoamide dehydrogenase from *Azotobacter vinelandii* at 2.2 A resolution. A comparison with the structure of glutathione reductase. *J Mol Biol* **220**, 975-994.

Mattevi, A., Obmolova, G., Sokatch, J. R., Betzel, C. & Hol, W. G. (1992). The refined crystal structure of *Pseudomonas putida* lipoamide dehydrogenase complexed with NAD+ at 2.45 A resolution. *Proteins* **13**, 336-351.

Maugeri, D. A., Cazzulo, J. J., Burchmore, R. J. S., Barrett, M. P. & Ogbunude, P. O. J. (2003). Pentose phosphate metabolism in *Leishmania mexicana*. *Molecular and Biochemical Parasitology* **130**, 117.

Mazumdar, J., E, H. W., Masek, K., C, A. H. & Striepen, B. (2006). Apicoplast fatty acid synthesis is essential for organelle biogenesis and parasite survival in *Toxoplasma gondii*. *Proc Natl Acad Sci U S A* **103**, 13192-13197.

McConville, M. J. & Ferguson, M. A. (1993). The structure, biosynthesis and function of glycosylated phosphatidylinositols in the parasitic protozoa and higher eukaryotes. *Biochem J* **294 (Pt 2)**, 305-324.

McConville, M. J., Mullin, K. A., Ilgoutz, S. C. & Teasdale, R. D. (2002). Secretory pathway of trypanosomatid parasites. *Microbiol Mol Biol Rev* **66**, 122-154; table of contents.

McConville, M. J., de Souza, D., Saunders, E., Likic, V. A. & Naderer, T. (2007). Living in a phagolysosome; metabolism of *Leishmania* amastigotes. *Trends Parasitol* **23**, 368-375.

McManus, E., Luisi, B. F. & Perham, R. N. (2005). Structure of a putative lipoate protein ligase from *Thermoplasma acidophilum* and the mechanism of target selection for post-translational modification. *J Mol Biol*.

McMillan, P. J., Stimmler, L. M., Foth, B. J., McFadden, G. I. & Muller, S. (2005). The human malaria parasite *Plasmodium falciparum* possesses two distinct dihydrolipoamide dehydrogenases. *Mol Microbiol* **55**, 27-38.

Michels, P. A. M., Bringaud, F., Herman, M. & Hannaert, V. (2006). Metabolic functions of glycosomes in trypanosomatids. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1763**, 1463.

Mikolajczyk, S. & Brody, S. (1990). *De novo* fatty acid synthesis mediated by acyl-carrier protein in *Neurospora crassa* mitochondria. *Eur J Biochem* **187**, 431-437.

Miller, J. R., Busby, R. W., Jordan, S. W., Cheek, J., Henshaw, T. F., Ashley, G. W., Broderick, J. B., Cronan, J. E., Jr. & Marletta, M. A. (2000). *Escherichia coli* LipA is a lipoyl synthase: in vitro biosynthesis of lipoylated pyruvate dehydrogenase complex from octanoyl-acyl carrier protein. *Biochemistry* **39**, 15166-15178.

Miranda, M. R., Canepa, G. E., Bouvier, L. A. & Pereira, C. A. (2008). *Trypanosoma cruzi*: multiple nucleoside diphosphate kinase isoforms in a single cell. *Exp Parasitol* **120**, 103-107.

Misslitz, A., Mottram, J. C., Overath, P. & Aebischer, T. (2000). Targeted integration into a rRNA locus results in uniform and high level expression of transgenes in *Leishmania* amastigotes. *Mol Biochem Parasitol* **107**, 251-261.

Mohebali, M., Fotouhi, A., Hooshmand, B., Zarei, Z., Akhoundi, B., Rahnema, A., Razaghian, A. R., Kabir, M. J. and Nadim, A. (2007). Comparison of miltefosine and meglumine antimoniate for the treatment of zoonotic cutaneous leishmaniasis (ZCL) by a randomized clinical trial in Iran. *Acta Trop* **103**, 33-40.

Moon, Y. A., Shah, N. A., Mohapatra, S., Warrington, J. A. & Horton, J. D. (2001). Identification of a mammalian long chain fatty acyl elongase regulated by sterol regulatory element-binding proteins. *J Biol Chem* **276**, 45358-45366.

Mooney, B. P., Miernyk, J. A. & Randall, D. D. (2002). The complex fate of alpha-ketoacids. *Annu Rev Plant Biol* **53**, 357-375.

Morikawa, T., Yasuno, R. & Wada, H. (2001). Do mammalian cells synthesize lipoic acid? Identification of a mouse cDNA encoding a lipoic acid synthase located in mitochondria. *FEBS Letters* **498**, 16.

Morita, Y. S., Paul, K. S. & Englund, P. T. (2000). Specialized fatty acid synthesis in African trypanosomes: myristate for GPI anchors. *Science* **288**, 140-143.

Morris, T. W., Reed, K. E. & Cronan, J. E., Jr. (1994). Identification of the gene encoding lipoate-protein ligase A of *Escherichia coli*. Molecular cloning and characterization of the lplA gene and gene product. *J Biol Chem* **269**, 16091-16100.

Morris, T. W., Reed, K. E. & Cronan, J. E., Jr. (1995). Lipoic acid metabolism in *Escherichia coli*: the *lplA* and *lipB* genes define redundant pathways for ligation of lipoyl groups to apoprotein. *J Bacteriol* **177**, 1-10.

Mottram, J. C., McCready, B. P., Brown, K. G. & Grant, K. M. (1996). Gene disruptions indicate an essential function for the *Lmm*CRK1 cdc2-related kinase of *Leishmania mexicana*. *Mol Microbiol* **22**, 573-583.

Muller, K., van Zandbergen, G., Hansen, B., Laufs, H., Jahnke, N., Solbach, W. & Laskay, T. (2001). Chemokines, natural killer cells and granulocytes in the early course of *Leishmania major* infection in mice. *Med Microbiol Immunol* **190**, 73-76.

Murray, H. W., Berman, J. D., Davies, C. R. & Saravia, N. G. (2005). Advances in *leishmaniasis*. *Lancet* **366**, 1561-1577.

Myskova, J., Svobodova, M., Beverley, S. M. & Volf, P. (2007). A lipophosphoglycan-independent development of *Leishmania* in permissive sand flies. *Microbes Infect* **9**, 317-324.

Naderer, T., Ellis, M. A., Sernee, M. F., De Souza, D. P., Curtis, J., Handman, E. & McConville, M. J. (2006). Virulence of *Leishmania major* in macrophages and mice requires the gluconeogenic enzyme fructose-1,6-bisphosphatase. *Proc Natl Acad Sci U S A* **103**, 5502-5507.

Naderer, T. & McConville, M. J. (2008). The *Leishmania*-macrophage interaction: a metabolic perspective. *Cell Microbiol* **10**, 301-308.

Nakai, T., Ishijima, J., Masui, R., Kuramitsu, S. & Kamiya, N. (2003a). Structure of *Thermus thermophilus* HB8 H-protein of the glycine-cleavage system, resolved by a six-dimensional molecular-replacement method. *Acta Crystallogr D Biol Crystallogr* **59**, 1610-1618.

Nakai, T., Nakagawa, N., Maoka, N., Masui, R., Kuramitsu, S. & Kamiya, N. (2003b). Coexpression, purification, crystallization and preliminary X-ray characterization of glycine decarboxylase (P-protein) of the glycine-cleavage system from *Thermus thermophilus* HB8. *Acta Crystallogr D Biol Crystallogr* **59**, 554-557.

Nakai, T., Nakagawa, N., Maoka, N., Masui, R., Kuramitsu, S. & Kamiya, N. (2005). Structure of P-protein of the glycine cleavage system: implications for nonketotic hyperglycinemia. *Embo J* **24**, 1523-1536.

Nauseef, W. M. (2007). How human neutrophils kill and degrade microbes: an integrated view. *Immunol Rev* **219**, 88-102.

Neagle, J., De Marcucci, O., Dunbar, B. & Lindsay, J. G. (1989). Component X of mammalian pyruvate dehydrogenase complex: structural and functional relationship to the lipoate acetyltransferase (E2) component. *FEBS Lett* **253**, 11-15.

Nelson, L. M. & Cox, M. M. (2000). *Lehniger Principles of Biochemistry*, 3 edn. New York: Worth Publishers.

Nikolau, B. J., Ohlrogge, J. B. & Wurtele, E. S. (2003). Plant biotin-containing carboxylases. *Arch Biochem Biophys* **414**, 211-222.

Norton, L. E. & Layman, D. K. (2006). Leucine regulates translation initiation of protein synthesis in skeletal muscle after exercise. *J Nutr* **136**, 533S-537S.

O'Riordan, M., Moors, M. A. & Portnoy, D. A. (2003). *Listeria* intracellular growth and virulence require host-derived lipoic acid. *Science* **302**, 462-464.

Oh, C. S., Toke, D. A., Mandala, S. & Martin, C. E. (1997). ELO2 and ELO3, homologues of the *Saccharomyces cerevisiae* ELO1 gene, function in fatty acid elongation and are required for sphingolipid formation. *J Biol Chem* **272**, 17376-17384.

Oizumi, J. & Hayakawa, K. (1989). Liberation of lipoate by human serum lipoamidase from bovine heart pyruvate dehydrogenase. *Biochem Biophys Res Commun* **162**, 658-663.

Okamura-Ikeda, K., Ohmura, Y., Fujiwara, K. & Motokawa, Y. (1993). Cloning and nucleotide sequence of the gcv operon encoding the *Escherichia coli* glycine-cleavage system. *Eur J Biochem* **216**, 539-548.

Oliver, D. J., Neuburger, M., Bourguignon, J. & Douce, R. (1990). Interaction between the Component Enzymes of the Glycine Decarboxylase Multienzyme Complex. *Plant Physiol* **94**, 833-839.

Opperdoes, F. R. & Coombs, G. H. (2007). Metabolism of *Leishmania*: proven and predicted. *Trends Parasitol* **23**, 149-158.

Ouellette, M., Drummelsmith, J. & Papadopoulou, B. (2004). Leishmaniasis: drugs in the clinic, resistance and new developments. *Drug Resist Updat* **7**, 257-266.

Overath, P., Stierhof, Y. D. & Wiese, M. (1997). Endocytosis and secretion in trypanosomatid parasites - Tumultuous traffic in a pocket. *Trends Cell Biol* **7**, 27-33.

Pacheco-Alvarez, D., Solorzano-Vargas, R. S. & Del Rio, A. L. (2002). Biotin in metabolism and its relationship to human disease. *Arch Med Res* **33**, 439-447.

Packman, L. C., Hale, G. & Perham, R. N. (1984). Repeating functional domains in the pyruvate dehydrogenase multienzyme complex of *Escherichia coli*. *Embo J* **3**, 1315-1319.

Pares, S., Cohen-Addad, C., Sieker, L., Neuburger, M. & Douce, R. (1994). X-ray structure determination at 2.6-A resolution of a lipoate-containing protein: the H-protein of the glycine decarboxylase complex from pea leaves. *Proc Natl Acad Sci U S A* **91**, 4850-4853.

Perham, R. N., Duckworth, H. W. & Roberts, G. C. (1981). Mobility of polypeptide chain in the pyruvate dehydrogenase complex revealed by proton NMR. *Nature* **292**, 474-477.

Perham, R. N. (1991). Domains, motifs, and linkers in 2-oxo acid dehydrogenase multienzyme complexes: a paradigm in the design of a multifunctional protein. *Biochemistry* **30**, 8501-8512.

Perham, R. N. (2000). Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu Rev Biochem* **69**, 961-1004.

Peters, N. C., Egen, J. G., Secundino, N., Debrabant, A., Kimblin, N., Kamhawi, S., Lawyer, P., Fay, M. P., Germain, R. N. & Sacks, D. (2008). *In vivo* imaging reveals an essential role for neutrophils in leishmaniasis transmitted by sand flies. *Science* **321**, 970-974.

Podda, M., Tritschler, H. J., Ulrich, H. & Packer, L. (1994). Alpha-lipoic acid supplementation prevents symptoms of vitamin E deficiency. *Biochem Biophys Res Commun* **204**, 98-104.

Rachman, H., Strong, M., Ulrichs, T., Grode, L., Schuchhardt, J., Mollenkopf, H., Kosmiadi, G. A., Eisenberg, D. & Kaufmann, S. H. (2006). Unique transcriptome signature of *Mycobacterium tuberculosis* in pulmonary tuberculosis. *Infect Immun* **74**, 1233-1242.

Ralton, J. E., Naderer, T., Piraino, H. L., Bashtannyk, T. A., Callaghan, J. M. & McConville, M. J. (2003). Evidence that intracellular beta1-2 mannan is a virulence factor in *Leishmania parasites*. *J Biol Chem* **278**, 40757-40763.

Reche, P., Li, Y. L., Fuller, C., Eichhorn, K. & Perham, R. N. (1998). Selectivity of post-translational modification in biotinylated proteins: the carboxy carrier protein of the acetyl-CoA carboxylase of *Escherichia coli*. *Biochem J* **329 (Pt 3)**, 589-596.

Reche, P. & Perham, R. N. (1999). Structure and selectivity in post-translational modification: attaching the biotinyl-lysine and lipoyl-lysine swinging arms in multifunctional enzymes. *Embo J* **18**, 2673-2682.

Reche, P. A. (2000). Lipoylating and biotinylating enzymes contain a homologous catalytic module. *Protein Sci* **9**, 1922-1929.

Reed, K. E. & Cronan, J. E., Jr. (1993). Lipoic acid metabolism in *Escherichia coli*: sequencing and functional characterization of the lipA and lipB genes. *J Bacteriol* **175**, 1325-1336.

Reed, K. E., Morris, T. W. & Cronan, J. E., Jr. (1994). Mutants of *Escherichia coli* K-12 that are resistant to a selenium analog of lipoic acid identify unknown genes in lipoate metabolism. *Proc Natl Acad Sci U S A* **91**, 3720-3724.

Reed, L. J., De, B. B., Gunsalus, I. C. & Hornberger, C. S., Jr. (1951). Crystalline alpha-lipoic acid; a catalytic agent associated with pyruvate dehydrogenase. *Science* **114**, 93-94.

Reed, L. J. & Hackert, M. L. (1990). Structure-function relationships in dihydrolipoamide acyltransferases. *J Biol Chem* **265**, 8971-8974.

Richarme, G. & Heine, H. G. (1986). Galactose- and maltose-stimulated lipoamide dehydrogenase activities related to the binding-protein-dependent transport of galactose and maltose in toluenized cells of *Escherichia coli*. *Eur J Biochem* **156**, 399-405.

Rittig, M. G. & Bogdan, C. (2000). *Leishmania*-host-cell interaction: complexities and alternative views. *Parasitol Today* **16**, 292-297.

Robinson, K. A. & Beverley, S. M. (2003). Improvements in transfection efficiency and tests of RNA interference (RNAi) approaches in the protozoan parasite *Leishmania*. *Mol Biochem Parasitol* **128**, 217-228.

Roche, T. E., Baker, J. C., Yan, X., Hiromasa, Y., Gong, X., Peng, T., Dong, J., Turkan, A. & Kasten, S. A. (2001). Distinct regulatory properties of pyruvate dehydrogenase kinase and phosphatase isoforms. *Prog Nucleic Acid Res Mol Biol* **70**, 33-75.

Roche, T. E., Hiromasa, Y., Turkan, A., Gong, X., Peng, T., Yan, X., Kasten, S. A., Bao, H. & Dong, J. (2003). Essential roles of lipoyl domains in the activated function and control of pyruvate dehydrogenase kinases and phosphatase isoform 1. *Eur J Biochem* **270**, 1050-1056.

Roche, T. E. & Hiromasa, Y. (2007). Pyruvate dehydrogenase kinase regulatory mechanisms and inhibition in treating diabetes, heart ischemia, and cancer. *Cell Mol Life Sci* **64**, 830-849.

Rock, C. O. & Jackowski, S. (2002). Forty years of bacterial fatty acid synthesis. *Biochem Biophys Res Commun* **292**, 1155-1166.

Rodriguez-Contreras, D., Feng, X., Keeney, K. M., Bouwer, H. G. & Landfear, S. M. (2007). Phenotypic characterization of a glucose transporter null mutant in *Leishmania mexicana*. *Mol Biochem Parasitol* **153**, 9-18.

Rogers, M. E., Chance, M. L. & Bates, P. A. (2002). The role of promastigote secretory gel in the origin and transmission of the infective stage of *Leishmania mexicana* by the sandfly *Lutzomyia longipalpis*. *Parasitology* **124**, 495-507.

Rosenberg, H. R. & Culik, R. (1959). Effect of alpha-lippoic acid on vitamin C and vitamin E deficiencies. *Arch Biochem Biophys* **80**, 86-93.

Rosenzweig, D., Smith, D., Opperdoes, F., Stern, S., Olafson, R. W. & Zilberstein, D. (2008). Retooling *Leishmania* metabolism: from sand fly gut to human macrophage. *Faseb J* **22**, 590-602.

Roy, S. & Packer, L. (1998). Redox regulation of cell functions by alpha-lipoate: biochemical and molecular aspects. *Biofactors* **8**, 17-21.

Russell, D. G., Xu, S. & Chakraborty, P. (1992). Intracellular trafficking and the parasitophorous vacuole of *Leishmania mexicana*-infected macrophages. *J Cell Sci* **103**, 1193-1210.

Rutter, G. A., McCormack, J. G., Midgley, P. J. & Denton, R. M. (1989). The role of Ca2+ in the hormonal regulation of the activities of pyruvate dehydrogenase and oxoglutarate dehydrogenase complexes. *Ann N Y Acad Sci* **573**, 206-217.

Sacks, D. & Anderson, C. (2004). Re-examination of the immunosuppressive mechanisms mediating non-cure of *Leishmania* infection in mice. *Immunol Rev* **201**, 225-238.

Sacks, D. L., Hieny, S. & Sher, A. (1985). Identification of cell surface carbohydrate and antigenic changes between noninfective and infective developmental stages of *Leishmania major* promastigotes. *J Immunol* **135**, 564-569.

Sacks, D. L., Modi, G., Rowton, E., Spath, G., Epstein, L., Turco, S. J. & Beverley, S. M. (2000). The role of phosphoglycans in *Leishmania*-sand fly interactions. *Proc Natl Acad Sci U S A* **97**, 406-411.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, 2nd edition edn: Cold Spring Harbor Laboratory Press.

Sanderson, S. J., Khan, S. S., McCartney, R. G., Miller, C. & Lindsay, J. G. (1996a). Reconstitution of mammalian pyruvate dehydrogenase and 2-oxoglutarate dehydrogenase complexes: analysis of protein X involvement and interaction of homologous and heterologous dihydrolipoamide dehydrogenases. *Biochem J* **319 (Pt 1)**, 109-116.

Sanderson, S. J., Miller, C. & Lindsay, J. G. (1996b). Stoichiometry, organisation and catalytic function of protein X of the pyruvate dehydrogenase complex from bovine heart. *Eur J Biochem* **236**, 68-77.

Santrich, C., Moore, L., Sherwin, T., Bastin, P., Brokaw, C., Gull, K. & LeBowitz, J. H. (1997). A motility function for the paraflagellar rod of *Leishmania* parasites revealed by PFR-2 gene knockouts. *Mol Biochem Parasitol* **90**, 95-109.

Sassetti, C. M., Boyd, D. H. & Rubin, E. J. (2003). Genes required for mycobacterial growth defined by high density mutagenesis. *Mol Microbiol* **48**, 77-84.

Schlein, Y. (1986). Sandfly diet and *Leishmania*. *Parasitol Today* **2**, 175-177.

Schonauer, M. S., Kastaniotis, A. J., Hiltunen, J. K. & Dieckmann, C. L. (2008). Intersection of RNA processing and the type II fatty acid synthesis pathway in yeast mitochondria. *Mol Cell Biol* **28**, 6646-6657.

Schwartz, E., Hatz, C. and Blum, J. (2006). New world cutaneous leishmaniasis in travellers. *Lancet Infect Dis* **6**, 342-349.

Scott, B. C., Aruoma, O. I., Evans, P. J., O'Neill, C., Van der Vliet, A., Cross, C. E., Tritschler, H. & Halliwell, B. (1994). Lipoic and dihydrolipoic acids as antioxidants. A critical evaluation. *Free Radic Res* **20**, 119-133.

Scott, D. A., Hickerson, S. M., Vickers, T. J. & Beverley, S. M. (2008). The role of the mitochondrial glycine cleavage complex in the metabolism and virulence of the protozoan parasite *Leishmania major*. *J Biol Chem* **283**, 155-165.

Scott, E. M., Duncan, I. W. & Ekstrand, V. (1963). Purification And Properties Of Glutathione Reductase Of Human Erythrocytes. *J Biol Chem* **238**, 3928-3933.

Searls, R. L. & Sanadi, D. R. (1960). alpha-Ketoglutaric dehydrogenase. 8. Isolation and some properties of a flavoprotein compnent. *J Biol Chem* **235**, 2485-2491.

Segal, A. W. (2005). How neutrophils kill microbes. *Annu Rev Immunol* **23**, 197-223.

Shaw, N. M., Birch, O. M., Tinschert, A., Venetz, V., Dietrich, R. & Savoy, L. A. (1998). Biotin synthase from *Escherichia coli*: isolation of an enzyme-generated intermediate and stoichiometry of S-adenosylmethionine use. *Biochem J* **330 (Pt 3)**, 1079-1085.

Smith, S. (1994). The animal fatty acid synthase: one gene, one polypeptide, seven enzymes. *Faseb J* **8**, 1248-1259.

Sofia, H. J., Chen, G., Hetzler, B. G., Reyes-Spindola, J. F. & Miller, N. E. (2001). Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res* **29**, 1097-1106.

Soto, J., Arana, B. A., Toledo, J., Rizzo, N., Vega, J. C., Diaz, A., Luz, M., Gutierrez, P., Arboleda, M., Berman, J. D., *et al.* (2004). Miltefosine for New World cutaneous leishmanisis. *Clin Infect Dis* **38**, 1266-1272.

Soto, J., Toledo, J., Valda, L., Balderrama, M., Rea, I., Parra, R., Ardiles, J., Soto, P., Gomez, A., Molleda, F., *et al.* (2007). Treatment of Bolivian mucosal leishmaniasis with miltefosine. *Clin Infect Dis* **44**, 350-356.

Spath, G. F., Epstein, L., Leader, B., Singer, S. M., Avila, H. A., Turco, S. J. & Beverley, S. M. (2000). Lipophosphoglycan is a virulence factor distinct from related glycoconjugates in the protozoan parasite *Leishmania major*. *Proc Natl Acad Sci U S A* **97**, 9258-9263.

Stephens, J. L., Lee, S. H., Paul, K. S. & Englund, P. T. (2007). Mitochondrial fatty acid synthesis in *Trypanosoma brucei*. *J Biol Chem* **282**, 4427-4436.

Stierhof, Y. D., Bates, P. A., Jacobson, R. L., Rogers, M. E., Schlein, Y., Handman, E. & Ilg, T. (1999). Filamentous proteophosphoglycan secreted by *Leishmania* promastigotes forms gel-like three-dimensional networks that obstruct the digestive tract of infected sandfly vectors. *Eur J Cell Biol* **78**, 675-689.

Suh, J. H., Zhu, B. Z., deSzoeke, E., Frei, B. & Hagen, T. M. (2004). Dihydrolipoic acid lowers the redox activity of transition metal ions but does not remove them from the active site of enzymes. *Redox Rep* **9**, 57-61.

Sulo, P. & Martin, N. C. (1993). Isolation and characterization of LIP5. A lipoate biosynthetic locus of *Saccharomyces cerevisiae*. *J Biol Chem* **268**, 17634-17639.

Sundar, S., More, D. K., Singh, M. K., Singh, V. P., Sharma, S., Makharia, A., Kumar, P. C. & Murray, H. W. (2000). Failure of pentavalent antimony in visceral leishmaniasis in India: report from the center of the Indian epidemic. *Clin Infect Dis* **31**, 1104-1107.

Sundar, S. and Rai, M. (2002). Advances in the treatment of leishmaniasis. *Curr Opin Infect Dis* **15**, 593-598.

Sundar, S., Jha, T. K., Thakur, C. P., Engel, J., Sindermann, H., Fischer, C., Junge, K., Bryceson, A. and Berman, J. (2002). Oral miltefosine for Indian visceral leishmaniasis. *N Engl J Med* **347**, 1739-1746.

Sunkin, S. M., Kiser, P., Myler, P. J. & Stuart, K. (2000). The size difference between *leishmania major* friedlin chromosome one homologues is localized to sub-telomeric repeats at one chromosomal end. *Mol Biochem Parasitol* **109**, 1-15.

Suzuki, Y. J., Tsuchiya, M. & Packer, L. (1991). Thioctic acid and dihydrolipoic acid are novel antioxidants which interact with reactive oxygen species. *Free Radic Res Commun* **15**, 255-263.

Suzuki, Y. J., Tsuchiya, M. & Packer, L. (1993). Antioxidant activities of dihydrolipoic acid and its structural homologues. *Free Radic Res Commun* **18**, 115-122.

Teague, W. M., Pettit, F. H., Wu, T. L., Silberman, S. R. & Reed, L. J. (1982). Purification and properties of pyruvate dehydrogenase phosphatase from bovine heart and kidney. *Biochemistry* **21**, 5585-5592.

Tetaud, E., Lecuix, I., Sheldrake, T., Baltz, T. & Fairlamb, A. H. (2002). A new expression vector for *Crithidia fasciculata* and *Leishmania*. *Mol Biochem Parasitol* **120**, 195-204.

Thevelein, J. M. & Hohmann, S. (1995). Trehalose synthase: guard to the gate of glycolysis in yeast? *Trends Biochem Sci* **20**, 3-10.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). Improved sensitivity of profile searches through the use of sequence weights and gap excision. *Comput Appl Biosci* **10**, 19-29.

Thomsen-Zieger, N., Schachtner, J. & Seeber, F. (2003). Apicomplexan parasites contain a single lipoic acid synthase located in the plastid. *FEBS Lett* **547**, 80-86.

Thorpe, C. & Williams, C. H., Jr. (1976). Differential reactivity of the two active site cysteine residues generated on reduction of pig heart lipoamide dehydrogenase. *J Biol Chem* **251**, 3553-3557.

Toke, D. A. & Martin, C. E. (1996). Isolation and characterization of a gene affecting fatty acid elongation in *Saccharomyces cerevisiae*. *J Biol Chem* **271**, 18413-18422.

Turco, S. J., Spath, G. F. & Beverley, S. M. (2001). Is lipophosphoglycan a virulence factor? A surprising diversity between *Leishmania* species. *Trends Parasitol* **17**, 223-226.

Ugulava, N. B., Gibney, B. R. & Jarrett, J. T. (2001a). Biotin synthase contains two distinct iron-sulfur cluster binding sites: chemical and spectroelectrochemical analysis of iron-sulfur cluster interconversions. *Biochemistry* **40**, 8343-8351.

Ugulava, N. B., Sacanell, C. J. & Jarrett, J. T. (2001b). Spectroscopic changes during a single turnover of biotin synthase: destruction of a [2Fe-2S] cluster accompanies sulfur insertion. *Biochemistry* **40**, 8352-8358.

Vaisvila, R., Rasmussen, L. J., Lobner-Olesen, A., von Freiesleben, U. & Marinus, M. G. (2000). The LipB protein is a negative regulator of dam gene expression in *Escherichia coli*. *Biochim Biophys Acta* **1494**, 43-53.

Van Hellemond, J. J. & Tielens, A. G. M. (1997). Inhibition of the respiratory chain results in a reversible metabolic arrest in *Leishmania* promastigotes. *Molecular and Biochemical Parasitology* **85**, 135.

Van Hellemond, J. J., Van der Meer, P. & Tielens, A. G. (1997). *Leishmania infantum* promastigotes have a poor capacity for anaerobic functioning and depend mainly on respiration for their energy generation. *Parasitology* **114**, 351-360.

Van Hellemond, J. J., Opperdoes, F. R. & Tielens, A. G. (1998). Trypanosomatidae produce acetate via a mitochondrial acetate:succinate CoA transferase. *Proc Natl Acad Sci U S A* **95**, 3036-3041.

van Weelden, S. W., Fast, B., Vogt, A., van der Meer, P., Saas, J., van Hellemond, J. J., Tielens, A. G. & Boshart, M. (2003). Procyclic *Trypanosoma bruce*i do not use Krebs cycle activity for energy generation. *J Biol Chem* **278**, 12854-12863.

van Weelden, S. W., van Hellemond, J. J., Opperdoes, F. R. & Tielens, A. G. (2005). New functions for parts of the Krebs cycle in procyclic *Trypanosoma brucei*, a cycle not operating as a cycle. *J Biol Chem* **280**, 12451-12460.

van Zandbergen, G., Klinger, M., Mueller, A., Dannenberg, S., Gebert, A., Solbach, W. & Laskay, T. (2004). Cutting edge: neutrophil granulocyte serves as a vector for *Leishmania* entry into macrophages. *J Immunol* **173**, 6521-6525.

Vanden Boom, T. J., Reed, K. E. & Cronan, J. E., Jr. (1991). Lipoic acid metabolism in *Escherichia coli*: isolation of null mutants defective in lipoic acid biosynthesis, molecular cloning and characterization of the *E. coli* lip locus, and identification of the lipoylated protein of the glycine cleavage system. *J Bacteriol* **173**, 6411-6420.

Vassylyev, D. G. & Symersky, J. (2007). Crystal structure of pyruvate dehydrogenase phosphatase 1 and its functional implications. *J Mol Biol* **370**, 417-426.

Voet, D. & Voet, J. (2005). *Biochemistry*, 3rd edn: Wiley.

Voth, D. E. & Heinzen, R. A. (2007). Lounging in a lysosome: the intracellular lifestyle of *Coxiella burnetii*. *Cell Microbiol* **9**, 829-840.

Wada, H., Shintani, D. & Ohlrogge, J. (1997). Why do mitochondria synthesize fatty acids? Evidence for involvement in lipoic acid production. *Proc Natl Acad Sci U S A* **94**, 1591-1596.

Wakid, M. H. & Bates, P. A. (2004). Flagellar attachment of *Leishmania* promastigotes to plastic film *in vitro*. *Exp Parasitol* **106**, 173-178.

Wallis, N. G. & Perham, R. N. (1994). Structural dependence of post-translational modification and reductive acetylation of the lipoyl domain of the pyruvate dehydrogenase multienzyme complex. *J Mol Biol* **236**, 209-216.

Weinberg, M. B. & Utter, M. F. (1980). Effect of streptozotocin-induced diabetes mellitus on the turnover of rat liver pyruvate carboxylase and pyruvate dehydrogenase. *Biochem J* **188**, 601-608.

Wexler, I. D., Hemalatha, S. G. & Patel, M. S. (1991). Sequence conservation in the alpha and beta subunits of pyruvate dehydrogenase and its similarity to branched-chain alpha-keto acid dehydrogenase. *FEBS Lett* **282**, 209-213.

White, S. W., Zheng, J., Zhang, Y.-M. & Rock, C. O. (2005). The structural biology of type II fatty acid biosynthesis. *Annual Review of Biochemistry* **74**, 791-831.

Wiemer, E. A., Hannaert, V., van den, I. P. R., Van Roy, J., Opperdoes, F. R. & Michels, P. A. (1995). Molecular analysis of glyceraldehyde-3-phosphate dehydrogenase in Trypanoplasma borelli: an evolutionary scenario of subcellular compartmentation in kinetoplastida. *J Mol Evol* **40**, 443-454.

Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J. & Matthews, B. W. (1992). *Escherichia coli* biotin holoenzyme synthetase/bio repressor crystal structure delineates the biotin- and DNA-binding domains. *Proc Natl Acad Sci U S A* **89**, 9257-9261.

Witkowski, A., Joshi, A. K. & Smith, S. (2007). Coupling of the *de novo* fatty acid biosynthesis and lipoylation pathways in mammalian mitochondria. *J Biol Chem* **282**, 14178-14185.

Wrenger, C. & Muller, S. (2004). The human malaria parasite *Plasmodium falciparum* has distinct organelle-specific lipoylation pathways. *Mol Microbiol* **53**, 103-113.

Wynn, R. M., Chuang, J. L., Cote, C. D. & Chuang, D. T. (2000). Tetrameric assembly and conservation in the ATP-binding domain of rat branched-chain alpha-ketoacid dehydrogenase kinase. *J Biol Chem* **275**, 30512-30519.

Yasuno, R. & Wada, H. (2002). The biosynthetic pathway for lipoic acid is present in plastids and mitochondria in *Arabidopsis thaliana*. *FEBS Lett* **517**, 110-114.

Yi, X. & Maeda, N. (2005). Endogenous production of lipoic acid is essential for mouse development. *Mol Cell Biol* **25**, 8387-8392.

Yi, X., Kim, K., Yuan, W., Xu, L., Kim, H. S., Homeister, J. W., Key, N. S. & Maeda, N. (2009). Mice with heterozygous deficiency of lipoic acid synthase have an increased sensitivity to lipopolysaccharide-induced tissue injury. *J Leukoc Biol* **85**, 146-153.

Zhang, W. J. & Frei, B. (2001). Alpha-lipoic acid inhibits TNF-alpha-induced NF-kappaB activation and adhesion molecule expression in human aortic endothelial cells. *Faseb J* **15**, 2423-2432.

Zhao, X., Miller, J. R., Jiang, Y., Marletta, M. A. & Cronan, J. E. (2003). Assembly of the covalent linkage between lipoic acid and its cognate enzymes. *Chemistry & Biology* **10**, 1293.

Zhao, X., Miller, J. R. & Cronan, J. E. (2005). The reaction of LipB, the octanoyl-[acyl carrier protein]:protein *N*-octanoyltransferase of lipoic acid synthesis, proceeds through an acyl-enzyme intermediate. *Biochemistry* **44**, 16737-16746.