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

The Characterisation of Transketolase from *Leishmania mexicana*

Nicola Jean Veitch

Division of Infection and Immunity

Institute of Biomedical and Life Sciences

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

Faculty of Biomedical and Life Sciences

University of Glasgow

March 2002

Declaration

The results presented in this thesis are my own work, except where there is a statement to the contrary. The subcellular localisation experiments were done in collaboration with Dante Maugeri.

Nicola Veitch

Abstract

Transketolase is a key enzyme of the non-oxidative branch of the pentose phosphate pathway. The identification and characterisation of transketolase in *Leishmania mexicana* will aid in the understanding of the specific interactions of this enzyme and its products within this biochemical pathway. The pentose phosphate pathway may be a potential target for chemotherapy against *Leishmania* as it provides several key intermediates necessary for parasite survival within the host cell.

In this study, the transketolase gene from *L. mexicana* was identified and cloned using a variety of molecular methodologies. The transketolase was shown to be similar to transketolases from other organisms, and contained two conserved motifs, a thiamine-pyrophosphate binding motif and a transketolase motif. Analysis of the transketolase sequence revealed a carboxy-terminal peroxisome-targeting signal (PTS), suggestive of a glycosomal subcellular localisation. Subcellular fractionation of the promastigote *Leishmania* revealed that the protein was predominantly cytosolic, however a glycosomal localisation cannot be ruled out in other life cycle stages. Putative *L. major* and *T. brucei* transketolase protein sequences also contain a PTS at the C-terminal of the protein.

The cloned transketolase gene was expressed as a recombinant protein using a bacterial system with the enzyme containing an N-terminal histidine tag, allowing for efficient purification. Overexpression of the soluble transketolase protein occurred only at a low temperature (15 °C). Characterisation of this recombinant transketolase is described. The specific activity of the *L. mexicana* transketolase was 1.65 U/mg and the K_m for ribose 5-phosphate was 2.75 mM, which is similar to the K_m values for this substrate measured in other species.

Phylogenetic analysis of the *L. mexicana* protein sequence revealed an evolutionary relationship between the *L. mexicana* transketolase and the cyanobacterial and plant transketolases. This observation is consistent with the idea that several genes present in trypanosomatids are related to sequences of plant and cyanobacterial isoforms. The theory of the Kinetoplastida lineage losing a photosynthetic algal endosymbiont subsequent to plastid gene transfer into the nucleus of the parasite is discussed.

Targeted gene replacement experiments were initiated to investigate the role transketolase plays within the *Leishmania* parasite. Constructs were prepared to replace transketolase with an antibiotic resistance marker via homologous recombination. The construction of the vectors and initial experiments are described.

Acknowledgements

I would first like to thank my supervisor, Dr. Mike Barrett for the support and guidance he has given me over the course of my studies. Mike's encouragement over the past three years has been invaluable to me.

Thank you to all my friends in the lab, past and present, who have been brilliant to work with over the past three years. I would especially like to thank Sanya Sanderson whom has been an inspiration to me as a scientist and as a friend.

Thank you, to all those in the lab in Buenos Aires who made me feel welcome and made my trip a memorable and enjoyable experience.

I have received valuable help and advice during this project from Kirsten Flemming, Hubert Denise and Dante Maugeri, to whom I am grateful.

Thank you to all my friends who were always available for a pint and a chat just when I needed them, particularly Emma, Graham, Caroline, Andy, Jane, Angus, Graham and David.

My parents and my sister have provided me with endless support and love, and have always let me know that my best was always enough, thank you.

My greatest debt is to Mark, who has been there for me every step of the way and has shown me the humour in every situation. Thank you for keeping me sane for the past three years, and for your confidence in me.

Table of contents

List of Figures

Chapter III. Cloning of transketolase genes from *Escherichia coli* and *Leishmania mexicana*

Chapter V. Localisation and phylogenetic analysis of a transketolase protein in *Leishmania mexicana*

List of Tables

List of Abbreviations

ADP	Adenosine diphosphate
ATP	Adenosine triphosphate
AU	Absorbance Units
bp	Base pair(s)
cDNA	Complementary DNA
CO_2	Carbon dioxide
ddd	Double distilled de-ionised
DMSO	Dimethylsulphoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EC	Enzyme Comission
EC_{50}	half effective concentration
EDTA	Ethylenediamine tetraacetic acid
EU	European Union
FCS	Foetal calf serum
HIV	Human Immuno-Deficiency Virus
kDa	Kilo Daltons
LB	Luria-Bertani Medium
Mb	Megabase
µg	Microgram(s)
mg	Milligram(s)
µl	Microlitre(s)
ml	Millilitre(s)
mm	Millimetre(s)

mRNA	Messenger RNA
NAD	Nicotinamide Adenine Dinucleotide (oxidised form)
NADH	Nicotinamide Adenine Dinucleotide (reduced form)
NADP	Nicotinamide Adenine Dinucleotide Phosphate (oxidised form)
NADPH	Nicotinamide Adenine Dinucleotide Phosphate (reduced form)
NMR	Nuclear Magnetic Resonance
ORF	Open Reading Frame
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
RNA	Ribonucleic acid
rpm	Rotations Per Minute
SBRI	Seattle Biomedical Research Institute
SDM	Semi-Defined Medium
SDS	Sodium Dodecyl Sulphate
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
USP	University of San Paulo
UV	Ultraviolet
WHO	World Health Organisation
WWW	World Wide Web

Chapter I

Introduction

1.1. *Leishmania*

Leishmania is a protozoon parasite from the family Trypanosomatidae, of the order Kinetoplastida. Trypanosomatids are exclusively flagellated obligate parasites and *Leishmania* is specifically intracellular, with a digenetic life cycle in insects and vertebrates.

The orders Kinetoplastida along with the order Euglenida comprise the phylum Euglenozoa. Members of the Euglenida possess chloroplasts, whereas the Kinetoplastids do not (Figure 1.1). It is of interest to determine whether the Euglenida acquired a chloroplast after the division from the Kinetoplastida or if a common ancestor possessed a chloroplast that was secondarily lost by Kinetoplastida. The Kinetoplastida are defined by the kinetoplast, a network of DNA that lies within the single mitochondrion (Clarke and Wallace, 1960). This order is classified into two families, the Trypanosomatidae and the Bodonidae, the later representing free-living organisms. The family Trypansomatidae parasitise various organisms including insects, plants, fish and animals.

1.2. Leishmaniasis

Leishmania parasites cause the globally prevalent disease leishmaniasis. It is estimated that 12 million people suffer from leishmaniasis, with 350 million at risk of contracting the disease (WHO, 2001). Leishmaniasis has a spectrum of clinical manifestations that are caused by a variety of different species of *Leishmania* (reviewed

in a Report of a WHO Expert Committee, 1984). The disease can be divided into four different forms:

Cutaneous leishmaniasis (CL) can cause skin ulcers at the site of the sandfly bite, usually on an exposed region of the skin. This can leave the patient scarred and disabled.

Diffuse cutaneous leishmaniasis (DCL) is characterised by nodules spread over the skin, which do not heal spontaneously and are not easily treated.

Mucocutaneous leishmaniasis (MCL), also known as espundia, can produce lesions that may lead to the erosion of the mucous membranes of the nose and the palate. This disease can be extremely disfiguring and untreated patients may die of secondary infections to the open wound.

Visceral leishmaniasis (VL) can cause swelling of the liver and the spleen, as the parasite invades these organs and the bone marrow. Irregular fever, weight loss and anaemia can also occur. A darkening of the skin can develop, most commonly in the Indian subcontinent, lending the name kala-azar (black sickness). If VL is not treated it will result in death.

The four different forms of leishmaniasis are caused by several different species of *Leishmania* (Table 1.1). *Leishmania mexicana*, the organism studied during this project, is responsible for cutaneous leishmaniasis in the New World, more specifically Mexico and Northern Argentina. This parasite causes 'chiclero ulcer', that often heals spontaneously. However chronic progressive lesions can develop, most commonly around the ear. Cutaneous leishmaniasis is the most common form of all the leishmaniases and represents 50-75 % of new cases (WHO, 2001). The global distribution of cutaneous leishmaniasis is shown in Figure 1.2. 90 % of cutaneous leishmaniasis occurs in Afghanistan, Brazil, Iran, Peru, Saudi Arabia and Syria.

Form of Leishmaniasis	**Causative agent**	**Distribution (Old world/New world)**
Cutaneous	*L. major, L. tropica, L. aethiopica*	Old
	L. mexicana, L. guyanensis, L. panamensis	New
Diffuse cutaneous	*L. aethiopica*	Old
	L. amazonensis	New
Mucocutaneous	*L. brasiliensis, L. panamensis*	New
Visceral	*L. infantum, L. donovani*	Old
	L. chagasi	New

Table 1.1. The causative species of the four forms of leishmaniasis and their distribution.

The emergence of *Leishmania*/HIV co-infection has been rapid and a cause for great concern in certain geographical areas such as south-western Europe and Brazil (Fenech, 1997). Individuals with HIV become immunosuppressed and leishmaniasis can quickly develop in endemic areas.

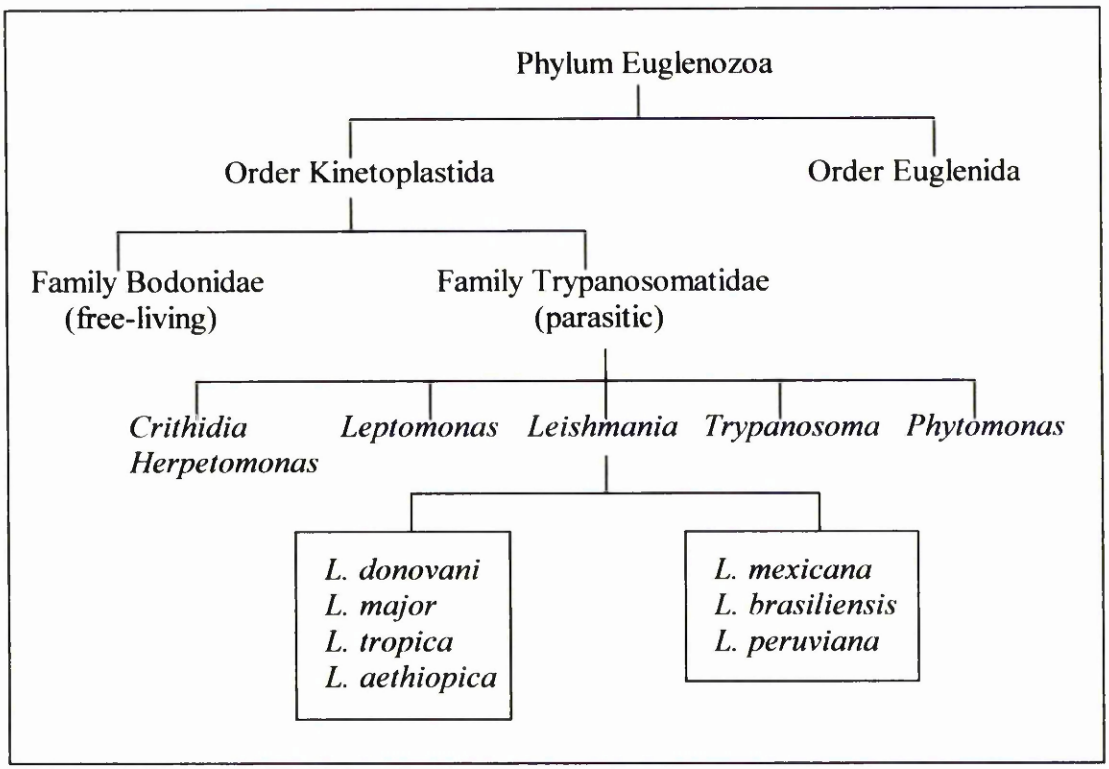


Figure 1.1. Taxonomy of the *Leishmania* species

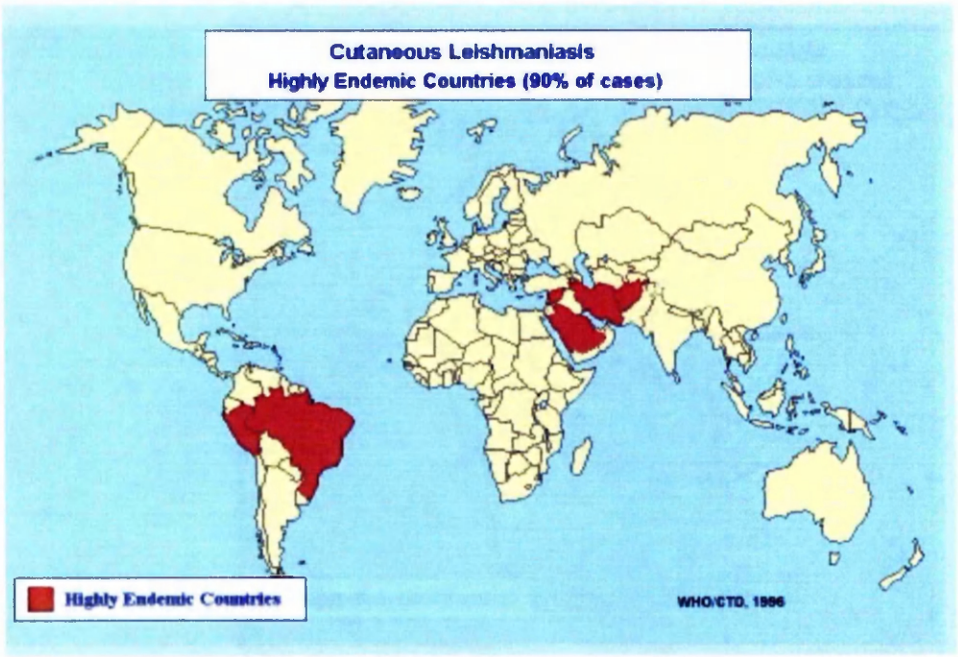


Figure 1.2. Global distribution of cutaneous leishmaniasis. Copied from WHO web site, 2002.

1.3. Life cycle

The *Leishmania* parasite exists in two forms: the promastigote form that survives in the insect vector, the sandfly, and the amastigote form that occurs in the macrophage of a mammalian host. The infected sandfly transmits the parasite by biting the host and the parasite is then injected into the skin, entering macrophage cells at the site of infection. The cycle is continued when a sandfly takes a bloodmeal from an infected host, contracting the *Leishmania* parasite (Figure 1.3).

Transmission occurs through the female sandfly from the genus *Phlebotomus* and *Lutzomyia* (in the Old World and the New World, respectively). The amastigotes are initially present in the mid-gut of the sandfly within the peritrophic membrane that contains the bloodmeal (reviewed by Schlein, 1993). The parasites transform into multiplicative promastigotes that migrate to the thoracic mid-gut and pharynx where they transform into non-dividing infective promastigotes. The parasites are regurgitated during the uptake of a bloodmeal, allowing the parasites to be injected into the host (Schlein *et al.*, 1992).

Once inside the host, macrophage cells at the site of infection phagocytose the metacyclic promastigotes. The parasite enters the parasitophorous vacuole of the macrophage, and subsequently fusion of the endosome and lysosomes creates a phagolysosome (Antoine *et al.*, 1998). It is within this vacuole that the promastigotes transform into multiplicative, nonmotile amastigotes. The macrophage eventually ruptures, releasing amastigotes that will infect other macrophage cells. *L. donovani* amastigotes can infect cells of the reticuloendothelial system present in the blood, spleen, lymph and bone marrow. This can lead to visceral leishmaniasis.

Leishmaniasis is a zoonosis, infecting forest rodents, marsupials, foxes, dogs, sloths and guinea pigs (reviewed by Lainson, 1983). These animals act as reservoirs for the parasite, with humans acquiring the disease when coming into contact with a forest environment. Deforestation is a major contributory factor to the increasing prevalence of leishmaniasis in Central and South America (Walsh *et al.*, 1993). Occupation of cleared areas results in people living in close contact with the forest, and concurrently contracting leishmaniasis and other tropical diseases. This disease is intrinsically linked to poverty in developing countries, where immigration, limited irrigation and exploitation of the land all contribute to its prevalence.

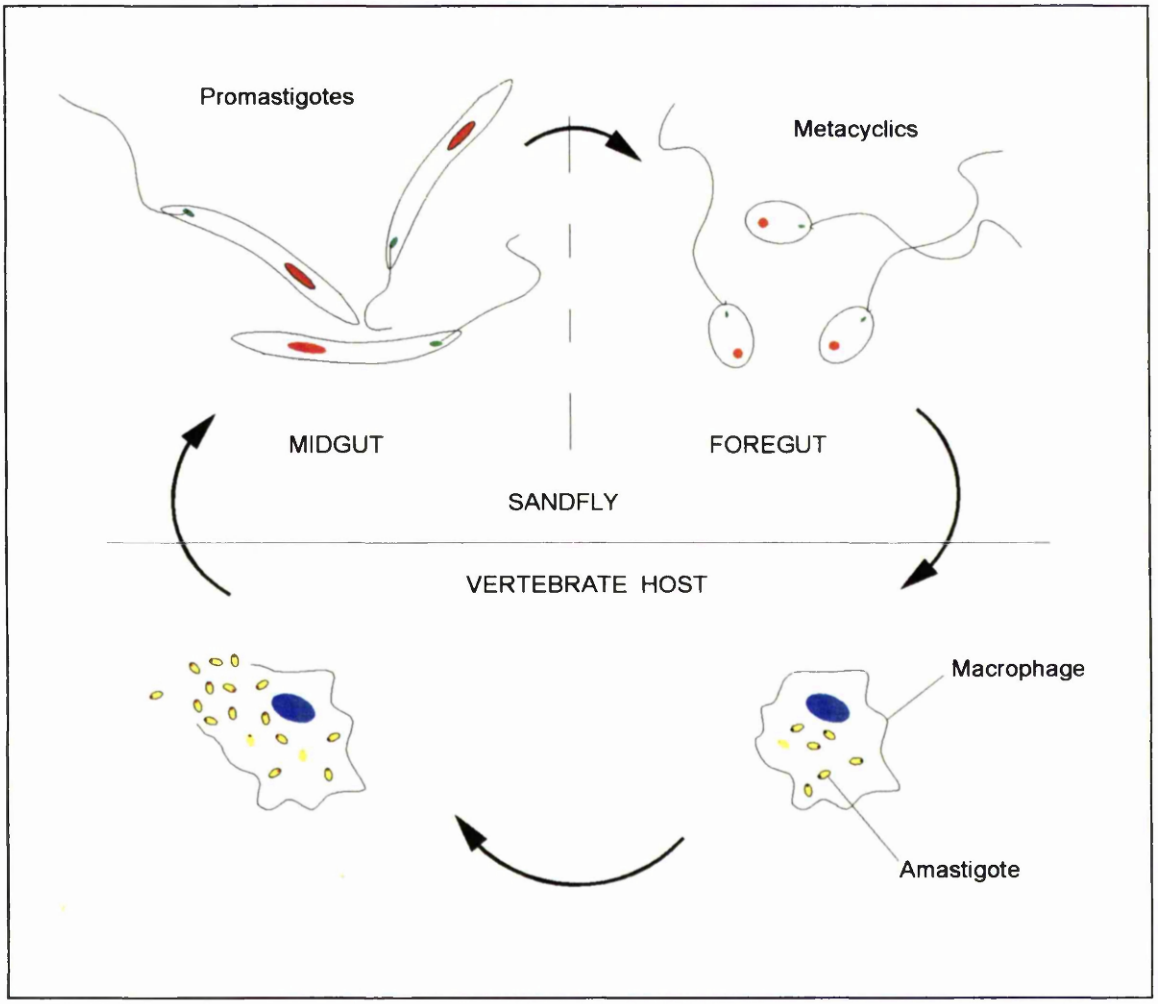


Figure 1.3. Life cycle of *Leishmania*

1.4. Virulence factors

The *Leishmania* parasite has evolved to survive in the mammalian host before and after the invasion of a macrophage. Upon entry to the host, the *Leishmania* promastigotes become opsonised by complement C3b, causing the parasite to be taken up by macrophage cells via receptors present on the macrophage cell surface (Brittingham and Mosser, 1996). Lipophosphoglycan (LPG) and the major surface protein gp63 are both highly abundant molecules present on the surface of the promastigote and are thought to be involved in the activation of complement. On the surface of infective metacyclic promastigotes, LPG is elongated, and is thought to act as a physical barrier to complement mediated lysis by preventing the attachment of the membrane attack complex to the parasite surface (Puentes *et al.*, 1990). However, it must be noted that recent experiments have shown that LPG was not essential for invasion and survival of *L. mexicana* in mice (Ilg, 2000). This may reflect virulence factors that are species specific. Folate and biopterin metabolic pathways have been shown to be important in maintaining a viable parasite (Cruz *et al.*, 1991; Cunningham *et al.*, 2001). Biopterin levels in *L. major* were shown to be important for the virulence of the parasites. An *L. major* strain devoid of the gene encoding for pteridine reductase has low levels of tetrahydrobiopterin and showed an increase in metacyclogenesis, which led to a corresponding increase in parasite burden and lesion size in a mammalian host (Cunningham *et al.*, 2001). This suggests that biopterin is involved in controlling the rate of metacyclogenesis and therefore is involved in the control of the virulence of the parasite.

For *Leishmania* to survive within the phagolysosome of the host it must possess virulence factors to evade reactive oxygen intermediates, nitric oxide and lysosomal enzymes, all of which are present in the host cell. LPG, cysteine proteinases and gp63

are molecules thought to be involved in the protection of the *Leishmania* parasite within the phagolysosome (reviewed by Matlashewski, 2001). Cysteine proteinases are known to be active in amastigotes of *Leishmania* species (Galvao-Quintao *et al.*, 1990). An *L. mexicana* strain lacking the cysteine proteinase B (CPB) gene array was shown to have decreased infectivity in macrophages and only produced small lesions (Mottram *et al.*, 1996). This observation suggests that CPB's are important virulence factors for *Leishmania* to establish a successful infection. Attenuated *L. donovani* promastigotes exhibiting a decrease in parasite burden in hamsters and mice, in comparison to infective parasites, also show a marked decrease in the amount of gp63 exposed on the cell surface of the parasite (Wilson *et al.*, 1989). This study suggests that gp63 expression on the parasite cell surface is important for the adhesion of promastigotes to macrophages and intracellular amastigotes survival.

1.5. Chemotherapy

The drugs of choice to treat all forms of leishmaniasis are pentavalent antimonials. Pentavalent antimony comes in two forms: sodium stibogluconate (Pentostam) and meglumine antimoniate (Glucantime). These drugs are far from exemplary as they are difficult to administer and are expensive. Antimonial resistance does occur. Moreover the drugs are toxic, causing a plethora of side effects including fever, nausea, anorexia, headache and cough (reviewed by Arana *et al.*, 2001). Cardiac toxicity can occur if a high dose of antimonial is administrated, causing fatal arrhythmias. Pentavalent antimonials have been shown to inhibit glycolytic enzymes and fatty acid oxidation in *Leishmania* amastigotes (Berman, 1988), however antimonials are heavy metals and may have other modes of action. The kidney rapidly excretes antimonials and a prolonged duration of treatment is usually required. This may be interrupted by common side effects resulting in discontinuous drug exposure

(Olliaro and Bryceson, 1993). Resistance to antimonials is thought to have arisen for this reason and because of sub-optimal drug regimes.

Pentamidine is a second line drug, used only in cases of antimony resistance. This drug is also used to treat sleeping sickness (Human African Trypanosomiasis) and *Pneumocystis carinii* pneumonia infections. Pentamidine is highly toxic and is excreted slowly by the liver and kidney. Amphotericin B has also been used to treat antimony resistant leishmaniasis, but is expensive and potentially very toxic, causing anaphylaxis, convulsions, fever and anaemia. However, this drug was used at a low dosage over a short period of time and achieved 100 % cure rate with little toxicity in India to treat antimony resistant kala-azar (Mishra *et al.*, 1992).

New approaches and chemotherapy are required to aid in the treatment of leishmaniasis for a number of reasons including the increasing number of travellers entering endemic areas, the emergence of Leishmania/HIV co-infection and parasite resistance to pentavalent antimonials. A new drug candidate currently under clinical development for visceral and cutaneous leishmaniasis is miltefosine, an anticancer drug that can be given orally. Clinical trials in India and Columbia have shown hopeful initial results, with 96 % and 94 % cure rate, respectively (Fischer *et al.*, 2001). The main side effects were gastrointestinal, such as vomiting and diarrhoea. The development of this drug has been a collaborative effort by WHO/TDR (special programme for research and training in tropical diseases) and Zentaris/ASTA Medica.

1.6. Parasite genome initiatives

The sequencing of several parasite genomes is underway. Parasite genomes being sequenced include including *Plasmodium* species, *Leishmania major*, *Trypanosoma brucei*, *Trypanosoma cruzi*, *Schistosoma mansoni* and *Brugia malayi*.

The substantial amount of genome sequencing data being produced from this massive effort is revolutionising the study of molecular biology and producing a new discipline of bioinformatics. Rational drug design is based on the identification and characterisation of novel drug targets in parasites and this method for identifying new antiparasitic drugs will be greatly accelerated by the availability of the gene sequences. The availability of the sequences of whole parasite genomes will yield information regarding novel genes, the products of which are parasite specific and have not been previously recognised. The sequencing of the first draft of the human genome was completed in 2001 (Lander *et al.*, 2001) and will therefore allow a comparison to be made between this and parasite genomes.

1.6.1. The *Leishmania* genome project

The *Leishmania* Genome Network (LGN) was set up by a collaborative group to sequence the *Leishmania major* Friedlin genome in 1994. The laboratories involved in the sequencing are primarily at Seattle Biomedical Research Institute and the Sanger Institute, with groups in Brazil and Europe also participating. This project is currently funded by Beowulf Genomics, the European Commission, the WHO/TDR (Special Programme for Research and Training in Tropical Diseases), the Burroughs Wellcome Fund and the U.S. National Institutes of Health. A number of approaches have been taken to sequence the ~34 Mb genome that is divided between 36 chromosomes. One approach has been to make chromosomal libraries, initially separating the chromosomes by pulse-field gel electrophoresis, shearing the DNA into 1-2 kb fragments and cloning into vectors that are then sequenced using vector specific primers. A shuttle cosmid vector library of *L. major* genomic DNA was created using the cLHYG vector (Ryan *et al.*, 1993). The individual clones have been assembled into overlapping contigs and

assigned chromosome map location (Ivens *et al.*, 1998) and are currently being shotgun sequenced. The sequencing data are assembled into contigs and are available to search via the LGN web site.

Approximately 25 % of the sequencing of the *L. major* genome has been completed (LGN web site, updated on 15th January, 2002). Chromosomes 1, 3 and 4 have recently been completed and the whole genome is expected to be sequenced by 2003. A summary of the current status of the genome project is shown in Table 1.2

Chromosome	Lab	Sequencing Complete	Sequencing in progress
1, 3, 7, 10, 27, 35	SBRI	1, 3	7, 10, 27, 35
4, 12, 15, 16, 17, 18, 20, 22, 24, 25, 26, 28, 31, 32, 33, 34, 36	Sanger Institute	4	12, 15, 16, 17, 18, 20, 22, 24, 25, 26, 28, 31, 32, 33, 34, 36
5,13,14,19,21,23	EULeish consortium	-	5,13,14, 19, 21, 23
2	USP, Brazil	-	2
6	Fiocruz, Brazil	-	6

Table 1.2. The status of the *L. major* genome sequencing project. The information shown was obtained from the *Leishmania* genome network web site, updated on the 15th January 2002.

Approximately 75 % of the genes identified by homology searching so far, have unknown function (Myler *et al.*, 2001). The rapid advancement of the parasite genome sequencing initiatives will speed up the initial step of rational drug design of identification of potential target proteins.

1.7. Gene organisation and expression

Trypanosomatids have a unique mode of gene organisation in comparison to higher eukaryotes. The genome is organised into polycistronic units of protein-coding genes with intergenic untranslated regions. No introns have been found in the *Leishmania* genome and transcription of individual genes occurs in conjunction with a

mechanism known as *trans*-splicing (detailed below). The sequencing of chromosomes 1, 3, and 4 have shown that one protein-coding gene exists for approximately every 3.26 kb of DNA (Myler *et al.*, 1999; Myler *et al.*, 2001). Chromosome one was the first chromosome to be fully sequenced and displays two 'head to head' long polycistronic units of tightly packed genes, one strand encoding 29 genes and the opposite strand encoding 50 genes (Myler *et al.*, 1999).

The *Leishmania* genome contains repetitive DNA sequences, including telomeric sequences and microsatellite DNA (Stiles *et al.*, 1999). Telomeric and sub-telomeric repetitive elements account for as much as 15 % of chromosome one of *L. major* (Myler *et al.*, 1999). Genes are present in Trypanosomatids as single-copy genes, paired genes and tandemly repeated, multi-copy genes. Multi-copy genes such as α- and β-tubulins and surface antigens including PARP and variant surface glycoproteins (VSG) in *T. brucei* are organised as direct tandem repeats (Stiles *et al.*, 1999).

Trypanosomatid protein-coding genes are usually present in the genome as polycistronic units that are transcribed as individual mRNAs (reviewed by Graham, 1995). The processing of the monocistronic genes is by 5' *trans*-splicing and 3' polyadenylation of mRNA transcripts. The initial cleavage of the pre-mRNA occurs when a capped 39 nucleotide spliced leader (SL) attaches to the intergenic region 5' of the gene (Agabian, 1990). Trypanosomatids possess a *trans*-splice acceptor site, characterised by the nucleotides AG, and an upstream polypyrimidine tract in the intergenic regions present between genes. This AG spliced leader acceptor site of mRNA is similar to the AG dinucleotide found in mammalian *cis*-splicing (Huang and Van der Ploeg, 1991). Polyadenylation is thought to occur within 1 kb upstream of the SL acceptor site. Recent experiments have suggested that *trans*-splicing and polyadenylation may be coupled, with the pyrimidine repeats and the SL acceptor site

playing an important role. Initial experiments altering the position of the splice acceptor site downstream of the dihydrofolate reductase-thymidylate synthase gene (DHFR-TS) in *L. major*, showed that the polyadenylation site moved 400-500 nucleotides upstream of the new acceptor (LeBowitz *et al.*, 1993). It was suggested that the positioning of the downstream splice acceptor site specifies the polyadenylation site and therefore the two processes must be linked. The procyclic acidic repetitive protein (PARP) gene A locus in *T. brucei* has also been examined for transcriptional processing signals present in the intergenic regions. A number of constructs have been made deleting blocks of sequence from the intergenic regions of the PARP gene from *T. brucei* (Schürch *et al.*, 1994; Hug *et al.*, 1994). In both studies, a polypyrimidine tract upstream of a splice acceptor site was shown to be crucial for accurate polyadenylation of the procyclin transcript. Similar experiments were done using the *T. brucei* α- and β-tubulin gene locus, where block deletions of the intergenic regions between the two genes identified that polypyrimidine tracts were essential for accurate trans-splicing of the downstream transcript and polyadenylation of the upstream transcript (Matthews *et al.*, 1994).

The regulation of gene expression is thought to be mainly post-transcriptional as genes on the same polycistronic unit are co-transcribed and cannot be individually regulated. Certain genes of trypanosomatids are developmentally regulated, with differential expression of mRNA transcripts during the life cycle. This is known to occur in VSG genes, PARP and glycolytic enzymes (Vanhamme and Pays, 1995; Graham, 1995).

1.8. Molecular biology as a tool

Molecular biology technologies can be used in combination with the vast amount of parasite genome sequencing data to investigate specific gene function and to characterise the analogous proteins. Identification of à gene's open reading frame (ORF) can lead to the isolation and cloning of genes of interest. Expression of recombinant proteins using expression vectors can lead to the rapid purification of proteins on a large scale. Gene replacement studies have been used as one approach to investigate and validate a specific gene (and gene-product) as a potential chemotherapeutic target in a parasitic organism, although it may be naïve to think that a gene essential to the parasite may automatically be a good drug target (Barrett *et al.*, 1999). Rational drug design is taking advantage of these new technologies in an attempt to make new viable antiparasitic drugs (Gutteridge, 1997). This project utilises a number of molecular biology approaches to begin to characterise transketolase in *L. mexicana*.

1.8.1. Gene replacement studies

Gene knockout experiments can be used to investigate the biochemical and molecular function of a single gene. This approach can yield interesting information regarding the mechanisms of parasite metabolic pathways, and potentially identify if a gene is essential to the parasite. A gene can be replaced by homologous recombination, a process that naturally occurs in trypanosomatids (Tobin *et al.*, 1991; Cruz *et al.*, 1991; Coburn *et al.*, 1991). To replace a gene, a DNA construct is made containing an antibiotic resistant gene flanked by the 5' and 3' intergenic sequences belonging to the gene to be deleted. This construct is transfected into the parasite, and antibiotic resistant

cells can be selected for in culture or on solid medium containing the selectable marker (LeBowitz, 1994). Individual clones can be characterised for genotype to ensure correct gene targeting and subsequently the phenotype can be examined.

Several selectable markers have been developed, as sequential gene disruption is essential to create null mutant parasites, as the *Leishmania* organisms are diploid. Hygromycin B phosphotransferase (*HYG*), neomycin phosphotransferase (*NEO*) and streptothricin acetyltransferase (*SAT*) can all confer resistance to stably transfected parasites (Cruz *et al.*, 1991; Laban *et al.*, 1990; Joshi *et al.*, 1995; Freedman and Beverley, 1993).

Vectors designed to knockout the DHFR-TS gene in *L. major* were further developed to create shuttle vectors (pX vectors) that can express exogenous genes and remain episomal (Ryan *et al.*, 1993; Kelly *et al.*, 1992). Vectors carrying reporter genes have been successfully transfected can used to examine the molecular machinery of the parasites. Chloroamphenicol acetyltransferase (CAT) can be detected by radiolabel and has been used to show the occurrence of homologous recombination in *L. enriettii* (Tobin *et al.*, 1991).

Folate metabolism was the first target area to be examined by gene-replacement studies. DHFR-TS null mutants in *L. major* were shown to grow only in the presence of thymidine, suggesting that this gene is essential to the cell (Cruz *et al.*, 1991). The cysteine proteinases (cp) from *L. mexicana* have been targeted by gene disruption creating *cpa*, *cpb* and *cpc* null mutants. *L. mexicana cpa* and *cpc* are single copy genes and sequential gene disruption showed that they are not essential for infectivity or virulence in the parasite (Souza *et al.*, 1994; Bart *et al.*, 1997). *Lmcpc* null mutants did exhibit reduced infectivity *in vitro*, however lesions were still able to form in mice (Bart *et al.*, 1997). *Lmcpb* are encoded on a tandem array of 19 genes that were successfully

replaced to create an *Lmcpb* null mutant (Mottram *et al.*, 1996). The *Lmcpb* knockout cells exhibited a marked difference in the infectivity of the parasite, with only 20 % of the host cells becoming infected *in vitro* in comparison to a wild-type population. The null mutants of the *cpb* array showed no difference in growth or differentiation of the parasite, and lesions were still able to grow *in vitro*. Using gene-targeting experiments, the cysteine proteinases have been shown not to be essential for *Leishmania* to survive and cause lesions within a mammalian host. However, these studies have shown that members of this gene family are important in the virulence and infectivity of the parasite.

1.9. Carbohydrate metabolism in trypanosomatids

Trypanosomatids catabolise glucose as a major energy source via the glycolytic pathway. Within the insect stage of the life cycle, amino acids, specifically proline, are used to fuel the parasites. The degradation of glucose occurs mainly in glycosomes, peroxisome-like organelles unique to Kinetoplastida (Visser *et al.*, 1981). Carbohydrate metabolism varies between the trypanosomatids, with the mitochondrion playing a role in only certain life cycle stages of the parasites, utilising end products from glycolysis and amino acid metabolism (Tielens and Van Hellemond, 1998). Trypanosomatids exhibit high rates of glucose consumption, with bloodstream *T. brucei* exhibiting a 10-fold increase in comparison to mammalian cells (Cazzulo, 1992). Carbohydrate metabolism can be thought of as a potential drug-target in trypanosomatids as it is unique in comparison to the host for several reasons including its unusual compartmentalisation (reviewed by Opperdoes and Michels, 2001).

1.9.1. *Trypanosoma brucei*

T. brucei has a biochemical physiology that is adapted to the varying environments to which it is exposed during the life cycle (Opperdoes, 1987). The long-slender trypanosomes present in the bloodstream of the mammalian host are extra-cellular, with a depleted mitochondrion and an inactive Krebs cycle with no respiratory chain or cytochromes. Pyruvate is the main end product and is excreted into the cytosol as the last reaction of glycolysis. ATP production and consumption within the glycosome is balanced, with 2 moles of ATP being produced per molecule of glucose in both life cycle stages of the parasite. The NADH produced during glycolysis is reoxidised via a mitochondrial glycerol-3-phosphate oxidase, thus maintaining redox balance within the glycosome. Transformation into the infective short-stumpy form is linked with the development of the mitochondrion and an increase in Krebs cycle activity. The procyclic insect stage of *T. brucei* has a fully functional mitochondrion and an active Krebs cycle. Pyruvate is oxidised to acetyl-CoA and is further degraded to CO_2 by the Krebs cycle. Acetyl-CoA is also converted to acetate by an acetate:succinate CoA transferase and is recycled to succinate producing ATP from ADP using a mitochondrial succinyl CoA synthetase (Van Hellemond *et al.*, 1998). Procyclic *T. brucei* exhibits a classical cytochrome-containing respiratory chain that is coupled to oxidative phosphorylation. The insect stage utilises proline as a primary energy source as it is abundant in the mid-gut of the sandfly vector.

1.9.2. *Leishmania*

Carbohydrate metabolism in *Leishmania* promastigotes and amastigotes occurs in the glycosome (Coombs *et al.*, 1982; Hart and Opperdoes, 1984), with end products further catabolised by the Krebs cycle in the mitochondrion (Figure 1.4). A functional

electron transport chain is present in *L. mexicana* with oxygen being the final electron acceptor (Hart and Coombs, 1982). This is vital for the re-oxidation of NADH generated during glycolysis. Pyruvate, the end product of glycolysis, enters the Krebs cycle in the mitochondrion, producing succinate and CO_2. Acetate is also an end product that is made via acetyl CoA. The major end products of aerobic carbohydrate metabolism are CO_2, succinate, acetate, pyruvate, alanine and lactate (Hart and Coombs, 1982; Cazzulo, 1992; Blum, 1993). Intracellular pools of alanine in *Leishmania* promastigotes are released under hypo-osmotic stress, an adaptation thought to be due to the changing habitat occurring in the sandfly (Blum, 1993). Glucose is thought to be the primary source of free energy in both stages of the *L. mexicana* parasite, with proline being catabolised to CO_2 at a low rate (Hart and Coombs, 1982). Both life cycle stages of *L. mexicana* do exhibit similar energy metabolism, although amastigotes do catabolise fatty acids up to 10-fold more than promastigotes, suggesting that it is an available substrate within the macrophage host cells (Hart and Coombs, 1982).

L. mexicana and *L. infantum* promastigotes do not have a facultative anaerobic energy metabolism, as in a decreased oxygen environment, the parasites do not divide and their motility is debilitated (Hart and Coombs, 1981; Van Hellemond *et al.*, 1997). However, *Leishmania* promastigotes can reversibly lower their metabolic rate under anaerobic conditions by inhibition of the respiratory chain (Van Hellemond and Tielens, 1997). This arrest in metabolism would allow the survival of *Leishmania* promastigotes within the sandfly midgut.

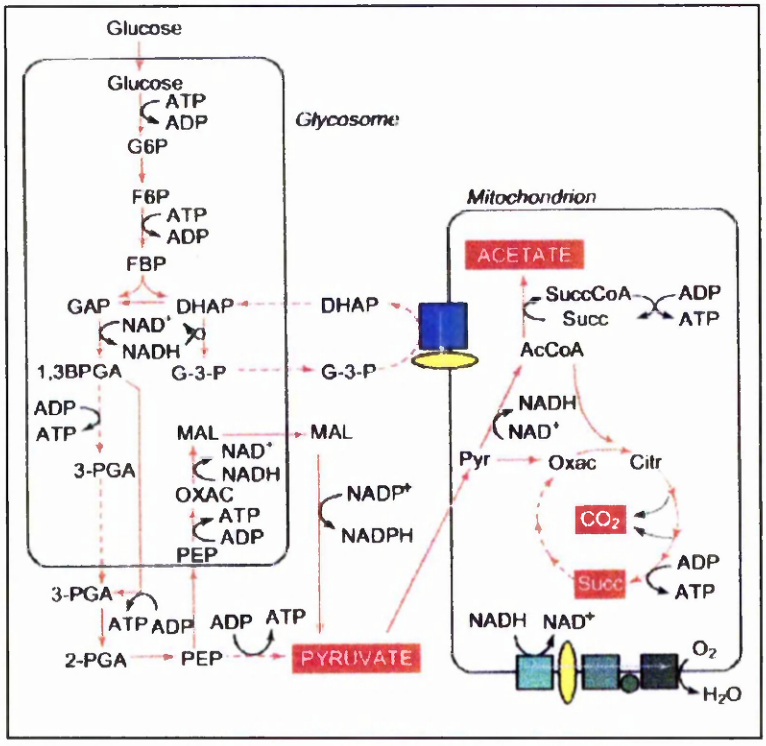


Figure 1.4: Schematic representation of *Leishmania* promastigote glucose metabolism. This diagram is taken from Tielens and Van Hellemond review, 1998. The end products of glycolysis are shown in red boxes and dashed lines show minor pathways. The electron transport complexes are in blue and green boxes and the ubiquinone/ubiquinol pool is in yellow elipses. Abbreviations: AcCoA, acetyl-CoA; Citr, citrate; DHAP, dihydroxyacetone phosphate; FBP, fructose 2,6-bisphosphate; F6P, fructose 6-phosphate; GAP, glyceraldehyde 3-phosphate; G-3-P, glycerol 3-phosphate; G6P, glucose 6-phosphate; MAL, malate; Oxac, oxaloacetate; PEP, phosphoenolpyruvate; Pyr, pyruvate; Succ, succinate; SuccCoA, succinyl-CoA, 1,3BPGA, 1,3-bisphosphoglycerate; 2-PGA, 2-phosphoglycerate; 3-PGA, 3-phosphoglycerate.

1.9.3. Regulation of glycolysis

Regulation of glycolysis in trypanosomatids is not similar to its mammalian counterpart. Hexokinase and phosphofructokinase do not regulate glycolysis in *T. brucei* (Opperdoes, 1987). It has been suggested that glycolysis may be regulated by the transport of glucose over the plasma membrane, or the subcellular compartmentation in the glycosome may be regulatory (Opperdoes, 1987; Cazzulo, 1992; Bakker *et al.*, 2000). Regulation of pyruvate kinase by fructose 2,6-bisphosphate has been seen to occur and is thought to be unique to trypanosomatids.

1.9.4. Glycosomes

Glycolytic enzymes are mostly sequestered in organelles unique to the order Kinetoplastida, named glycosomes (Opperdoes and Borst, 1977). The occurrence of glycosomes in several *Leishmania* species has been shown by cell fractionation and glycolytic enzyme latency (Hart and Opperdoes, 1984). Some glycolytic enzymes do exist in the cytosol, although it has been shown that glycolysis mainly takes place in the glycosome (Visser *et al.*, 1981).

Glycosomes are organelles thought to belong to the peroxisome family. Mammalian peroxisomes are involved in the production or degradation of hydrogen peroxide and harbour the enzyme catalase. Glycosomes in *T. brucei* and *Leishmania* do not contain this enzyme and are primarily involved in glycolysis (Borst, 1989). Both peroxisomes and glycosomes have a single-phospholipid bilayer membrane and contain no DNA. The glycosome's main function is glucose metabolism, however other metabolic pathways are known to be present that are typical in peroxisomes (reviewed by Michels *et al.*, 2000). Enzymes involved in ether lipid synthesis and β-oxidation of fatty acids have been shown to be present in *Leishmania* species (Hart and Opperdoes, 1984; Opperdoes, 1984). Subsequently, the initial steps of the ether lipid biosynthetic pathway have been shown to be associated with glycosomes (Heise and Opperdoes, 1997; Zomer *et al.*, 1999). Some of the enzymes present in purine metabolism pathways and *de novo* pyrimidine biosynthesis have been shown to be localised inside the glycosome (Hassan *et al.*, 1985; Hammond *et al.*, 1981). Certain enzymes of the pentose phosphate pathway in *L. mexicana*, *T. brucei* and *T. cruzi* are thought to be partially present in the glycosome (Duffieux *et al.*, 2000; Cazzulo, personal communication, 2000).

1.9.4.1. Peroxisome targeting signals

Glycosomal proteins are synthesised on free cytoplasmic ribosomes and are imported post-translationally into the organelle. Topogenic signals have been identified that allow such proteins into both peroxisomes and glycosomes. Three peroxisome targeting signals (PTS) are thought to exist: PTS-1, a C-terminal tripeptide based on the residues serine-lysine-leucine (SKL), PTS-2, an N-terminal extension with a loosely conserved motif and PTS-3, short internal peptides that will be on the surface of the fully-folded protein and available to bind to receptors on the glycosome. The firefly luciferase gene catalyses a light reaction in the lantern organelle, and has been used as a reporter gene to examine a PTS in mammals, plants, yeasts and parasites. It was shown that the C-terminal tripeptide SKL or those with similar physico-chemical characteristics were necessary for the import of the luciferase protein into the peroxisome (Gould *et al.*, 1989). Subsequently Wang and co-workers stably transfected trypanosomes with the luciferase gene, where it was expressed and accumulated within the glycosomes (Sommer *et al.*, 1992). This study also mutated the C-terminal PTS and showed that other residues were also efficient for importation (Table 1.3). A similar study using the reporter gene chloroamphenicol acetyltransferase (CAT) fused to 20 amino acids of the C-terminal sequence from the glycosomal enzyme phosphoglycerate kinase-C (Blattner *et al.*, 1992) suggested similar findings. However, amino acids upstream to the C-terminal PTS were important for efficient protein import into the glycosome. The dihydroxyacetone synthase protein (DHAS) is a transketolase of the methanol-utilising yeast *Hansenula polymorpha* and resides in the peroxisome of this organism. A C-terminal PTS was demonstrated to be necessary for the importation of this protein into the peroxisome (Hansen *et al.*, 1992).

Amino acid residue	Redundancy	Property of amino acid
Serine (S)	A/C/G/H/N/P/T	small neutral
Lysine (K)	H/M/N/Q/R/S	hydrogen bonding
Leucine (L)	I/M/Y	hydrophobic

Table 1.3: C-terminal PTS-1 amino acid residues able to import proteins into the glycosome. The individual residues SKL were replaced by other amino acids shown in the table with similar physico-chemical properties and were all able to import proteins into the glycosomes (Sommer *et al.*, 1992).

An N-terminal PTS-2 was shown to exist on the fructose bisphosphate aldolase in *T. brucei*, a glycosomal enzyme present in the glycosome but lacking a PTS-1. A 20 amino acid extension at the N-terminal of the fructose bisphosphate aldolase aligned with a predicted PTS-2 sequence of mammalian aldolase, and experiments were done with fusion proteins containing CAT to elucidate its function (Blattner *et al.*, 1995). It was shown that the aldolase N-terminal signal could import CAT into the glycosome. A third PTS has been suggested, as there are several proteins localised to the glycosome that do not contain the PTS sequences previously described. It was proposed that two positively charged clusters of amino acids on the surface of the protein are required for importation into the glycosome (Wierenga *et al.*, 1987), although this theory now seems unlikely (Kendall *et al.*, 1990).

The glycolytic enzymes in *Trypanosoma* and *Leishmania* species have been investigated for their molecular and biochemical properties. Hexokinase and fructose 1,6 bisphosphate aldolases from *T. brucei* and *L. mexicana* are mainly glycosomal and have an N-terminal PTS-2 (Clayton, 1985; De Walque *et al.*, 1999; Opperdoes and Michels, 2001). Glucose-6-phosphate isomerase (G6PI) was shown to be bicompartmental, being present in both the cytosol and the glycosome. G6PI in *L. mexicana* and *T. brucei* has the C-terminal PTS-1 AHL and SHL, respectively (Nyame *et al.*, 1994; Marchand *et al.*, 1989). Glyceraldehyde-3-phosphate dehydrogenase

(GAPDH) is present in the *L. mexicana* genome in 2 isoforms, one encoding a glycosomal protein that has the PTS-1, SKM, and the other gene encoding a cytosolic enzyme (Hannaert *et al.*, 1992). *T. brucei* and *T. borelli* glycosomal GAPDH has the PTS-1, AKL (Wiemer *et al.*, 1995; Misset *et al.*, 1987). Phosphoglycerate kinase (PGK) in *T. brucei* has three isoforms: PGK-A has a minor glycosomal compartmentation, PGK-B is entirely cytosolic and PGK-C is glycosomal and has the C-terminal PTS-1, SSL (Misset and Opperdoes, 1987). PGK-A was shown to have an internal region essential for glycosomal targeting (Peterson *et al.*, 1997). Triosephosphate isomerase (TIM) from *L. mexicana* has been purified from glycosomes however both the *L. mexicana* and the *T. brucei* TIM do not contain a PTS-1 or PTS-2 (Kohl *et al.*, 1994).

Recently, some of the enzymes of the pentose phosphate pathway (PPP) have been cloned and sequenced in *T. brucei* and *L. mexicana*. Glucose-6-phosphate dehydrogenase (G6PDH), the first enzyme in the PPP, was shown to be partially glycosomal (~ 45 %) using subcellular fractionation experiments (Heise and Opperdoes, 1999). Subsequent molecular cloning and characterisation of this gene has shown that G6PDH does not contain a PTS-1 or a PTS-2 (Duffieux *et al.*, 2000). No isoenzymes have been found thus far. The G6PDH from *L. mexicana* was shown to be partially localised to the glycosome (Mottram and Coombs, 1985), however the gene sequence has not been published as yet. 6-phosphogluconolactonase (6PGL) from *T. brucei* was also sequenced and does contain the PTS-1, C-terminal tripeptide AKF (Duffieux *et al.*, 2000). This enzyme was shown to be bicompartmentalised, with ~ 10 % of the 6PGL activity related to the glycosome (Duffieux *et al.*, 2000). G6PDH and 6PGL genes are thought to be bifunctional proteins in mammals and *Plasmodium falciparum* forming a fusion protein (Collard *et al.*, 1999; Clarke *et al.*, 2001). It is possible that the two

proteins in *T. brucei* fuse before importation into the glycosome, using the PTS from the 6PGL to enter the microbody. The 6-phosphogluconate dehydrogenase (6PGDH) gene was cloned and overexpressed from *T. brucei* and was shown not to contain a PTS on the C- or N-terminal (Barrett and Le Page, 1993; Barrett *et al.*, 1994). This enzyme was subsequently shown to be entirely cytosolic in *T. brucei* (Heise and Opperdoes, 1999). The 6PGDH genes from *L. major* and *L. tropica* have recently been sequenced by Greenblatt and co-workers (accession numbers AF242436 and AY045763, respectively) however no further characterisation has been completed. The sequences do not contain a PTS-1 or PTS-2, however a PTS-1 type sequence has been found within the protein sequence and crystal structure studies suggest that these residues may lie on the surface of the protein, therefore may have access to the glycosome (Dr. M.P. Barrett, personal communication). Ongoing studies by Cazzulo and co-workers suggest that all the PPP enzymes in *L. mexicana* are associated at a low level with the glycosome, but are mainly cytosolic (Prof. J. J. Cazzulo, personal communication).

A partial PPP may be present in the glycosome, as the initial two enzymes of the pathway have been shown to be bicompartmental. The enzymes found to be within the glycosomes are in the oxidative branch of the pathway that is primarily responsible for NADPH production. It is possible that the glycosomes have a requirement for NADPH, as it is necessary in reductive biosynthetic pathways such as the ether lipid and sterol synthesis, known to be present in the glycosome. All of the enzymes in the PPP in *L. mexicana* are mainly cytosolic, but a small percentage has been shown to be present in the glycosomes. The PPP may play an important role, in maintaining NADPH levels and ribose 5-phosphate availability, as purine and pyrimidine biosynthetic pathways are also present in the glycosome.

1.9.4.2. Evolution of the glycosome and glycolytic enzymes

Glycosomes are not present in Euglenids, suggesting that these organelles must have been lost from the Euglenida lineage and remained in the Kinetoplasts (Michels *et al.*, 2000). It is unclear as to the origin of glycosomes in trypanosomatids. As the overall organisation and structure of peroxisomes and glycosomes are similar, a common origin of these microbodies is likely. This may have occurred with the entry of an endosymbiont and subsequent gene transfer to the nucleus, or a specialisation of a pre-existing compartment (Borst, 1989; Latruffe and Vamecq, 2000).

The origin of the enzymes of glucose metabolism in trypanosomatids has been of interest recently as more DNA sequencing data is obtained. GAPDH has two isoforms encoding a glycosomal enzyme and a cytosolic enzyme in trypanosomatidae that are arranged in a tandem repeat. The two enzymes in *T. brucei* and *L. mexicana* are distantly related and have evolved slowly (Hannaert *et al.*, 1992). It has been postulated that an acquisition of a foreign gene has occurred by horizontal gene transfer in a common ancestor, before the divergence of *Trypanosoma* and *Leishmania*, thus allowing the original enzyme to be involved exclusively in glycolysis in both organisms (Hannaert *et al.*, 1992; Wiemer *et al.*, 1995). Euglenida and Kinetoplastida are closely related organisms thought to have diverged from a single lineage (see Section 1.1). The cytosolic GAPDH in *Euglena gracilis* is orthologous to the glycosomal GAPDH from trypanosomatids (Henze *et al.*, 1995). The PTS is not present on the GAPDH of *Euglena* suggesting that this is a modification that has occurred subsequent to the split in lineage and perhaps due to the presence of a peroxisome-like organelle. The sequencing of chloroplast 6PGD from spinach has led to the phylogenetic analysis of homologues of 6PGD and G6PDH (Krepinsky *et al.*, 2001). *T. brucei* 6PGDH and G6PDH both branch with plants, with 6PGD specifically clustering with cyanobacteria.

This suggests a common ancestry of some genes in Kinetoplastids and higher plants with a cyanobacterial origin. The 6PGDH genes with a cyanobacterial affinity are also been shown to be present in several nonphotosynthetic protists (Andersson and Roger, 2002). Cyanobacteria are thought to be closely related to the chloroplasts in plants and algae (Martin and Schnarrenberger, 1997). Many other genes involved in glucose metabolism of trypanosomatids are shown to cluster phylogenetically with plants and cyanobacterial (Hannaert *et al.*, unpublished, F.R. Opperdoes, personal communication). There are also several plant-like genes encoded for in the nuclear genome of trypanosomatids such as sedoheptulose-1, -7, -bisphosphatase. It has been postulated that an organism ancestral to Euglenida and Kinetoplastida contained an algal endosymbiont that was subsequently lost from the Kinetoplast lineage, leaving remnants of genetic information within the nuclear genome of the Kinetoplastids (Hannaert *et al.*, unpublished, F.R. Opperdoes, personal communication).

1.10. Pentose Phosphate Pathway

The pentose phosphate pathway (PPP) utilises glucose to produce NADPH and phosphorylated sugars via a series of reactions. The products of this pathway have several important functions and therefore, not surprisingly, the pathway has been found in all tissues investigated so far. The PPP can be divided into two distinct branches, the oxidative branch and the non-oxidative branch (Figure 1.5). Two NADP-requiring dehydrogenases and a lactonase converts glucose 6-phosphate (G6P) to ribulose 5-phosphate (Ru5P), providing the cell with two molecules of NADPH. The enzymes of the non-oxidative branch of the PPP catalyse the interconversion of 3- to 7- carbon sugar phosphates. Transketolase and transaldolase reversibly catalyse the production of ribose 5-phosphate (R5P), erythrose 4-phosphate (E4P), xylulose 5-phosphate (X5P), fructose 6-phosphate (F6P) and glyceraldehyde 3-phosphate (GA3P).

NADPH is required for many enzymatic reactions such as fatty acid biosynthesis, maintenance of reduced thiols (eg. glutathione), drug detoxification, proline biosynthesis and structural integrity of cells (reviewed by Wood, 1986 (a)). The intracellular thiol, glutathione, is maintained by glutathione reductase, an NADPH-dependent enzyme, and plays an important role in protecting the cell against oxidative stress (reviewed by Krauth-Siegel and Coombs, 1999). Trypanosomatids have a unique glutathione metabolism with two glutathione molecules conjugated to the polyamine spermidine, termed trypanothione [N^1,N^8 - bis (glutathionyl) spermidine]. The NADPH-linked enzyme trypanothione reductase maintains reduced trypanothione and therefore maintains an intracellular reducing environment. Reactive oxygen intermediates such as hydrogen peroxide and hydroxyl radicals will be encountered by the parasite via internal cofactors and drug metabolism and externally by the host's immune system. Trypanothione can trap hydroxyl radicals and is subsequently

converted to a free thiol state using an NADPH oxidoreductase (reviewed by Fairlamb and Cerami, 1992). Trypanothione has been localised to the cytosol in trypanosomatids and the presence of the PPP in the cytosol suggests that a major role of the pathway is providing NADPH to maintain the reducing environment.

Transketolase and pentosephosphate isomerase reversibly catalyse the production of R5P, a component necessary for the synthesis of ribonucleotides that are incorporated into nucleic acids. In plants, R5P and X5P are both utilised in the Calvin cycle as well as the PPP. E4P is a precursor in the synthesis of aromatic amino acids and vitamins in microorganisms and plants via the shikimic acid pathway (Ganem, 1978; Draths *et al.*, 1992). S7P is a component of some bacterial cell walls, specifically the lipopolysaccharide layer. The interconversion of phosphorylated sugars is reversible and is dependent upon cellular conditions and requirements (Stryer, 1988).

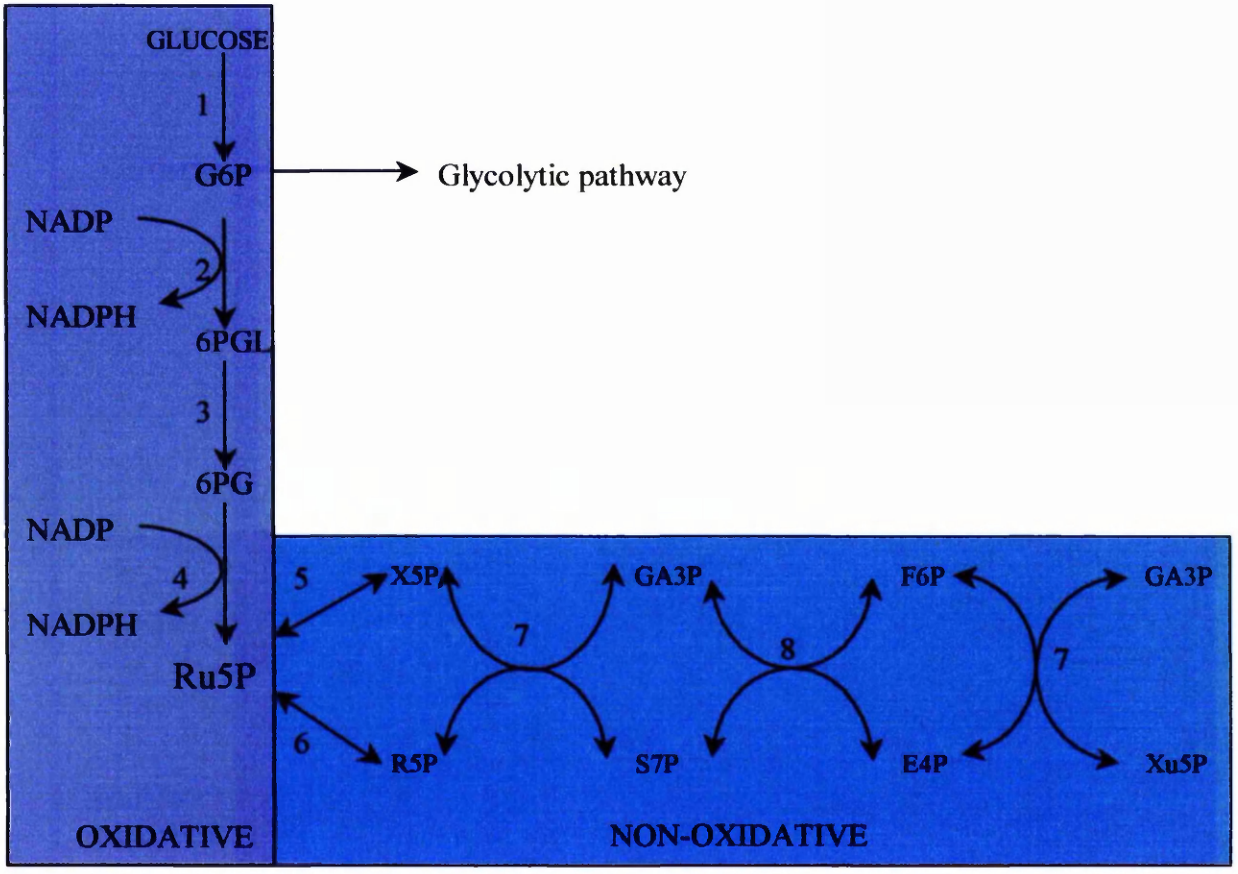


Figure 1.5: The pentose phosphate pathway. The pathway is split into the oxidative and non-oxidative branches. Enzymes from 1 to 8 are: hexokinase, glucose-6-phosphate dehydrogenase, 6-phosphogluconolactonase, 6-phosphogluconate dehydrogenase, ribulose-5-phosphate epimerase, pentosephosphate isomerase, transketolase and transaldolase. Abbreviations: G6P, glucose 6-phosphate; 6PGL, 6-phosphogluconolactone; 6PG, 6-phosphogluconate; Ru5P, ribulose 5-phosphate; X5P, xylulose 5-phosphate; R5P, ribose 5-phosphate; GA3P, glyceraldehyde 3-phosphate; S7P, Sedoheptulose 7-phosphate; F6P, fructose 6-phosphate; E4P, erythrose 4-phosphate.

1.10.1. Occurrence of the Pentose Phosphate Pathway

The distribution of the PPP has been measured in mammalian tissue, fish, birds, fungus, bacteria and parasites. The pathway is most active in tissue such as liver, adipose tissue and tumour cells that all possess high proliferation rates (Wood, 1986 (b)). The PPP is present in various parasitic protozoans, including *Leishmania*, *Trypanosoma*, *Plasmodium*, *Entamoeba* and *Giardia* (reviewed by Barrett, 1997). The PPP was initially identified in *L. donovani* promastigotes as it was shown that ribose could be used as a carbohydrate source in place of glucose (Ghosh and Datta, 1971). In this initial study of the PPP, 6PGDH was not detected in crude lysates however, subsequent studies did detect the presence of this enzyme in various species of *Leishmania* (Martin *et al.*, 1976; Janovy, 1972). A significant amount of glucose was shown to be metabolised via the PPP (Berens *et al.*, 1980; Keegan *et al.*, 1987) (see Section 1.10.3). It was also shown that as *Leishmania* promastigotes entered stationary phase in culture, a decrease in the flux of glucose through the PPP occurred concurrent with a decrease in the overall rate of oxidation of ribose (Keegan *et al.*, 1987). This suggests that varying life-cycle stages of the parasite have different requirements for carbohydrate metabolism. Subcellular localisation studies of *L. mexicana* promastigotes suggested that G6PDH was mainly cytosolic, however a small but significant percentage was present in the glycosome (Coombs *et al.*, 1982; Mottram and Coombs, 1985). G6PDH was present in the amastigotes stage however only exhibited approximately 50 % of the promastigote enzyme activity. More recently, all of the enzymes of the PPP in *L. mexicana* have been studied and shown to be partially associated with glycosomal microbodies (Cazzulo, personal communication).

All of the enzymes of the PPP have been identified in *T. brucei* procyclic cells (insect stage), however in the bloodstream forms both transketolase and ribulose 5-

phosphate epimerase could not be detected (Cronin *et al.*, 1989). As transaldolase was shown to be present in bloodstream forms it was suggested that a novel mechanism of carbon transfer might exist to complete the non-oxidative pathway (Cronin *et al.*, 1989). G6PDH and 6PGL have been recently cloned and characterised, and are both partially glycosomal (Heise and Opperdoes, 1999; Duffieux *et al.*, 2000). The 6PGDH from *T. brucei* has been studied in detail with the crystal structure resolved to 2.8 Å (Phillips *et al.*, 1998; Phillips *et al.*, 1993). 6PGDH is thought to be a potential drug target in *Trypanosoma* due to differences between the *T. brucei* enzyme and the mammalian isoform (Hanau *et al.*, 1996). Analogues of suramin, a trypanocidal drug known to inhibit 6PGDH, are currently being developed and have been shown to inhibit this enzyme (Opperdoes and Michels, 2001). These studies suggest that a functional PPP is present in *Trypanosoma* and may be bicompartmental, with a small percentage of activity being associated with glycosomes.

1.10.2. Intracellular distribution

The enzymes of the PPP are thought to reside in most organisms in the cytosol (Wood, 1986 (b)). However pentose phosphate enzyme activity has been recorded in other compartments of the cell. All of the enzymes in the PPP in rat tissue are known to be present in the cytosol and the endoplasmic reticulum (ER) (Bublitz and Steavenson, 1988). G6PDH has recently been identified in the ER of cells present in the small intestine of rabbits (Ninfali *et al.*, 2001).

Transketolase and ribulose-5-phosphate epimerase have both been isolated from spinach chloroplasts (Teige *et al.*, 1998). These enzymes have a role in the Calvin cycle in this organelle as well as the PPP.

G6PDH and 6PGDH are present in the peroxisomes of rat liver and have been associated with a large particulate fraction from other rat tissue (Antonenkov, 1989; Zaheer Baquer and McLean, 1972). The non-oxidative enzymes were present only in the cytosol. The main biological significance suggested for this localisation was the intra-peroxisomal requirement for NADPH, as several NADPH-dependent enzymes and pathways exist in this location. The transketolase in methanol-utilising yeast *H. polymorpha* is dihydroxyacetone synthase (DHAS) and has been shown to be present in the peroxisomes (Janowitz *et al.*, 1985). The PPP is thought to be bicompartmental in trypanosomatids, being partially distributed in the glycosomes that are thought to be specialised peroxisomes (see Section 1.9.3). However the PPP is thought to be mainly cytosolic in these parasites and may play an important role in maintaining NADPH levels for reducing trypanothione and other thiols. GAPDH in trypanosomatids does have an enzyme exclusively located in the cytosol, and might therefore be involved in metabolising glyceraldehyde 3-phosphate, made by the PPP in the cytosol (Hannaert *et al.*, 1992). During this study, subcellular localisation techniques were used to determine the distribution of transketolase in *L. mexicana* promastigotes.

1.10.3. Contribution of glucose flux through the PPP

The relative flux of glucose through the PPP can be measured by following the evolution of CO_2 from radiolabelled carbon. 1-[^{14}C] glucose and 6-[^{14}C] glucose both become 3-[^{14}C] glyceraldehyde 3-phosphate via glycolysis and release $^{14}CO_2$ from pyruvate. 1-[^{14}C] glucose is also incorporated into the PPP and is released as $^{14}CO_2$. The difference represents the conversion of 6PGL to Ru5P, which concurrently releases CO_2 and therefore the glucose entering the PPP can be measured. The glucose flux into the PPP is entirely dependent on culture conditions. The PPP in *T. cruzi* was shown to catabolise up to 40 % of available glucose, indicating that this pathway is relatively important in this parasite (Mancilla and Naquira, 1964). Using similar techniques, an appreciable flux of glucose through the PPP was shown to occur in *L. brasiliensis* and *L. donovani* (Berens *et al.*, 1980; Keegan *et al.*, 1987). ^{14}C-labelling experiments evaluating glucose flux in *L. mexicana* in culture conditions under oxidative stress are currently in progress (Cazzulo, poster, World Leish conference, Crete, May 2001). Oxidative stress can be induced in trypanosomatids by methylene blue that causes an increase in NADPH oxidation, thus activating the oxidative PPP. Preliminary data suggests that an increase of 13 % of glucose utilisation via the PPP occurred under oxidative stress, while the uptake of glucose via glycolysis did not alter. These results indicate that the PPP may play an important role in maintaining NADPH levels during oxidative stress encountered by the parasite. However this experimental method does not take into account reaction reversibility and recycling of the non-oxidative pathway products, therefore differences in data may arise (Follstad and Stephanopoulos, 1998).

1.10.4. The PPP as a potential drug target

Various studies indicate that the PPP plays an important role as a metabolic pathway in various organisms and therefore may be regarded as a potential drug target. The reduction of transketolase activity in plants resulted in decreased levels of aromatic amino acids and phenylpropanoid intermediates that are synthesised via the shikimic acid pathway (Henkes *et al.*, 2001). Photosynthesis was inhibited and a local loss of chlorophyll and carotene also occurred at high levels of transketolase inhibition. This study suggests that transketolase is important for photosynthesis and aromatic amino acid production. Peroxide sensitive yeast mutants were isolated and characterised and were shown to be deficient in genes encoding enzymes of the PPP (Juhnke *et al.*, 1996). Reduced glutathione reacts with hydrogen peroxide to prevent oxidative stress, and reduced levels of NADPH may inhibit this protective response. The non-oxidative pathway mutants show the same phenotype and it was suggested that feed-back inhibition might be occurring (Juhnke *et al.*, 1996).

The PPP is upregulated in proliferative cells and tumour tissue and is thought to be correlated to an increase in a cellular requirement for nucleic acids as 85 % of ribose 5-phosphate synthesised in tumour cell RNA was shown to be synthesised by transketolase (Boros *et al.*, 1997). Oxythiamine, a competitive inhibitor of transketolase, inhibits tumour cell proliferation by up to 39 %. When used in conjunction with a G6PDH inhibitor (dehydroepiandrosterone-sulphate) inhibition was increased to 60 % (Boros *et al.*, 1997). Therefore the enzymes of the PPP are potential drug targets and the PPP is currently being investigated as an anticancer drug target.

1.11. Transketolase

Transketolase is a thiamine pyrophosphate-dependent enzyme that catalyses the reversible transfer of a 2-carbon glycoaldehyde unit from a ketone donor to an aldehyde acceptor (reviewed by Schenk *et al.*, 1998 (a)) (Figure 1.6.). Transketolase catalyses two reactions in the non-oxidative branch of the PPP (Figure 1.5.) that are linked by transaldolase with the purpose of interchanging phosphorylated sugars. This enzyme is also involved in the photosynthetic Calvin cycle in plants and bacteria where it utilises the same substrates as in the PPP (reviewed by Martin and Schnarrenberger, 1997). Within a cell, transketolase transfers a 2-carbon unit from xylulose 5-phosphate to either ribose 5-phosphate or erythrose 5-phosphate, which accepts the residues to make glyceraldehyde 3-phosphate, sedoheptulose 7-phosphate and fructose 6-phosphate. Mammalian transketolase can only utilise a specific set of substrates, all of which are mentioned above. Certain transketolases can utilise a broad range of substrates including those mentioned above and dihydroxyacetone phosphate, hydroxypyruvate and glycoaldehyde. Transketolase can also utilise unphosphorylated sugars as substrates, however they have a much lower affinity for binding as the phosphate group of the substrates interact directly with the enzyme (Villafranca and Axelrod, 1971; Nilsson *et al.*, 1997). A 100-fold increase in the K_m values of D-ribose over ribose 5-phosphate of transketolase from spinach can be estimated, however the two assays cannot be directly compared from the two studies as different acceptor substrates were used (Villafranca and Axelrod, 1971; Teige *et al.*, 1998). Due to this broad substrate specificity, transketolase is used in industry to make a variety of products (see Section 1.11.5).

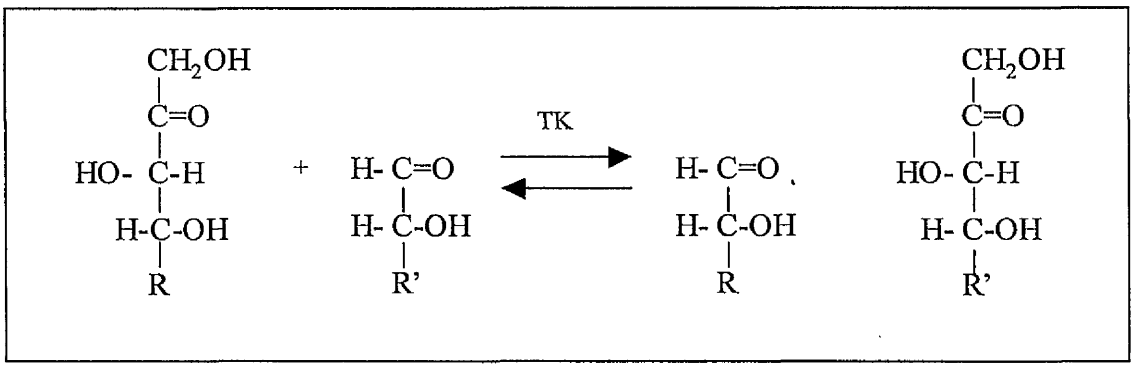


Figure 1.6. Scheme of the transketolase reaction. The R group represents a phosphorylated carbon group of one or more carbons.

Transketolase has been isolated from a wide variety of organisms including humans (McCool *et al.*, 1993; Schenk *et al.*, 1998 (b)), other mammals (Salamon *et al.*, 1998), yeast (Racker, 1961), plants (Bernacchia *et al.*, 1995), bacteria (Sprenger, 1993), and parasites (Cronin *et al.*, 1989). Due to the advance of DNA sequencing and the rapid progress of genome sequencing, several transketolase protein sequences have recently become available. During this study many of these available protein sequences were aligned and subjected to phylogenetic analysis to determine the relative ancestry of the proteins.

Transketolase plays a structural role as it is abundantly expressed in the corneal tissue of mice, humans and rabbits (Sax *et al.*, 1996; Jester *et al.*, 1999). Transketolase is thought to be an enzyme-crystallin playing a role in the transparency and refractive properties of the lens. An increase in transketolase expression has been linked to the maturation of the eye, with low levels being seen prior to the eye opening (Sax *et al.*, 1996). Reduced levels of transketolase were also shown in opaque scleral fibroblasts or from injured regions of the cornea in keratocytes (Jester *et al.*, 1999).

Transketolase requires both thiamine pyrophosphate (TPP) and divalent cations such as magnesium ions to be catalytically active. A TPP-binding consensus motif was identified in TPP-binding enzymes, including the transketolase protein sequence (Robinson and Chun, 1993; Hawkins *et al.*, 1989). The yeast transketolase crystal structure showed that this motif does bind TPP in the cleft of the dimer (Lindqvist *et al.*, 1992). Other TPP-dependent enzymes include pyruvate dehydrogenase of the E1 complex, pyruvate decarboxylase, pyruvate oxidase, acetolactate and dihydroxyacetone synthase. Transketolase is functional in a dimer form, with individual subunits of 65-74 kDa, depending on the organism the protein is from. Crystal structures have been derived for the *S. cerevisiae* and the *E. coli* transketolases and the three dimensional structures of the *S. cerevisiae* has been determined (Lindqvist *et al.*, 1992; Nikkola *et al.*, 1994; Littlechild and Turner, 1995).

1.11.1. Reaction mechanism of transketolase

The transketolase protein from yeast has been crystallised and the structure examined (Lindqvist *et al.*, 1992; Nikkola *et al.*, 1994). The transketolase monomer is a V-shaped subunit consisting of three domains; an N-terminal domain involved in co-factor binding, a middle domain also involved in TPP-binding and substrate binding, and a C-terminal domain that has no known function as yet (Lindqvist *et al.*, 1992). The N-terminal domain and the middle domain interact at the subunit-subunit interface of the dimer, forming the active site of the enzyme. Two molecules of TPP bind at the dimer interface and are necessary for the formation of the functional dimer. The diphosphate group of the cofactor forms hydrogen bonds at the N-terminus, and binds to divalent cations (Ca^{2+} or Mg^{2+}). The thiazolium and pyrimidine rings (Figure 1.7) are

bound deep in the cleft of the active site, where a number of hydrogen bonds form with conserved residues of the transketolase (discussed in Section 4.7).

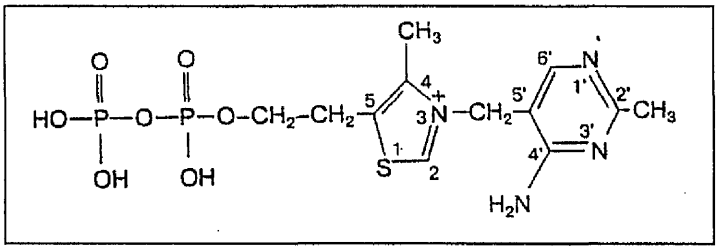


Figure 1.7: Thiamine pyrophosphate structure. Taken from Kochetov, 2001.

The reaction begins as the TPP interacts with the transketolase monomer, forming a catalytically inactive complex. A slow conformational rearrangement occurs, where the two subunits interact, producing an active holoenzyme (Heinrich *et al.*, 1972; Egan and Sable, 1981). This interaction of apotransketolase with the cofactor TPP has been shown to be the reason for the lag phase exhibited at the beginning of the reaction (Booth and Nixon, 1993). The donor substrate is bound to TPP in the active site pocket, that is a deep cleft formed between the two transketolase monomers. The C-2 of the thiazolium ring of TPP is deprotonated, leading to the cleavage of the ketose donor substrate at the carbonyl group. This substrate is covalently bound to the cofactor, where the C-3 hydroxyl group is deprotonated, yielding an aldose product. The glycoaldehyde unit attached to the TPP (dihydroxyethyl TPP) reacts with the aldose acceptor substrate and is released from the thiazolium ring producing a ketose.

It has been shown that a transketolase reaction can occur with only one substrate. Using xylulose 5-phosphate as a donor substrate, erythrulose is formed as individual glycoaldehyde units are released into the solvent and collide with other units (Fiedler *et al.*, 2001; Bykova *et al.*, 2001; Solov'eva *et al.*, 2001). This shows that

transketolase can cleave the donor ketose in the absence of an acceptor molecule. The transketolase reaction is thought to be consistent with the Bi-Bi Ping-Pong mechanism of action, with the initial substrate being released prior to the binding of the second substrate (Calvier and Sable, 1972; Gyamerah and Willetts, 1997; Fiedler *et al.*, 2001).

1.11.2. Multiple forms of transketolase

Multiple copies of the transketolase gene have been found in various organisms. Three forms of transketolase from yeast have been reported (Kuimov, 1990). The multiple forms of transketolase were identified based on the differential elution patterns when separated by phosphocellulose chromatography and thermostability. However, only two transketolase genes have been isolated from yeast (Schaff-Gerstenshläger and Zimmermann, 1993; Sundström *et al.*, 1993). The third form of transketolase is possibly due to post-translational modifications of the protein, or perhaps a third divergent gene is present in the yeast genome.

Three transketolase isoenzymes have been found in the *Craterostigma plantagineum*, two of which are upregulated during the rehydration process (Bernacchia *et al.*, 1995). One of the genes *tkt3* is constitutively expressed in all conditions in both roots and leaves, whereas *tkt7* accumulates preferentially during the rehydration phase in both roots and leaves. *tkt10* was highly expressed during the rehydration phase, specifically present in the leaves of the plant (Bernacchia *et al.*, 1995).

Two copies of the transketolase gene have been identified and characterised in *E. coli*. Mutants defective in transketolase were isolated and were unable to grow on any pentose as a sole carbon source (Josephson and Fraenkel, 1969; Josephson and Fraenkel, 1974). In addition, the mutants required a supplement of aromatic amino acids or shikimic acid for normal growth, however this requirement was 'leaky', leading

to the speculation of a second transketolase. Subsequently *tktA* was cloned, sequenced and thoroughly characterised (Sprenger, 1993; Sprenger *et al.*, 1995). A second transketolase gene, *tktB*, was isolated and shown to be responsible for only minor transketolase activity in comparison to *tktA* (Iida *et al.*, 1993). *tktA tktB* double mutants displayed a phenotype that required pyridoxine, aromatic amino acids and vitamins, suggesting that transketolase plays an important role in the production of erythrose 4-phosphate, a precursor in aromatic amino acid synthesis in *E. coli* (Zhao and Winkler, 1994).

1.11.3. Alternative transketolases

1-deoxy-D-xylulose-5-phosphate synthase and dihydroxyacetone synthase are both novel transketolases as they catalyse reactions not normally catalysed by transketolase.

Dihydroxyacetone synthase (DHAS) is a formaldehyde transketolase of the methanol-utilising yeast *Hansenula polymorpha* (Janowicz *et al.*, 1985). DHAS is a key enzyme in methanol metabolism and resides in the peroxisomes of the cell (Hansen *et al.*, 1992). The enzyme transfers a 2-carbon unit from xylulose 5-phosphate to formaldehyde, the first metabolite of methanol oxidation. It has been described in other methylotrophic organisms including *Candida boidinii* and *Acinetobacter* species.

Another unique family of transketolases has been identified that synthesise isoprenoids via a non-mevalonate pathway in eubacteria, plant chloroplasts and green algae (Lange *et al.*, 1998). Isoprenoids are present in plants as hormones, photosynthetic pigments, electron carriers, and structural components of cell membranes and also play a role in communication and defence of the cell. Isoprenoids are known to be synthesised by a mevalonate pathway and a recently described non-

mevalonate pathway (Rohmer *et al.*, 1993; Lichtenthaler *et al.*, 1997; Eisenreich *et al.*, 1997). The non-mevalonate pathway synthesises isopentenyl diphosphate (IPP) from pyruvate and GA3P, with the first step in the pathway being catalysed by the TPP-dependent enzyme 1-deoxy-D-xylulose-5-phosphate synthase (DXPS). 1-deoxy-D-xylulose 5-phosphate (DOXP) is then converted to 2-C-methyl-D-erythritol 4-phosphate via DOXP reductoisomerase. DXPS has now been cloned from a number of organisms including *E. coli* (Lois *et al.*, 1998), peppermint (Lange *et al.*, 1998), *Rhodobacter capsulatus* (Hahn *et al.*, 2001) and *Plasmodium falciparum* (Jomaa *et al.*, 1999). The protein sequences of DXPS do show similarity to transketolases and are therefore thought of as a novel family of transketolases (Lange *et al.*, 1998). In a study to identify key residues of DXPS in *E. coli*, histidine^{49} was shown to be essential for catalytic activity (Querol *et al.*, 2001). This residue in yeast (histidine^{30}) has been previously shown to be present in the active site of the enzyme and is involved in proton transfer during catalysis (Lindqvist *et al.*, 1992; Wikner *et al.*, 1997). Inhibitors of the non-mevalonate pathway have been described and this metabolic pathway is thought to represent an interesting drug target (Altincicek *et al.*, 2000; Jomaa *et al.*, 1999). In *P. falciparum*, fosmidomycin, an inhibitor of DOXP reductoisomerase was shown to suppress multi-drug resistant parasites (Jomaa *et al.*, 1999).

1.11.4. Transketolase-linked disease

Abnormal biochemical characteristics of transketolase have been related to the neurophychiatric disorder Wernicke-Korsakoff-syndrome (WKS). WKS is characterised by chronic memory disorder and is linked to thiamine-deficiency caused by malnutrition (often brought on by alcoholism). Transketolase isolated from patients with WKS exhibit a reduced affinity (an increased apparent K_m) for TPP-binding (Blass

and Gibson, 1977). However it has been shown that there is no difference in the protein sequence at the transcriptional level, and therefore must be a post-transcriptional modification of the enzyme (McCool *et al.*, 1993).

Transketolase is thought to be involved in the contribution of RNA ribose synthesis in tumour cell proliferation (Boros *et al.*, 1997; Boros *et al.*, 1998). Treatment of adenocarcinoma cells with the transketolase inhibitor oxythiamine, resulted in inhibition of cellular proliferation (Boros *et al.*, 1997). Cancer patients exhibit a depleted thiamine status even under normal dietary requirements, suggesting that thiamine-dependent enzymes may be utilising the co-factor (Basu and Dickerson, 1976). It has been proposed that thiamine supplementation in Western diets could enhance tumour proliferation as it is the co-factor for transketolase reactions (Boros, 2000).

1.11.5. Use of transketolase in industry

Transketolase is used in industry for the synthesis of chemicals and stereo-specific C-C bond synthesis. Transketolase has been purified from several organisms and expressed as recombinant proteins to be used in industry. Spinach transketolase has been used to synthesise 6-deoxy-L-sorbose, a precursor of furaneol, that is an aromatic product with a caramel-like flavour (Hecquet *et al.*, 1996). Transketolase is used in the industrial biosynthesis of aromatic amino acids via the shikimic acid pathway. Erythrose 4-phosphate and phosphoenolpyruvate (PEP) combine to produce 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP), a precursor of chorismic acid that acts as a branch point of various biosynthetic pathways (Ganem, 1978). The natural products L-phenylalanine, L-tryptophan and L-tyrosine are all made in industry from glucose using the *E. coli* transketolase (Draths *et al.*, 1992). L-phenylalanine is used as the artificial sweetner aspartame, L-tryptophan can be used as the dye indigo and L-tyrosine is used

as eumelanin, a unique UV-absorbing substance. Transketolase can also be used as a biocatalyst in biotransformations to make products such as erythrulose from hydroxypyruvate (Lilly *et al.*, 1996). Hydroxypyruvate is a useful synthetic ketol donor, as CO_2 is produced during the reaction, deeming it irreversible. During this project, xylulose 5-phosphate was synthesised using fructose 1, 6-bisphosphate and hydroxypyruvate as the initial starting products, with fructose 1, 6-bisphosphate aldolase and transketolase to catalyse the reaction (see Section 4.5).

1.11.6. Transketolase in trypanosomatids

Transketolase activity has been detected in the promastigote forms of several *Leishmania* species, including *L. brasiliensis*, *L. donovani*, *L. mexicana* and *L. tropica* (Martin *et al.*, 1976). Transketolase activity was also detected in *T. brucei* procyclic form (insect stage) during an investigation of all the pentose phosphate enzymes (Cronin *et al.*, 1989). However, no transketolase or ribulose-5-phosphate epimerase activity was detected in the bloodstream forms of the parasite. The authors suggested that a novel mechanism of 2-carbon transfer occurs, as transaldolase was present in both life-cycle stages (Cronin *et al.*, 1989).

During this study, the transketolase from *L. mexicana* was identified and characterised. The promastigote crude lysates were assayed for transketolase activity and subcellular fractionation identified the location of the protein within the cell. The transketolase gene has been identified as part of this study, leading to the preparation of a recombinant transketolase that has been characterised. This study was done as part of a project to characterise the pentose phosphate pathway in the *Leishmania* parasite.

Aims of this study

This investigation of transketolase from *Leishmania mexicana* was part of a study to characterise the pentose phosphate pathway in *Leishmania* species. Several laboratories worldwide were involved in this project, with groups examining one specific enzyme within the pathway, or investigating the role of the pentose phosphate pathway in *Leishmania*. This collaboration is ongoing, and its main goal is to identify if the pentose phosphate pathway can act as a chemotherapeutic target.

The main aims of this study are:

1) To identify and clone transketolase from the parasite *Leishmania mexicana*
2) To characterise the transketolase gene in terms of gene copy number and compare the *L. mexicana* gene to transketolases from other organisms
3) To overexpress and purify a recombinant *L. mexicana* transketolase
4) To characterise the recombinant transketolase
5) To examine the native transketolase from crude lysates and identify the subcellular localisation of the enzyme
6) To knock-out transketolase from *L. mexicana* and identify if the gene is essential to the parasite

These aims are the basis of the investigation of transketolase in *L. mexicana* and may begin to determine if transketolase can act as a chemotherapeutic target.

Chapter II

Materials and Methods

2.1. Cell Culture Techniques

2.1.1. Culturing of *Leishmania mexicana* promastigotes

Handling of *Leishmania* parasites took place in a class II sterile hood. *Leishmania mexicana mexicana* (MNYC/BZ/62/M379) promastigotes were grown axenically *in vitro* at 25 °C in HOMEM (Berens *et al.*, 1976) (Gibco, Life Technologies) (see Section 2.5) with 10 % heat-inactivated foetal calf serum (FCS) (Labtech International). The culture was initiated at a density of 2 x 10^5 cells/ml. Log phase cultures were reached after 2-3 days when the cell density was approximately 5 x 10^6 cells/ml. After 7 days growth the culture had reached stationary phase, with a cell density of 1-2 x 10^7 cells/ml, and were sub-passaged into fresh medium.

2.1.2. Culturing of *Leishmania mexicana* amastigotes

Amastigotes recovered from *L. mexicana* lesions in mice and axenic amastigotes were obtained from David Laughland. The amastigotes were obtained from the rump of BALB/C mice, 2-3 months after the initial infection. The lesion material was treated with saponin to lyse the red blood cells and purified using a sephadex-CM-25/cellulose column. The axenic amastigotes originate from lesions from BALB/C mice and are maintained in SDM (Gibco, Life Technologies) and 20 % FCS, pH 5.5.

2.1.3. Harvesting of *Leishmania mexicana*

The cell density of cultures was calculated using an improved Neubauer haemocytometer. The culture was harvested in 15 ml tubes (Greiner) by centrifugation for 5-10 min at 1000 g at 4 °C (unless specified) and the supernatant removed.

2.1.4. Cell lysate preparation

2 x 10^8 cells of *L. mexicana* log-phase promastigotes were harvested and the pellet washed three times in iso-osmotic buffer (see Section 2.5) and resuspended in crude lysate lysis buffer (see Section 2.5). The suspension was incubated at 4 °C for one hour on a shaker and harvested at 11,000 g for 90 min at 4 °C. The supernatant was regarded as the cytosolic protein fraction.

2.1.5. Stabilate procedure of *Leishmania mexicana*

To stabilate *Leishmania* parasites, 1 ml of log-phase culture (7 x 10^6 – 2 x 10^7 cell/ml) was stored in 5 % DMSO in a cryotube. The vials were frozen at - 70°C overnight, and then transferred into liquid nitrogen for long-term storage.

2.1.6. Culturing of *Trypanosoma cruzi* epimastigotes

T. cruzi epimastigotes were cultured by Dr. Cazzulo, Buenos Aires, using monophasic medium (see Section 2.5) and 10 % FCS. The epimastigotes were grown at 28 °C on a shaker at 100 rpm.

2.2. Molecular Methods

2.2.1. Bacterial strains and plasmids

Bacterial strains *Escherichia coli* JM109 (Promega) and *E. coli* BL21(DE3) were used throughout this project. *E. coli* JM109 was used as a standard sub-cloning strain of *E. coli*, whereas *E. coli* BL21(DE3) was used to express recombinant proteins. The various plasmids used in this study are listed and described in Table 2.1.

Plasmid name	Uses	Manufacturer
pUC18	Sub-cloning vector	Amersham
pGEM-T	Sub-cloning vector	Promega
pET-16b	Expression vector (N-terminal His-tagged)	Novagen (Studier and Moffat, 1986)
pGL345	Confers Hygromycin B resistance	Dr. Hubert Denise (Glasgow) / Dr. Steven Beverley (St. Louis, USA.)
PGL520	Confers Nourseothricin resistance	Dr. Hubert Denise (Glasgow) / Dr. Steven Beverley (St. Louis, USA.)
PXG102	Re-expression vector used in *Leishmania*. Confers Neomycin resistance	Dr. Hubert Denise (Glasgow) / Dr. Steven Beverley (St. Louis, USA.)

Table 2.1: Summary of plasmids used in study.

2.2.2. Storage of bacterial strains

Bacteria and bacteria containing plasmids requiring long-term storage were kept as glycerol stocks. Individual colonies were selected from an LB agar plate and were grown overnight at 37 °C in 5 ml LB broth with or without ampicillin (100 μg/ml). 1 ml of overnight culture was used to inoculate a culture of LB broth and was grown at 37 °C until the optical density (OD) at 600 nm reached 0.6-0.8 Au. 0.6 ml of culture was then mixed with 0.4 ml of LB containing 30 % glycerol and was stored at -70°C. To inoculate a culture from a glycerol stock, a scraping was taken from the frozen vial and

streaked onto an LB agar plate with or without antibiotic. The plate was incubated overnight at 37 °C and an individual colony was picked to inoculate liquid medium.

2.2.3. Isolation of genomic DNA from *Escherichia coli*

100 ml of *E. coli* JM109 was grown in LB to an OD_{600} of 0.8 Au was reached. The culture was centrifuged for 10 minutes at 2,500 x g at 4 °C and the cells washed in 40 ml of TE. The cell pellet was resuspended in iso-osmotic buffer (see Section 2.5) containing 10 mg/ml lysozyme that caused lysis of the bacteria cell walls. This was incubated on ice for 5 minutes, then 1 % SDS was added on ice for a further 5 minutes. 0.5 mg/ml of proteinase K was added to the lysed cells and incubated at 51 °C for 3 hours to digest the protein component of the sample. The DNA was extracted from the sample by phenol-chloroform and precipitated using 0.1 volumes of ammonium acetate (3 M) and 2 volumes of 100 % ethanol. The nucleic acid was collected by centrifugation and air-dried for 10 minutes before being re-suspended in water.

2.2.4. Nucleic acid preparation from *Leishmania mexicana*

2.2.4.1. Isolation of genomic DNA

The protocol used to isolate DNA was based on the method of Medina-Acosta and Cross (Medina-Acosta and Cross, 1993). 1 x 10^8 parasites were harvested and re-suspended in 150 µl of TELT lysis buffer (see Section 2.5). The sample was incubated for 5 minutes at 25 °C. 150 µl of phenol/chloroform mixture (1:1 v/v) was added and mixed gently by inversion. Two phases were separated by centrifugation at 12,500 g for 5 minutes and the nucleic acids were precipitated from the upper phase by adding 0.1 volumes of sodium acetate (3 M) and 2 volumes of 100% ethanol. The sample was

gently mixed for 15 seconds and incubated on ice for 5 minutes. The nucleic acid pellet was collected by centrifugation at 13,000 g for 15 min and washed in 70% ethanol. The pellet was air dried and re-suspended at 200 ng/µl in TE buffer.

2.2.4.2. Isolation of total RNA

Total RNA was isolated from *L. mexicana* promastigotes using TRIzol reagent (Life Technologies, GibcoBRL), which was a mono-phasic solution of phenol and guanidine isothiocyanate. 2 x 10^8 cells were lysed in 1 ml of TRIzol reagent. The lysate was incubated at room temperature for 5 minutes and the RNA extracted with 0.2 ml of chloroform, incubating at room temperature for 2 minutes. Centrifugation at 12,000 g for 15 minutes at 4 °C separates the solution into organic and aqueous phases. The RNA remained exclusively in the clear upper aqueous phase and was recovered by precipitation with 0.5 ml of isopropyl alcohol at room temperature for 10 minutes followed by centrifugation at 12,000 g for 10 minutes at 4 °C. The RNA pellet was washed in 1 ml of 75% ethanol, air-dried for 10 minutes and re-suspended in DEPC-treated sterile ddd water. The RNA was stored at -70°C.

2.2.4.3. Handling RNA

When working with RNA, all reagents and equipment must be RNase free. RNase free water was prepared by adding 0.01 % (v/v) diethylpyrocarbonate (DEPC) to ddd water and leaving the bottle overnight in a fume hood at room temperature before autoclaving. All solutions were made with fresh material and DEPC water. Sterile plasticware and sterile filter pipette tips were used. Rubber gloves were worn during all RNA work as skin was a source of nucleases.

2.2.4.4. cDNA synthesis

One method used to characterise the 5' and 3' regions of the transketolase gene involved the amplification of DNA sequence from mRNA. 5 μg of total RNA was used to generate full-length first-strand cDNA using an oligo $d(T)_n$ primer and reverse transcriptase using Ready-To-Go T-Primed First-Strand Kit (Amersham Pharmacia Biotech). The manufacturer's instructions were followed. This reaction is summarised in Figure 2.1. The total RNA was heated to 65 °C for 5 minutes and then placed at 37 °C for 5 minutes. The First-Strand Reaction mix (Amersham Pharmacia Biotech) contained dATP, dCTP, dGTP, dTTP, Murine Reverse Transcriptase, BSA and Not I-$d(T)_{18}$ primer (5'-AAC TGG AAG AAT TCG CGG CCG CAG $GAAT_{18}$), and appeared as a translucent pellet. The reaction mix was equilibrated to 37 °C for 5 minutes and the RNA added. The reaction was gently vortexed and incubated at 37 °C for 60 minutes. The RNA-cDNA duplex was denatured at 90 °C for 5 minutes and used directly for PCR amplification with the addition of 2.5 units of Taq polymerase and gene specific primers.

2.2.5. Polymerase Chain Reaction

The polymerase chain reaction (PCR) amplified specific fragments of DNA from various nucleic acid templates. The PCR reaction components were: 1X PCR buffer (Promega), 1.5 mM $MgCl_2$ (Promega), 0.5 mM dNTPs (ABbioscience), 2.5 units of DNA Taq Polymerase (Promega/ABbioscience), 10-100 pmol of each primer (MWG-BIOTECH) and 1-100 ng DNA. 5 % DMSO was included in the PCR reaction mixture when *L. mexicana* DNA was used as a template to improve amplification. The

PCR cycle temperatures and times varied depending on DNA template and size of product required. PCR was performed using a GeneAmp 2400 PCR System (Perkin Elmer) or a DNA Engine PTC-200 Peltier Thermal Cycler (MJ Research).

2.2.6. Primer design

Oligonucleotides were designed to known regions of DNA to amplify various fragments using PCR. One primer was designed in the 5'-3' direction and the other primer was designed downstream to the complementary strand of DNA. When required, primers were designed to incorporate restriction sites so the PCR product could subsequently be cloned into a vector. Degenerate oligonucleotides were designed when the sequencing data available were amino acid sequences. The primers were synthesised by MWG-Biotech.

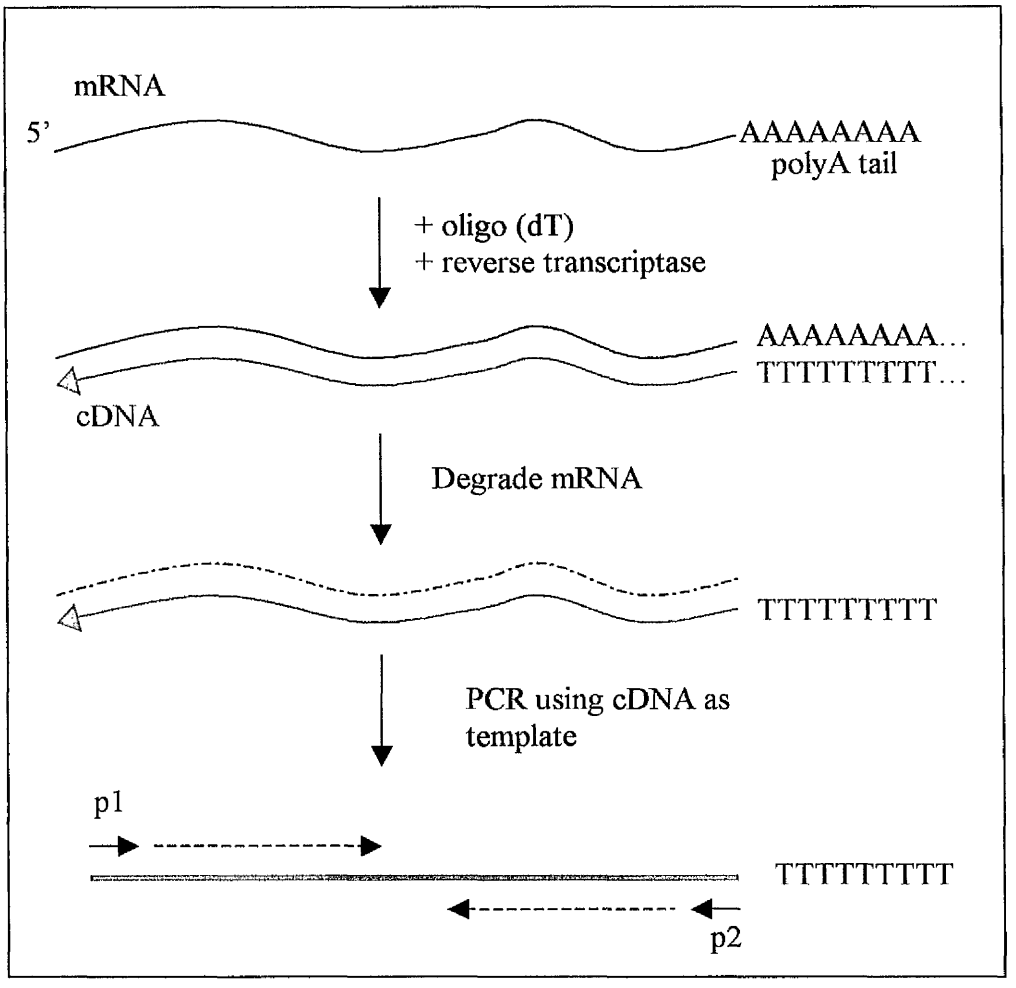


Figure 2.1: cDNA synthesis from mRNA to amplify 5' and 3'-ends of *TKT*. The primer oligo d(T) acted as an anchor to the poly A tail of the mRNA transcript and reverse transcriptase transcribes DNA from the RNA template. The mRNA was then degraded and the cDNA was used as a template in a PCR. Primer p1 with oligo dT can be used to amplify the 3'-end of a gene, whereas, primer p2 and a primer to the spliced leader sequence at the 5'-end of mRNA transcripts could be used to amplify the 5'-region of a gene.

2.2.7. Rapid Amplification of cDNA Ends (RACE)

RACE was a method used based on that previously described in section 2.2.4.4 (cDNA synthesis). However, both the 5' and 3'-RACE kits were commercially available from GIBCO BRL, Life Technologies and contained adaptor primers for the more specific amplification of cDNA-ends. To amplify the 3'-region of a gene, 5 μg of total RNA was incubated with the adaptor primer (see Table 2.2 for sequence) for 10 minutes at 70 °C. 5'-RACE required the RNA to be incubated with a gene specific primer to amplify cDNA. The sample was mixed with 1 x PCR buffer (Tris-HCl, pH 8.4 (200 mM), KCl (500 mM), $MgCl_2$ (25 mM), dNTP (500 μM each), DTT (10 mM) and 200 U of SUPERSCRIPT II Reverse Transcriptase. The reaction was incubated at 42 °C for 50 minutes to allow for the extension of the cDNA from the poly-A tail (3'-5'). The reaction was terminated by incubation at 70 °C for 15 minutes and 50 ng of RNase H was mixed into the tube to degrade the RNA in the sample.

Primer name	Amplification target	Primer sequence 5'-3'
Adaptor primer	3'	GGC CAC GCG TCG ACT AGT ACT TTT TTT TTT TTT TTT T
Abridged Anchor Primer	5'	GGC CAC GCG TCG ACT AGT ACG GGI IGG GII GGG IIG
Abridged Universal Amplification Primer	3' or 5'	GGC CAC GCG TCG ACT AGT AC

Table 2.2: Sequences of primers used in the RACE kit.

5'-RACE cDNA was then purified using a GLASSMAX® Spin Cartridge to exclude unincorporated dNTPs, gene specific primers and proteins. A homopolymeric tail was then added to the 3'-end of the cDNA to act as the binding site for the abridged anchor primer during PCR. The purified cDNA was incubated with 1x tailing buffer (Tris-HCl, pH 8.4 (10 mM), KCl (25 mM) and $MgCl_2$ (1.5 mM)), dCTP (2mM) and 1 µl of Terminal deoxynucleotidyl transferase (TdT) in a total volume of 24 µl for 10 minutes at 37 °C. Heating for 10 minutes at 65 °C terminated the reaction. To amplify the target sequence from cDNA, gene specific primers designed to known gene sequence were required and were used with the abridged anchor primer and the abridged universal amplification primer to amplify gene specific products. Nested PCR was recommended to amplify specific products.

2.2.8. Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate DNA fragments of various sizes according to the standard protocol (Sambrook *et al.*, 1989). Agarose powder (Sigma) was dissolved in 1 x TAE buffer by heating in a microwave oven. An agarose concentration of 0.8 - 2 % (w/v) was used depending on the size of DNA fragments in the sample to be separated. The solution was cooled to 50-60 °C and the gel was poured into a gel tray containing a comb. After the gel set the comb was removed and 1 x TAE buffer was poured into the electrophoresis tank chamber to cover the surface of the gel. The DNA samples were mixed with 6 x gel loading buffer (Promega) to a final concentration of 1 x, and loaded into the wells in the gel. To determine the size of the DNA fragments a 1 kb ladder (Promega) was loaded into the first well. The gel apparatus was connected to an electrical power supply and a voltage of 1-5V/cm was

applied allowing the DNA to migrate to the anode. The gel was stained in 0.5 μg/ml ethidium bromide (Sigma) for 15-30 min and washed in ddd water. Ethidium bromide interchelates with the DNA and fluoresces under ultraviolet light (260 nm). To view the DNA the gel was placed on an ultraviolet transilluminator and photographed by a gel imager (Ultra-Violet Products Ltd, Cambridge).

2.2.9. Nucleic acid concentration determination

The concentration of DNA/RNA in a sample was determined using a GeneQuant II (Pharmacia Biotech). This determined the difference in the absorbance of the DNA/RNA sample at A_{260}-A_{280}.

2.2.10. DNA extraction from agarose gels

Purification of individual bands of DNA from agarose gels was carried out using QIAquick gel elute kit (Qiagen). The DNA was briefly visualised on a UV transilluminator to prevent nicking of the DNA and was excised from the agarose gel using a sterile scalpel blade minimising the surrounding agarose. The nucleic acid was extracted from the agarose gel and the ethidium bromide using a QIAquick Spin column (Qiagen). The gel was dissolved and loaded into the spin column where the DNA was absorbed onto the column. The column was washed and the DNA eluted in sterile water.

2.2.11. Sub-cloning procedures

2.2.11.1. Blunt-ended ligation into pUC-18

Purified PCR products were routinely cloned into vector pUC18 using a blunt ended ligation process (SureClone Ligation Kit, Amersham Pharmacia Biotech). The DNA was blunt-ended and phosphorylated by Klenow fragment and T4 polynucleotide kinase, respectively. Following phenol/chloroform extraction and MicroSpin Column Purification the fragment was ligated into a blunt-ended vector (pUC18 Sma 1/BAP).

2.2.11.2. pGEM®-T vector cloning

pGEM®-T vectors (Promega) were pre-cut with Eco RV and contained a 3' terminal thymidine on both strands. This created a compatible overhang for PCR products, as Taq polymerase adds a single deoxyadenosine to the 3'-ends of the amplified fragments. This vector was also used routinely to clone PCR fragments. The reaction contained 1 x rapid ligation buffer (Promega), 50 ng of pGEM®-T vector, 1-50 ng of PCR product, T4 DNA ligase (3 Weisse units). The ligation mixture was incubated overnight at 4 °C.

2.2.12. Restriction site cloning

Vectors were cut at specific cloning sites using chosen restriction enzymes to produce a plasmid with 'sticky-ends'. This allowed the directional ligation of insert DNA containing complementary restriction sites. 3-5 μg of vector was digested with 10 U of restriction enzyme in a total volume of 50 μl for 4 hours at 25/37 °C. When a double digest was required often a change of buffer occurred. The first digestion was

de-salted using a MicroSpin column (Amersham Pharmacia Biotech) and digested with the second enzyme in the appropriate buffer.

2.2.13. Dephosphorylation of digested vector

To prevent self-ligation of a digested vector the 5' end phosphate group was removed using Calf Intestinal Alkaline Phosphatase (CIAP). This process decreases the number of non-recombinant vectors that occur when using a single enzyme to digest a vector, or when the restriction sites are close together on the vector. 0.01 U CIAP (Boehringer Mannheim) was used per pmol end of vector in a final volume of 100 µl. The reaction was incubated for 30 minutes at 37 °C, followed by the addition of another 0.01 U CIAP/pmol ends vector for a further 30 minutes incubation at 37 °C. The DNA was gel purified prior to ligation as CIAP may inhibit the ligation reaction.

2.2.14. DNA ligation

The vector and insert DNA are ligated at a vector: insert DNA ratio between 3:1 and 1:3. 1 U of T4 DNA Ligase (Boehringer Mannheim) was incubated in the presence of 1X DNA Ligase buffer (Boehringer Mannheim) in a total volume of 10 µl at 4/16 °C for 16 hours (overnight). To determine the extent of recombination, a control containing no insert in the reaction mix was also performed.

2.2.15. Preparation of competent cells

E. coli JM109 competent cells were available commercially from Promega. *E. coli* BL21(DE3) competent cells were prepared immediately before use using a standard calcium chloride method (Sambrook *et al.*, 1989). *E. coli* BL21(DE3) cells were streaked onto an LB agar plate from a glycerol stock. The plate was incubated overnight at 37 °C to allow the growth of individual colonies. One colony was used to inoculate a 5 ml culture of LB that was left to grow for 3 hours at 37 °C on a shaker at 225 rpm. The starter culture was transferred into 50 ml of LB and left to grow until the OD at 600 nm reached 0.4-0.6 AU. The cells were placed in a chilled polypropylene tube (Greiner), left on ice for 10 minutes then washed in 10 ml of ice-cold 0.1 M $CaCl_2$. The cells were harvested and resuspended in 2 ml 0.1 M $CaCl_2$. 200 µl aliquots of the competent cells were placed on ice and used immediately.

2.2.16. Transformation into *Escherichia coli*

E. coli JM109 (Promega) and *E. coli* BL21(DE3) competent cells were both transformed using a standard heat-shock method (Sambrook *et al.*,1989). 100-200 µl of competent cells were slowly thawed on ice. 1-50 ng of DNA was mixed into the cells and left on ice for 10 minutes. The cells were then heat-shocked in a water bath at 42 °C for 45-50 seconds then immediately put on ice for 2 minutes. Chilled SOC medium (see Section 2.5) was added to the reaction to make a final volume of 1 ml. Transformation efficiency of the competent cells was determined by using 1 µg of control supercoiled DNA. The cells were pelleted by centrifugation at 10,000 g for 3 minutes then re-suspended in 200 µl of LB and spread on an LB plate plus antibiotic.

2.2.17. Blue/white colony screening

Plasmids pUC18 and pGEM®-T both carry the β-galactosidase gene (*lacZ*) and are therefore capable of α-complementation. Blue colonies will form in the presence of X-gal if no DNA has been ligated into the cloning sites within this reading frame. Plates for blue/white selection contained 0.1 mM IPTG (Melford Laboratories Ltd.) and 40 μg/ml X-Gal (Roche Diagnostics Ltd.). Individual white colonies were selected from the plates.

2.2.18. Plasmid purification

Plasmids were purified from *E. coli* cells using a QIAprep Miniprep or Midiprep kit (Qiagen). The method suggested by the manufacturer was followed. The *E. coli* cells containing the plasmid were grown overnight in LB medium containing 100 μg/ml ampicillin at 37 °C shaking at 250 rpm. The cells were harvested and resuspended in alkaline lysis buffer (Qiagen). The lysate was neutralised and centrifuged to separate the plasmids from precipitated cellular debris, proteins and chromosomal DNA. The supernatant was applied to a QIA prep column and plasmids were selectively absorbed onto the membrane. The column was washed to remove RNA and contaminating proteins and the plasmids were eluted from the column with ddd water.

2.2.19. Cosmid DNA purification

To prepare cosmid DNA from a culture of bacteria, a protocol from Dr. Al Ivens, Sanger Centre, was followed. 1.5 ml of saturated culture of *E. coli* containing the relevant cosmid was centrifuged at 12,000 x g for 1 minute at room temperature. The pellet was resuspended in 200 μl of cold cosmid solution 1 (see Section 2.5), and

allowed to incubate at room temperature for 5 minutes. 500 µl of cosmid solution 2 (see Section 2.5) was added to the mix, the tube inverted 3 times and left to stand for 10 minutes. The 250 µl of cold cosmid solution 3 (see Section 2.5) was mixed into the suspension and incubated at –20 °C for 10 minutes. The vial was centrifuged at 12,000 x g for 10 minutes at room temperature. The supernatant was removed and 540 µl of propan-2-ol added. The pellet was allowed to precipitate on ice for 10 minutes and was centrifuged at 12,000 x g for 10 minutes at room temperature. The pellet was resuspended in 100 µl of water by incubating the tube at 65 °C for 10 minutes. 100 µl of ice-cold LiCl (5 M) was mixed into the solution and was incubated on ice for 10 minutes. The tube was centrifuged for 10 minutes at 12,000 g and the supernatant removed to a fresh tube. The cosmid DNA was precipitated in 2 volumes of 100 % ethanol at –20 °C for 10 minutes. The solution was centrifuged and the pellet washed in 70 % ethanol and then suspended in 50 µl TE containing RNase (10 µg/ml).

2.2.20. Restriction enzyme digestion

Restriction enzymes were bought from Promega and Boehringer Mannheim. The manufacturer recommended an appropriate buffer and incubation temperature for complete digestion of the DNA. The concentration of the restriction enzyme varied depending on the DNA template. A standard plasmid restriction digest was composed of 1-2 µg plasmid DNA, 5 units of restriction enzyme and 1 x enzyme buffer in a total volume of 30 µl. The reaction was incubated for 1 hour at 37 °C, or until complete digestion was obtained.

2.2.21. DNA Sequencing

Vectors containing the cloned PCR fragments were sent to MWG-BIOTECH (Milton Keynes, UK.) for double stranded confirmed sequencing. To analyse the data the double stranded sequence was confirmed by comparing chromatograms using Vector NTI version 6 (Informax, Inc.). The consensus sequence was put through a database search using the BLASTX algorithm (Altschul *et al.*, 1997) that compares a translated product of an unknown sequence against translated nucleic acid and protein databases. The databases used included the protein database SwissProt (Centre Medical Universitaire, Geneva, Switzerland), NCBI (National Centre for Biotechnology Information, Maryland, USA.), The Sanger Centre database (Cambridge, England).

2.2.22. Transfection of *Leishmania*

DNA was transfected into parasites for chromosomal or episomal integration using a method originally described by Cruz (Cruz *et al*, 1991). Approximately 4 x 10^7 *Leishmania mexicana* log-phase promastigotes were used for each transfection. Parasites were harvested at 1,000 g for 5 minutes, washed in 10 ml chilled transfection buffer and resuspended at 10^8 cells/ml of transfection buffer. 5 µg of DNA was placed in chilled cuvettes containing the parasites. The cells were transfected by electroporation using a Bio-Rad Gene Pulser set at 0.45 kV/cm and a capacitance of 500 µF. The cells were allowed to recover from the transfection in 10 ml complete HOMEM, for 24 hours at 25 °C. 1 ml of transfected cells was then placed into 9 ml of complete HOMEM containing the appropriate antibiotic. The remaining 9 ml was harvested and plated onto 1 % bactoagar plates containing complete HOMEM and antibiotic. Plates were incubated for approximately 10 days at 25 °C to obtain colonies.

The details of antibiotics used, including the concentration of drugs used to select for resistant parasites, are summarised in Table 2.3.

Antibiotic	Manufacturer	EC_{50} (μg/ml)	Drug concentration for selection (μg/ml)
Hygromycin B	CalBiochem	10	25-50
Nouseothricin	ClonAt	10	25-50
Neomycin	CalBiochem	10	10-25

Table 2.3:Details of antibiotics used to select for parasites containing resistance markers.

2.2.23. Southern Blotting

2.2.23.1. Genomic DNA preparation and digestion

Genomic DNA was isolated from the parasites mid-log and incubated with proteinase K (100 μg/ml) (Promega) and 5 mM EDTA (Fisons) for 40 minutes at 50 °C. This extra step degraded any proteins contaminating the nucleic acid sample thus allowing for complete digestion of the DNA. 5-10 μg of DNA was digested in a final volume of 100 μl, including 10 μl of appropriate buffer. 50 units of the restriction enzyme was added to the sample and incubated at 37 °C for 2 hours and then a further 50 units of enzyme was added for an overnight incubation at 37 °C. The extent of the digestion was analysed by agarose gel electrophoresis. Once the DNA was fully digested the sample was precipitated with pellet paint (Novagen), a dye-labelled carrier, and was resuspended in 20 μl of sterile water.

2.2.23.2. Gel preparation

A 0.8 % agarose gel was used to separate the digested DNA fragments. For reference, a 1 Kb DNA ladder (Promega) was run beside the samples. 100 pg of the template DNA complementary to the probe was run as a positive control to ensure complete transfer and hybridisation of the DNA. The gel was run at a low voltage (~50 volts) until the dye front reached the end of the gel. The gel was then stained with ethidium bromide and viewed under UV light. The agarose gel was incubated in depurination buffer for 10 minutes to ensure the larger DNA fragments on the gel were fully fragmented. The gel was then placed in denaturation buffer for 30 minutes to produce single stranded DNA, and finally washed in neutralisation buffer for 30 minutes. Between each step the gel was rinsed in ddd water.

2.2.23.3. Capillary blotting to transfer DNA onto nylon membrane

The digested DNA was transferred from the agarose gel to Hybond-N nylon membrane (Amersham Pharmacia Biotech). 20 x SSC transfer buffer was the reservoir and 3 x 3 mm filter paper acted as a wick to absorb the buffer through the agarose gel and nylon membrane. Paper towels were placed on top of the gel to allow the buffer to be absorbed through the gel, transferring the DNA onto the membrane (see Figure 2.2). Capillary blotting took place overnight at room temperature to ensure that all of the fragments were transferred. After blotting, the DNA was crosslinked to the nylon membrane by optimised UV crosslinking and stored at room temperature in Saran wrap.

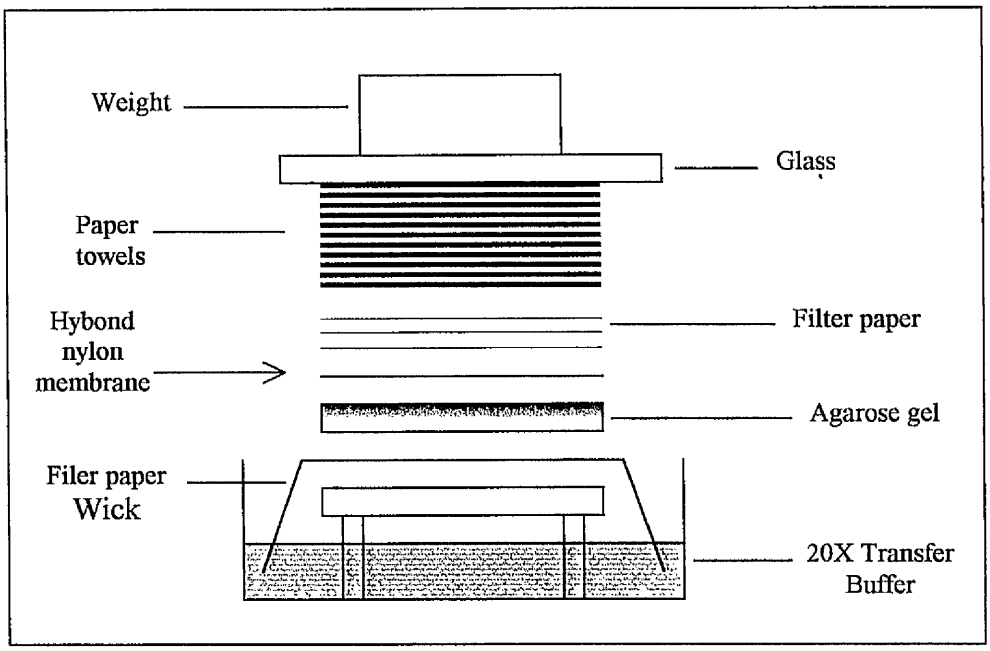


Figure 2.2: Capillary blotting to transfer digested DNA from an agarose gel to a nylon membrane for Southern blot analysis.

2.2.23.4. Hybridisation

The membrane was pre-hybridised in 20 ml Church-Gilbert's solution for 4 hours at 65 °C in a tube in a rotating hybridisation oven (Stratagene). The ^{32}P-labelled probe was made using Prime-It II Random Primer Labelling Kit (Stratagene). The protocol recommended by the manufacturer was followed. Random oligonucleotides anneal to multiple sites on the DNA template and this provides a substrate for the Klenow fragment to synthesise a complementary strand of DNA. Within the reaction dATP is substituted with radiolabelled dATP (α-^{32}P) (NEN Life Science Products) producing a radioactive template that can be used as a probe once denatured. The labelled probe was then added to 5 ml of Church-Gilberts solution and was placed in a fresh tube with the blot overnight at 65 °C. The blot was washed in 2 x SSC, 0.1 %

SDS at 65 °C for 10 minutes. A further wash of 1 x SSC, 0.1 % SDS at 65 °C for 2 x 10 minutes was done if required. The membrane was wrapped in saran wrap and then exposed to film at -70 °C.

2.3. Biochemical Methods

2.3.1. Enzymatic assays

2.3.1.1.Transketolase assay (1)

The transketolase (EC. 2.2.1.1) assay was based on the coupled assay described by Josephson and Fraenkel (1969). A schematic representation is shown in Figure 2.3. Xylulose 5-phosphate (X5P) was the donor substrate, with ribose 5-phosphate (R5P) being the acceptor molecule of the transferred 2-carbon glycoaldehyde unit. This reaction yields sedoheptulose 7-phosphate and glyceraldehyde 3-phosphate, the latter measured by the conversion to dihydroxyacetone phosphate by triosephosphate isomerase (TPI) (EC.5.3.1.1), which is reduced to glycerol 1-phosphate by glycerolphosphate dehydrogenase (GPDH) (EC.1.1.1.8) and NADH. The reaction mixture contained, per ml: Tris-HCl, pH 7.5 (100 mM) (Fisons) or glycylglycine (Sigma) pH 8.5 (50 mM), D-ribose 5-phosphate (10 mM) (Sigma), D-xylulose 5-phosphate (2 mM) (Sigma/preparation in laboratory, see Section 2.3.2), thiamine pyrophosphate (10 μM) (Sigma), $MgCl_2$ (1.2 mM) (Fisons), NADH (100μM) (Sigma), triosephosphate isomerase (1 unit/ml) (Boehringer Mannheim) and glycerol-3-phosphate dehydrogenase (1 unit/ml) (Boehringer Mannheim). Both coupling enzymes were desalted by gel filtration using Sephadex G-10 (Sigma) before use. NADH oxidation was monitored using a spectrophotometer (Kayak XA, Hewlett Packard) at 340 nm and the rate of decrease in absorbance was measured. The first absorbance

reading was taken following the addition of NADH into the assay mixture and the assay allowed to run until stable. The reaction was initiated by the addition of the enzyme/cellular fraction unless stated. The standard control lacks the substrate ribose 5-phosphate, thus measuring background cellular NADH oxidation already present in the cell. The linear rate was measured as the absorbance change per second (AU/sec). One unit is defined as the amount of enzyme that catalyses the formation of 1 μmol of glyceraldehyde 3-phosphate per minute. The specific activity is expressed as units per mg of protein (Units/mg).

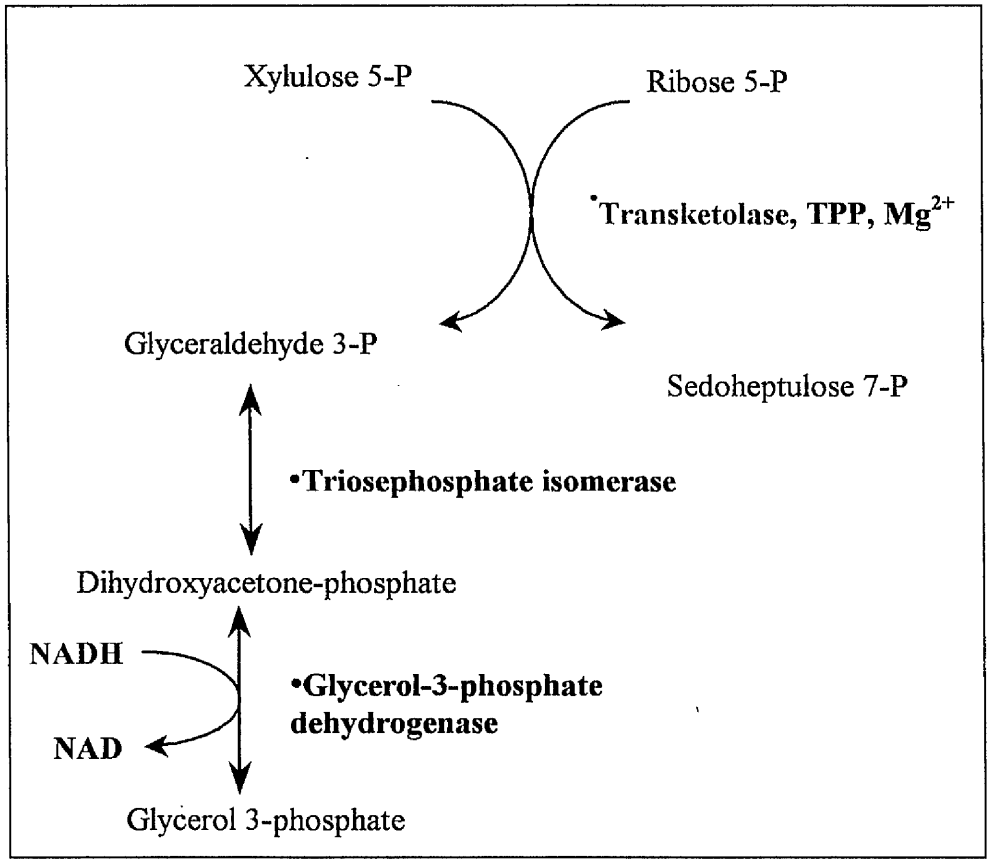


Figure 2.3: Schematic representation of the assay used to measure TKT activity using R5P and X5P as substrates.

2.3.1.2. Transketolase assay (2)

The second assay to measure transketolase activity was using hydroxypyruvate and erythrose 4-phosphate as the donor and acceptor substrates, respectively (Sprenger *et al.*, 1995) (Figure 2.4). The product of the reaction, fructose 6-phosphate, is converted to glucose 6-phosphate and then 6 phosphogluconolactonate by the coupling enzymes phosphoglucose isomerase (EC.5.3.1.9) and glucose-6-phosphate dehydrogenase (G6PDH) (EC.1.1.1.49). The G6PDH reaction is coupled to the reduction of NADP to NADPH that can be measured using a spectrophotometer at the absorbance 340 nm. The components of the assay in 1 ml were: glycylglycine buffer (0.5 M), pH 8.5, hydroxypyruvate (50 mM) (Sigma), erythrose 4-phosphate (1 mM) (Sigma), $MgCl_2$ (5mM) (Fisons), thiamine pyrophosphate (1 mM) (Sigma), NADP (100 μM) (Sigma), phosphoglucose isomerase (1 unit/ml) (Boehringer Mannheim) and glucose-6-phosphate dehydrogenase (1 unit/ml) (Boehringer Mannheim). The commercial erythrose 4-phosphate was contaminated with glucose 6-phosphate, therefore the reaction was run until this contamination was exhausted and no NADPH was formed before the transketolase reaction was initiated. The hydroxypyruvate was dissolved in 1 M glyclyglycine, pH 8.5, to allow the reaction to take place at pH 8.5.

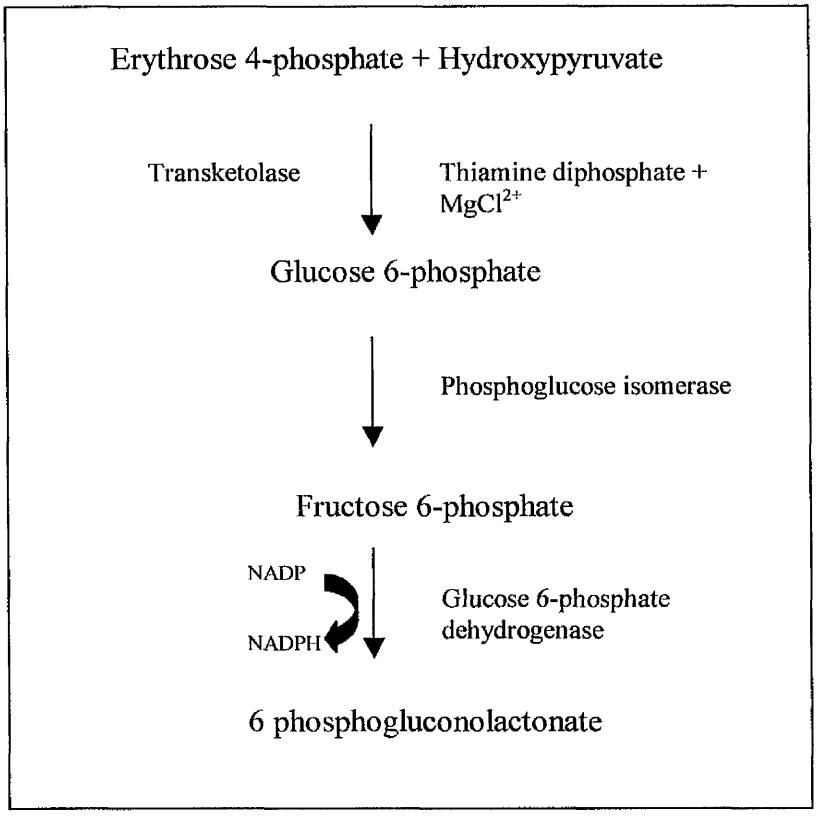


Figure 2.4. Schematic representation of the assay used to measure TKT activity using hydroxypyruvate and erythrose 4-phosphate as substrates.

2.3.1.3. Hexokinase assay

Hexokinase activity was measured by following NADPH production in a spectrophotometer at 340 nm (Martin *et al.*, 1976). The 1 ml reaction contained: Triethanolamine (0.2 M) pH 7.5, $MgCl_2$ (5 mM), NADP (0.5 mM), glucose (2 mM), ATP (2.5 mM), glucose-6-phosphate dehydrogenase (G6PDH) (1 unit/ml) and sucrose (0.25 M). The reaction was started with the addition of enzyme or parasite fraction.

The final volume of the reaction was 100 µl. Hexokinase was used as a marker of glycosomal fractions during subcellular fractionation experiments.

2.3.1.4. Pyruvate kinase assay

Pyruvate kinase (EC.2.7.1.40) activity was measured according to the method described by Bergmeyer (Bergmeyer, 1974). The reaction mix contained: Tris-Acetic acid buffer (0.2 M), pH 6.8, PEP (0.1 mM), $MgCl_2$ (0.5 mM), KCl (10 mM), ADP (0.01 mM), NADH (0.01 mM), lactate dehydrogenase (EC.1.1.1.27) (1 unit/ml) and sucrose (0.25 M). The reaction was initiated with the addition of parasite fraction. Pyruvate kinase catalyses the conversion of PEP to pyruvate in the presence of ADP. Lactate dehydrogenase will catalyse the production of lactate from pyruvate in the presence of NADH, the latter being measured using a spectrophotometer at the absorbance of 340 nm.

2.3.1.5. Fructose 1,6-bisphosphate assay

The presence of fructose 1, 6-bisphosphate was determined by using a coupled enzyme assay. Fructose-1, 6-bisphosphate aldolase (EC.4.1.2.13) catalyses the production of dihydroxyacetone phosphate and glyceraldehyde 3-phosphate from the substrate fructose 1, 6-bisphosphate. Dihydroxyacetone phosphate (DHAP) can be measured using the NADH requiring reaction with glycerol-3-phosphate dehydrogenase (EC.1.1.99.5) (G3PDH), making glycerol 3-phosphate.

2.3.2. Preparation of Xylulose 5-Phosphate

2.3.2.1. Xylulose 5-phosphate synthesis

It was necessary to synthesise xylulose 5-phosphate during this project, as Sigma discontinued this product. D-xylulose 5-phosphate (Na-salt) was synthesised according to the method devised by Zimmermann (Zimmermann *et al.*, 1999). The synthesis is described and illustrated in Figure 2.5. The preparation was initially scaled down to 20% of the suggested quantities, mainly due to the high concentration of transketolase required. The reaction took place over 48 hours in a flask containing: Fructose 1,6-bisphosphate (F1,6P) (BDH) (2.5 mmol), Li hydroxypyruvate (Sigma) (5.0 mmol), $MgCl_2(H_2O)_6$ (BDH) (0.38 mmol), thiamine diphosphate (Sigma) (0.19 mmol) and was adjusted to pH 7.0 with 01 M NaOH. The enzymatic reaction was started with the addition of: D-fructose-1,6-bisphosphate aldolase (Boehringer Mannheim) (FruA) (14.4 U), TPI (Boehringer Mannheim) (25 U), recombinant *E. coli* transketolase (TKA) (20 mg). After 6 hours, a further 7.2 U of FruA, 13 U of TPI and 5 mg of TKA were added, 8 hours into the reaction, additional TKA (5 mg) and Li hydroxypyruvate (3.0 mmol), and after 24 hours, 10 mg of TKA and Li-hydroxypyruvate (3.0 mmol) were added. Assaying fractions of the reaction for remaining DHAP monitored the reaction. To assay F1,6P and DHAP, 100 μM NADH (Sigma) and 1 U of G3PDH (Boehringer Mannheim) were mixed into a 1 ml cuvettes containing Tris pH 7.0 (100 mM). FruA was already present in the reaction mix being measured and therefore both substrates were assayed. 100 μl of the X5P reaction was mixed into the cuvette, and NADH oxidation was followed using a spectrophotometer (Kayak XA, Hewlett Packard) at 340 nm. During the reaction, an increase in pH occurred due to the consumption hydroxypyruvic acid. The pH was returned to pH 7.0 by 2.5 M HCl throughout the

reaction. The reaction was terminated after 48 hours with the addition of 4 g of Dowex AG 50W-X8, H^+ -form (BDH). The cation exchange resin was gently stirred for one hour at room temperature into the reaction mix and removed by filtering.

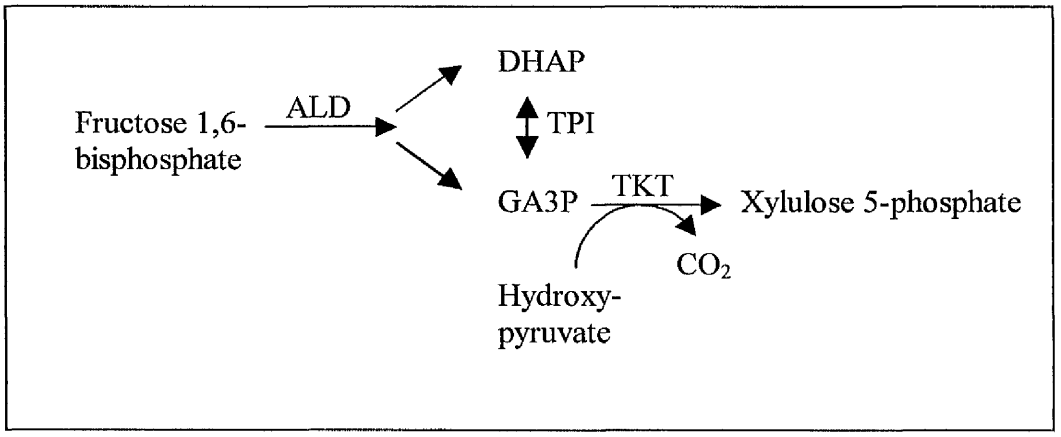


Figure 2.5: Xylulose 5-phosphate synthesis. Fructose 1,6-bisphosphate is converted to Dihydroxyacetone phosphate (DHAP) and glyceraldehyde 3-phosphate (GA3P) by fructose 1,6-bisphosphate aldolase (ALD). Triosephosphate isomerase (TPI) converts DHAP to GA3P allowing more GA3P to be produced. GA3P is converted by transketolase (TKT) to xylulose 5-phosphate. Hydroxypyruvate acts as the donor substrate releasing CO_2 thus making the reaction irreversible.

2.3.2.2. Purification of xylulose 5-phosphate

The sample containing X5P was purified using an anion exchange column made with Dowex AG 1-X8 strong anion exchange resin, Cl^- form (BDH). 35 g of resin were initially slurried in 200 ml of NaOH (1M) and poured into the column (QVF). The column was converted from Cl^- to formate form in two steps:

Step1. Cl- → OH-

Step 2. OH- → Formate.

Cl^- to OH^- ionic form required 22 bed volumes of 1 M NaOH (approximately 1.5 L) to be passed through the resin at a flow rate of 5 ml/min, using a HR Flow Inducer Pump, Watson-Marlow Ltd. To rinse the resin, two litres of ddd water was passed through the column at 5 ml/min. Ionic exchange was completed when the pH was less than 9.0. OH^- to formate form conversion required 2.5 bed volumes of 1 M formic acid (approximately 500 ml) to be passed through the column at 5 ml/min. This was rinsed with six litres of ddd water, until the pH was greater than 4.8. The change from OH^- to formate form was accompanied by a colour change from dark brown to yellow. The final column volume was 50 ml, as swelling and shrinking occurred while changing the ionic form. The column was stored and used subsequently at 4 °C for subsequent use.
The 100 ml sample was degassed and loaded onto the column at 2 ml/min. 1.3 M formic acid was used to elute the bound X5P at a rate of 4 ml/min. 2 x 100 ml fractions were initially collected, followed by 80 x 15 ml fractions using a fraction collector, over a period of 6 hours. Every fifth fraction was assayed for the presence of X5P using the transketolase assay previously described in Section 2.3.1.1. Fractions containing X5P were pooled and stores at 4 °C. The anion exchange column was re-equilibrated with eight litres of ddd water. One litre of Xu5P in formic acid was evaporated using a freeze-dryer. This was done in batches of 125 ml using a Flexi-Dry (FTS Systems, Inc.). The sample was placed into a round-bottomed flask and frozen in a thin layer around the flask using a dry ice/methanol slurry. This was attached to the freeze-dryer and left overnight to dry the sample. Residual substrate was resuspended in a small

volume of water and evaporated for a second time. The product appeared as a brown sticky material, thus was stored resuspended in Tris 100 mM, pH 7.0.

The purity and yield was determined by NMR (by Jim Gall, University of Glasgow). The presence of DHAP and hydroxypyruvate were determined enzymatically.

2.3.3. Protein concentration determination

The Bio-Rad protein assay, based on the Bradford method (Bradford, 1976) was used to determine the concentration of solubilised protein. This assay involved a differential colour change of an acidic dye in response to various concentrations of protein. The Bio-Rad dye was diluted in water 1:4 and filtered before use. Various dilutions of the protein sample were prepared and 10 µl placed into individual wells in a 96 well plate (Cellstar, Greiner Bio-one), including standards. 200 µl of the dye was mixed into each well and the absorbance was measured on a spectrophotometer (Titertek Multiskan® MCC/340) at 595 nm and compared to a standard of bovine serum albumin (BSA) (Sigma) (0.05-0.5mg/ml).

2.3.4. Polyacrylamide gel electrophoresis (PAGE)

SDS-PAGE was used to separate, visualise and analyse a variety of protein samples. The method used was originally described by Laemmli, 1970, and the protocol followed was described in Sambrook *et al.*, 1989. The resolving gel contained the following: 8-10 % bis/acrylamide (v/v) (Bio-Rad), 25 % resolving buffer (v/v) and 0.1 % SDS (v/v) (Fisons), 7.5 % ammonium persulphate (w/v) (Sigma) and 0.1 % TEMED (v/v) (Sigma). 10 ml were used to pour into two assembled mini-gel plates

(Bio-Rad), leaving space for the stacking gel, and was allowed to set. The stacking buffer (4 % bis/acrylamide (v/v), 2.5 % stacking buffer (v/v), 0.1 % SDS (v/v), 5 % ammonium persulphate (w/v) and 0.1 % TEMED (v/v)) was layered onto the resolving gel and a comb inserted. Once polymerisation had occurred, the wells were rinsed with ddd water and the gel was placed in 1 x reservoir buffer. 1 x sample buffer was mixed into the protein sample that was boiled for 5 minutes prior to loading. Benchmark™ Protein Ladder (GibcoBRL) was used as a molecular weight standard. The gel apparatus was run at a voltage of 15 V/cm until the bromophenol blue reached the bottom of the resolving gel. The gel was stained in coomassie blue stain for 1-2 hours and de-stained in de-stain buffer.

2.3.5.1. Expression of recombinant proteins

To clone and express the targeted proteins of interest in *E. coli*, the pET System (Studier and Moffatt, 1986) was chosen for this project due to the characteristics of the vectors. The targeted gene was cloned into unique restriction sites in a defined orientation. The vectors contained a T7 promoter and terminator, inducible by T7 RNA Polymerase, provided by a λDE3 lysogen. The chromosomal copy of T7 RNA Polymerase present in the host cell (BL21(DE3)) can be expressed by isopropyl-β-D-thiogalactopyranoside (IPTG), as it is under *lac*UV5 control, therefore selectively transcribing the targeted gene present in the pET vector. The chosen vector, pET-16b, contained an efficient ribosome binding site from phage T7 major capsid, as the gene to be cloned into the vector was eukaryotic. It contained a T7*lac* promoter that contained a *lac* operator sequence downstream of the T7 promoter and a natural *lacI* gene with independent promoter. The *lacI* gene represses natural T7RNA polymerase expression

that can occur in the absence of IPTG, thus preventing any non-inducible transcription. pET-16b was ampicillin resistant. The β-lactamase gene is in the same orientation as the T7 promoter therefore, expression of this gene could lead to degradation of the ampicillin in the presence of IPTG. Therefore a high level of ampicillin was present in all cultures during selection and induction. This vector contained an adjacent sequence encoding for an N-terminal 6-histidine tag for purification of the target protein.

2.3.5.2. Overexpression of recombinant transketolase protein

0.1 – 2 L of LB broth containing ampicillin (100 μg/ml) was inoculated with *E. coli* BL21(DE3) containing the appropriate expression plasmid. The culture was grown at 37 °C until the OD at 600 nm reached 0.6-0.8 AU, and 0.4 mM IPTG was mixed into the medium to induce expression of the recombinant protein. The flask was placed in a shaking incubator at 15 or 37 °C for 14-16 hours to induce protein expression.

2.3.6. Purification of protein preparation using Immobilised Metal Affinity Column

To purify the over-expressed *L. mexicana* and *E. coli* recombinant transketolase from *E. coli* crude lysates, immobilised metal affinity chromatography (IMAC) was used. The 6-histidine residues on the N-terminal of the protein bound to divalent cations (Ni^{2+}) immobilised to resin and could be eluted by imidazole that competes for binding to the column. A BioCAD 700E Workstation was used for the IMAC due to the speed at which the purification could be done (approximately one hour). Alan Scott, University of Glasgow, purified all protein preparations using a BioCAD 700E Workstation (PE Biosystems). The column used was a POROS MC 4.6 mmD/100

mmL that had a column size of 1.7 ml. The POROS beads were made of polystyrenedivinybenzene and were charged with nickel ions. The rapid purification was due to the design feature of the POROS beads (perfusion chromatography) (Figure 2.6). Two types of pores were present in these beads: large through-pores transect the bead and short diffuse pores branch from the larger pores. Increased speed of binding occurred due to the increased surface area available, the intra-particle flow and the increase of short diffusive pores in comparison to conventional chromatography.

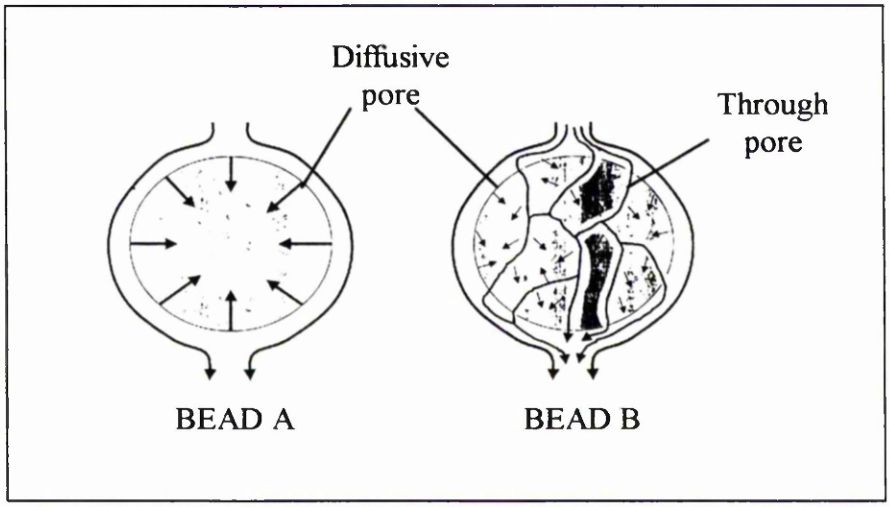


Figure 2.6: Diagram illustrating two bead types for chromatography. Bead A is used during conventional chromatography and bead B during perfusion chromatography (POROS beads). Beads illustrated by PE Biosystems.

2.3.6.1. Protein extract preparation

The induced culture was harvested by centrifugation at 5000 x g for 30 minutes at 4°C. The cells were lysed by sonication (Soniprep150, MSE) in lysis buffer 2 containing a proteinase inhibitor cocktail (4-(2-aminoethyl)benzenesulfonyl fluoride, pepstatin A, trans- epoxysuccinyl-L-leucylamido(4-guanidino) butane, bestatin,

leupeptin and aprotinin) (Sigma). The suspension was centrifuged at 12,000 x g for 30 minutes to isolate the soluble fraction. The sample was filtered using a 0.2 μm filter unit before use.

2.3.6.2. Purification method

The nickel-ion affinity column was initially regenerated with elution buffer (EB) and equilibrated with sonication buffer (SB) containing 0.5 mM imidazole. The sonicated extract was loaded onto the nickel-ion column at a flow rate of 5-10 ml/min. The column was washed with SB containing 0.5 mM imidazole and the flow through collected (approximately 18 ml). A second wash using SB containing 20 mM imidazole was used to remove tightly bound *E. coli* proteins from the column (approximately 15 ml). To elute the bound protein, a gradient of 20-500 mM imidazole in SB was used over 10 column volumes. 10 x 1.7 ml fractions were collected during the elution. SB + 500 mM imidazole was run through the column for a further 5 column volumes. Following purification, the column was stripped with 10 column volumes of strip buffer, and recharged. The absorbance at A_{280} was observed throughout the purification procedure. The success of the protein purification was analysed by the BioCAD trace and SDS-PAGE.

2.3.6.3. Dialysis

The collected fractions containing recombinant protein were pooled and dialysed overnight at 4 °C into 1 L 100 mM Tris-HCl (pH 7.0) to remove imidazole from the sample.

2.3.7. Western Blotting

Western blotting was used to detect transketolase protein expression in parasite crude lysates using an antibody obtained from the third bleed raised to the *Leishmania mexicana* recombinant transketolase.

2.3.7.1. Antibody production

Diagnostics Scotland raised antibody in rabbit to the purified recombinant *Leishmania mexicana* transketolase protein. 100 μg of recombinant transketolase was injected into a rabbit each month, for four months. Three bleeds were obtained.

2.3.7.2. Sample preparation

Leishmania cultures were prepared for western blotting as whole cell lysates. The parasites were harvested by centrifugation at 1,000 x g for 10 minutes and washed in iso-osmotic buffer (see Section 2.5) three times. Cells were re-suspended in 90 μl of lysis buffer (see Section 2.5) and 10 μl of proteinase cocktail inhibitors (Sigma) and allowed to lyse on ice for 30 minutes. The cells were centrifuged at 12,000 x g for 15 minutes and the supernatant used as soluble protein.

2.3.7.3. Dot blotting

To determine the optimal concentration of the primary and secondary antibody, dot blotting was initially performed on the first and second bleeds of the rabbit. Various protein concentrations (1- 0.125 μg/μl) were dotted onto ECL membrane (Enzyme-Linked Chemoilluminescence). Various concentrations of primary and secondary antibodies were incubated using the method below and optimal concentrations used in western blots.

2.3.7.4. Blotting

The protein samples were separated on a 0.8 % SDS-PAGE gel. The proteins were transferred onto Hybond™ ECL™ nitrocellulose membrane (Amersham Pharmacia Biotech) for 1 hour at 100 Volts in cold 1 x transfer buffer according to Sambrook *et al.*, 1989. Ponceau S dye (Sigma) is placed onto the membrane after transfer and the protein can be visualised. At this point, marks can be made on the membrane to orientate the gel. Non-specific binding sites were blocked by incubating the membrane in 0.2 % Tween-20 (v/v), 1 x TBS and 5 % milk (w/v) for 3 hours at room temperature on a slow shaker (100 rpm). The blot was incubated in diluted primary antibody (anti-transketolase) in 0.1 % Tween-20 (v/v), 1 x TBS and 1 % milk (w/v) overnight at 4 °C on a slow shaker. The blot was washed in 1 x TBS and 1 % milk (w/v) for 4 x 30 minutes at room temperature on a slow shaker. The secondary antibody (anti-rabbit) (Diagnostics Scotland) was diluted in 10 x TBS and 1 % milk (w/v) and incubated at room temperature for 2 hours. The blot was washed 3 x for 30 minutes in 1 x TBS and 1 % milk (w/v). Rinse the blot briefly in 10 x TBS.

2.3.7.5. Detection

An equal volume of Supersignal® West Pico Luminol/Enhancer Solution and Supersignal® West Pico Stable Peroxide Solution (Pierce) were mixed together and poured over the membrane. This was shaken for 2 minutes at room temperature and poured off. The blot was placed in saran wrap and put in a film cassette. In a dark room, Hyperfilm ECL was placed over the blot, the cassette closed and exposed for 15 – 60 seconds. The film was then removed and developed.

2.3.8. Subcellular fractionation

2.3.8.1. Parasite preparation and lysis

The parasites were harvested by centrifugation at 5,000 rpm for 10 mins. The cells were initially washed twice in isotonic buffer containing sucrose 0.25 M and KCl (5 mM). The cells were suspended in hypotonic buffer (see Section 2.5), the volume being determined by the following calculation.

$$\frac{\text{Total no. Parasites}}{1.2 \times 10^{9}} = \text{Volume of hypotonic buffer required}$$

The pellet was suspended in the required volume of hypotonic buffer and incubated on ice for 10 minutes. The parasites were monitored for swelling by microscopy. The cells were disrupted by passing them through a 25 gauge needle, immediately adding sucrose to the isotonic concentration of 0.25 M when the cells are broken. 500 µl of the total homogenate was stored at - 70 °C. 1 volume of isotonic buffer containing proteinase inhibitors was added to the suspension to allow for better separation of the

fractions. The homogenate was centrifuged at 5,000 rpm for 10 minutes at 4 °C and the supernatant removed.

2.3.8.2. Separation of fractions

The pellet was washed in 10 ml of isotonic buffer and the centrifugation repeated. The pellet was stored at – 70 °C and was the nuclear fraction. The supernatant was centrifuged at 15,000 rpm for 30 minutes at 4 ° C, the pellet washed in isotonic buffer and the step repeated. The supernatant from these washes was used to prepare the cytosolic and microsomal fractions. The pellet containing the large and small fractions was re-suspended in 800 µl of isotonic buffer and homogenised with a 25 gauge needle. The extent of the disruption was monitored under a microscope, and used when all granules were liberated from the membranes. This granular fraction was separated using a sucrose gradient.

2.3.8.3. Sucrose gradient

To prepare the sucrose gradient, various concentrations of sucrose were prepared in Tris-HCl (25-50 mM) and EDTA (1 mM). The gradient ranged from 2.5 M to 0.5 M sucrose, the various concentrations being carefully layered in a plastic tube. The gradient was placed at 4 °C for 16 hours before use. The sample was loaded onto the gradient, and then completed with buffer containing no sucrose. The gradient was then sealed with a heat gun and centrifuged at 300,000 x g for 3 hours at 4 °C (60,000 rpm). 12 x 1 ml fractions were collected from the sucrose gradient using a fraction collector and a peristaltic pump.

2.3.8.4. Cytosolic and microsomal fractions

To obtain the cytosolic fraction, the supernatant was centrifuged at 45,000 rpm for 1 hour. The pellet was then washed in iso-osmotic buffer and the spin repeated. The remaining pellet was the microsomal fraction. The cytosolic fraction was confirmed by repeating the centrifugation with the supernatant collected.

2.3.9. Indirect immunoflourescence

Cover slips were treated with 3 % 3-aminopropyltrethloxyl saline in acetone to allow the parasites to bind to the cover slip. 5-20 x 10^6 parasites / ml were washed in 1 ml of PBS, placed onto the treated cover slip and allowed to incubate for 10 minutes at room temperature. The parasites were removed and the slide covered with either, a paraformaldehyde fixative and left to incubate for 10 minutes or a methanol fixative that was allowed to incubate at -20 °C for 15 minutes. This was prepared by making a 4 % solution of paraformaldehyde in water at 60 °C, then adding 0.3 volumes of 3 x PBS and normalising the pH to 7.2. The fixative was filtered before use and cooled to room temperature. The fixative was removed from the slide and the slides checked for fixed parsites. The cells were permeabilised by incubating the parasites in blocking buffer (II) (see Section 2.5) for 30 minutes in a humid environment. The slips were then incubated in blocking buffer (II) containing the primary antibody diluted to an appropriate concentration for one hour at room temperature. The slides were then washed in PBS, and incubated in flouresceine isothiocyanate conjugated to rabbit immunoglobulin for one hour in the dark. The cells were washed in PBS, the cover slips mounted onto slides using ProLong Antifade Kit (Molecular Probes) and incubated in the dark for 1 hour. The slides were viewed using a Nikon E600 epifluorescence

microscope and photographs taken. The slides could also be used for confocal microscopy.

2.4. Sequence analysis tools

The analysis of nucleic acid and amino acid sequences during this study was done using the following web sites:

NCBI: http://www.ncbi.nlm.nih.gov

EMBL: http://www.ebi.ac.uk

BLAST: http://blast.wustl.ed

PSORT: http://psort.nibb.ac.jp/

CyanoBase: http://www.kazusa.orjp/cyano.html

PlasmoDB: http://PlasmoDB.org/

ProtParam Tool, ExPASy: http://ca.expasy.org/cgi-bin/protparam

The protein alignments and the phylogenetic trees created during this study were made using the following computer programmes:

Clustal W, Align X, Vector Nti (Thompson *et al.*, 1994)

Clustal X (1.5b) (Thompson *et al.*, 1997; Saitou and Nei, 1987))

Tree View Version 1.6.5. (Page, 1996)

2.5. Buffers and Reagents

Blocking buffer: 0.2 % Tween 20 (v/v) and 1 x TBS

Blocking buffer (II): PBS, 0.1 % saponin (v/v), 1 % BSA (w/v), 2 % swine normal serum(v/v)

Church-Gilberts solution: 340 mM Na_2HPO_4, 158 mM NaH_2PO_4, 240 mM SDS, 1 mM EDTA pH 8.0

Coomassie blue stain: 0.5 g/1 L Coomassie brilliant blue stain, 10 % acetic acid (v/v), 12.5 % isopropanol (v/v)

Cosmid solution 1: 25 % (w/v) glucose, 50 mM Tris-HCl (pH 8.0), 10 mM EDTA

Cosmid solution 2: 0.2 M NaOH, 1% SDS (w/v)

Cosmid solution 3: 3 M potassium acetate (pH 4.8) with acetic acid

Crude extract lysis buffer: 25 mM Tris-HCl, 1 % Triton X-100 (v/v), 0.25 M Sucrose, 0.1 mM EDTA and proteinase inhibitors cocktail

Denaturation buffer: 1.5 M NaCl, 0.5 M NaOH

Depurination buffer: 0.125 M HCl

De-stain solution: 10 % acetic acid (v/v), 12.5 % methanol (v/v)

Elution buffer (EB): 50 mM NaP, 300 mM NaCl, 500 mM imidazole, pH 8.0

HOMEM: MEM salts, 10 ml 50 x /L MEM amino acids, 10 ml 50 x /L MEM non essential amino acids, 5 mM glucose, 3.6 mM sodium bicarbonate, 1 % sodium pyruvate, 8 μM p-amino benzoic acid, 0.4 μM biotin, 20 mM HEPES, pH 7.0

Hypotonic buffer: 25 mM Tris-HCl, 1 mM EDTA, pH 8

Iso-osmotic buffer: 25 mM Tris-HCl, 1 mM EDTA, pH 8, 0.25 M Sucrose

LB Medium (Luria-Bertani Medium): 1 % Bacto-tryptone (BBL) (w/v), 0.5 % Bacto-yeast extract (Oxoid) (w/v), 170 mM NaCl, pH 7.0

LB Medium plates: LB Medium, 1.5 % Bacto-agar (Oxoid) (w/v), pH 7.0

Lysis buffer 2: 0.5 M NaCl, 0.1 M Tris-HCl, pH 8.0

Monophasic medium: 3.3 % Brain heart infusion, 0.3 % Tryptose, 23 mM PO_4HNa_2, 5.4 mM KCl, 1.5 mM Glucose.

Neutralisation buffer: 1.5 M NaCl, 0.5 M Trizma base, pH 7.5

PBS: 20 mM sodium phosphate buffer, 150 mM NaCl

Ponceau S: 3-hydroxy-4- (2-sulfo-4[4-sulfophenylazol]-phenyl-azo)-2,7-naphthalenedisulfonic acid.

Recharge solution: 0.1 M Nickel sulphate

Resevoir buffer (10x): .25 M Tris, 1.9 M Glycine, 28.8 M SDS

Resolving buffer: 1.5 M Tris-HCl, pH 8.8

Sample buffer (2x): 86.2 mM Tris-HCl, pH 6.8, 14% glycerol (v/v), 2.75 % SDS (w/v), 0.1 % bromophenol blue (v/v), 100 mM DTT

SOC medium: 2 % Bacto-tryptone (w/v), 0.5 % Bacto-yeast extract (w/v), 8.5 mM NaCl, 2.4 mM KCl, 100 mM $MgCl_2$, 20 mM Glucose, pH 7.0

Sonication buffer (SB): 50 mM NaP, 300 mM NaCl, pH 8.0

SSC buffer (x20): 3M NaCl, 0.33 M Tri-sodium citrate

Stacking buffer: 0.5 M Tris-HCl, pH 6.8

Strip solution: 50 mM EDTA, 1 M NaCl, pH 8.0

TBS (x10): 0.2 M Tris and 137 mM NaCl, pH 7.6

TE: 10 mM Tris-HCl and 1 mM EDTA, pH8.0

TELT lysis buffer: 50 mM Tris-HCl pH 8.0, 62.5 mM EDTA pH 9.0, 2.5 M LiCl, 4% Triton X-100 (v/v)

Transfection buffer: 21 mM HEPES pH 7.5, 137 mM NaCl, 5 mM KCl, 0.7 mM Na_2HPO_4, 5 mM Glucose

Transfer buffer (x10): 0.2 M Tris and 1.5 M Glycine

Transfer buffer (x1): 1/10 10x transfer buffer and 20 % methanol (v/v) (store at 4 °C 2 hours before use)

Chapter III

Cloning of transketolase genes from *Escherichia coli* and *Leishmania mexicana*

3.1. Introduction

The transketolase gene (*TKT*) from *Leishmania mexicana* was cloned and sequenced as the first step in this project. As *TKT* sequences were not available from the *Leishmania major* or *Trypanosoma brucei* sequencing databases at the outset of the study, several different molecular techniques had to be used to obtain part of the gene sequence. There have been several transketolase gene sequences published from various organisms, including mammals, plants, yeast and bacteria. The transketolase amino acid sequences have a high level of similarity and two consensus motifs. Transketolase is a thiamine pyrophosphate (TPP) dependent enzyme, and shares features common to all TPP-dependent enzymes, including pyruvate decarboxylase and the E1 component in the pyruvate dehydrogenase complex (Hawkins *et al.*, 1989). A TPP-binding consensus motif was identified in transketolase and other TPP-dependent enzymes (Schenk *et al.*, 1997). The publication of yeast transketolase crystal structure showed that this motif does bind TPP (Lindqvist *et al.*, 1992). The central domain of the transketolase sequence contains a well conserved consensus motif, named the transketolase motif (Schenk *et al.*, 1997). These conserved sequences were useful in cloning a large section of the *TKT* gene sequence from *L. mexicana.* Preliminary DNA sequence from the *L. major* database similar to *TKT* emerged in August 2000 and led to the completion of the sequence of the gene and the protein that it encodes from *L. mexicana.* The sequencing

of the *TKT* gene revealed the predicted protein to contain the conserved thiamine pyrophosphate (TPP) and transketolase (TKT) motifs, common to all transketolase proteins. A peroxisome-targeting signal, SKM, was present at the C-terminus of the protein, suggesting that some TKT may be localised to the glycosome. This was further investigated with localisation experiments (Chapter V). The untranslated region (UTR) of the gene was also sequenced. With this information, constructs were produced for use in experiments to try to replace the two alleles of *TKT* with antibiotic selectable markers.

3.2.1. Cloning of the transketolase A gene from *Escherichia coli*

A source of transketolase was required during this project for development of the assay system and other downstream applications. The *E. coli* transketolase A (*tktA*) gene sequence and translated amino acid sequence were previously determined (Schenk *et al.*, 1993). The availability of this information allowed for the cloning of the *tktA* gene into a vector and the expression of the recombinant protein. The in-house production of *E. coli tktA* represents a cost effective method of making a large quantity of protein that was useful during this project. Oligonucleotides were designed to amplify the 2 kb gene from *E. coli* genomic DNA and the primers included restriction sites for convenient cloning into the expression vector pET-16b that contained an N-terminal His-tag (Table 3.1).

Primer name	Primer sequence 5'→ 3'	Restriction site
pEcolitk1	TGG AGT CCA TAT GTC CTC ACG	*Nde*I
pEcolitk2	ACC CGG ATC CCT AAT TAC AGC AGT TCT	*Bam*HI

Table 3.1: Oligonucleotides used to amplify the *E. coli tktA* gene. The restriction sites for *Nde*I and *Bam*HI (highlighted in blue) were incorporated into the primer sequence, so the PCR product could be directionally cloned into the expression vector pET-16b.

To amplify *tktA* from *E. coli* the following PCR cycle was as used:

Pre-heat reaction:	60 seconds - 95 °C	
Melt and denature:	30 seconds - 95 °C	} x 30 cycles
Anneal:	30 seconds - 48 °C	
Extension:	60 seconds - 72 °C	
Final hold:	5 minutes - 72 °C	

The PCR reaction was analysed on a 1 % agarose gel that revealed that a 2 kb product had been amplified (Figure 3.1). This product was eluted from the agarose gel and digested with the restriction enzymes *Nde*I and *Bam*HI at 37 °C. The vector pET-16b (Novagen) was also digested with the same restriction enzymes and dephosphorylated to prevent self-ligation of partially digested plasmids. Both digested products were gel eluted to remove uncut supercoiled plasmid, and ligated overnight at 16 °C. The resulting plasmids were 7.8 kb in size, which indicated the ligation of plasmid and insert had been successful. The plasmid was transformed into competent *E. coli* BL21(DE3) cells that are capable of over-expressing the recombinant N-terminal 6 histidine-tagged *E. coli* TKA. A diagrammatic representation of the plasmid containing *E. coli tktA* (pET16bEctkt) can be seen in Figure 3.2.

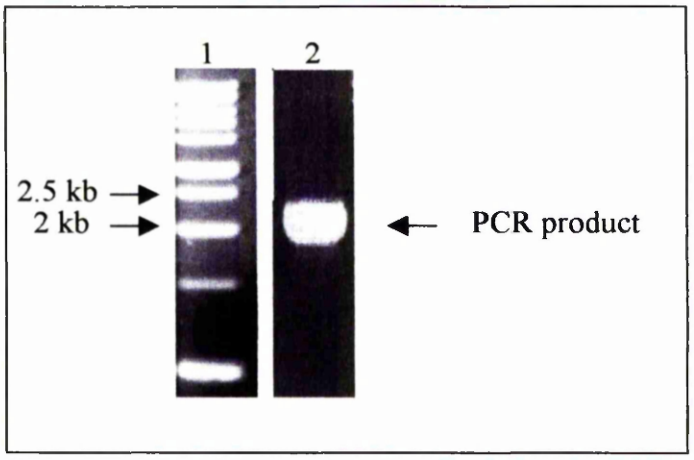


Figure 3.1: The PCR product obtained when amplifying *E. coli tktA*. The PCR product (Lane 2) is approximately 2 kb in size according to the 1 kb ladder (Lane 1).

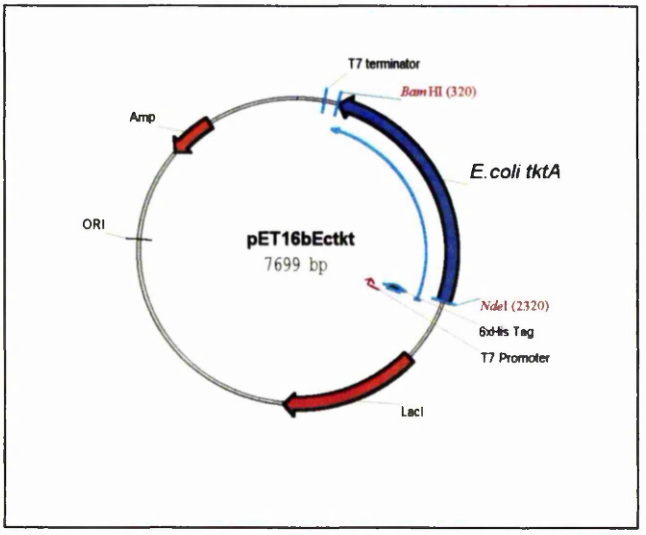


Figure 3.2: Diagrammatic representation of plasmid pET-16b containing the *E. coli tktA* gene. The ORF is represented with a blue line indicating the direction of the transcription of the *tktA* gene from pET-16b. The location of the T7 promoter and terminator, and the origin of replication (ORI) are noted on the plasmid. The histidine-tag (6x His Tag) resides upstream of the *Nde*I cloning site (blue block arrow inside plasmid), and is expressed on the N-terminal of the protein. The genes encoding for ampicillin resistance (Amp) and the *LacI* repressor gene (LacI) are both represented as red block arrows in the direction of translation.

3.2.2. Sequencing of the *Escherichia coli tktA* gene

To sequence the 2 kb *tktA* gene, the plasmid DNA was sent to MWG-BIOTECH for double-stranded sequencing. Various internal primers were designed to anneal to the sequence, as the length of the sequencing reads were approximately 800 bp. The chromatograms were analysed on Vector NTI™ V5.5, InforMax®, by aligning the various sequences as contigs (Figure 3.3), and a consensus sequence was produced.

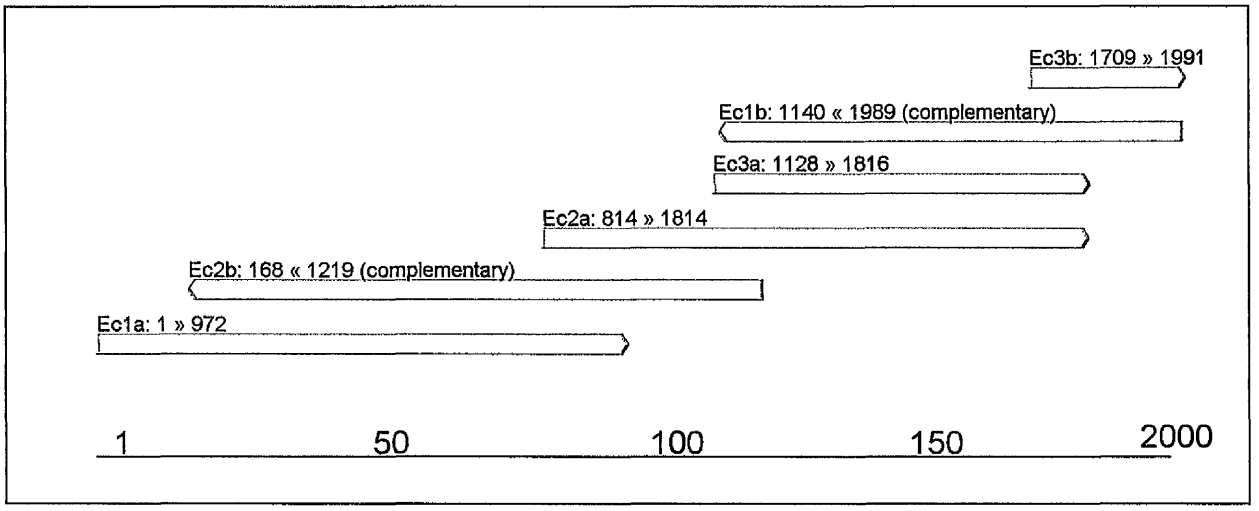


Figure 3.3: Illustration of DNA contigs aligned to obtain a consensus sequence of the cloned *E. coli tktA*. The 6 contigs of sequence obtained from MWG-BIOTECH were aligned using Contig Express (Vector NTI) to create a confirmed consensus DNA sequence. The numbers on the single line below the contigs represent the bp position of the transketolase gene.

The 1992 bp DNA sequence was translated using Vector NTI to produce an ORF of 664 amino acids. The predicted amino acid sequence of the cloned gene is shown in Figure 3.4, aligned with the published *E. coli tktA* ORF. The alignments were performed with AlignX (Vector NTI) that uses the Clustal W algorithm (Thompson *et al.*, 1994). The cloned and published sequences of *E. coli tkt* are 98.6 % identical. Six amino acid differences at positions 103, 105, 107, 108, 588 and 634, can be seen when comparing these sequences (Table 3.2). All of the changes in amino acids occur in non-conserved

areas of the sequence. These changes may have arisen due to PCR or sequencing error, as Taq was used to amplify the gene. The TKA his-tagged protein was over-expressed in *E. coli* as soluble active enzyme, and was used as a positive control for the assay during this project. The *E. coli* TKA was used in two biotransformations, (1) the production of xylulose 5-phosphate and (2) the synthesis of 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP) by Dr. Kirsten Fleming. Due to the fact that the enzyme was active and we wished to use it only in biotransformations and as a control to test that the assays were working, no further attention was paid to the discrepancies in sequence between our sequence and the published one.

```
                    1                                                  50
EcoliTKpub      (1) MSSRKELANAIRALSMDAVQKAKSGHPGAPMGMADIAEVLWRDFLKHNPQ
EcoliTKaa       (1) MSSRKELANAIRALSMDAVQKAKSGHPGAPMGMADIAEVLWRDFLKHNPQ
Consensus       (1) MSSRKELANAIRALSMDAVQKAKSGHPGAPMGMADIAEVLWRDFLKHNPQ
                    51                                                 100
EcoliTKpub     (51) NPSWADRDRFVLSNGHGSMLIYSLLHLTGYDLPMEELKNFRQLHSKTPGH
EcoliTKaa      (51) NPSWADRDRFVLSNGHGSMLIYSLLHLTGYDLPMEELKNFRQLHSKTPGH
Consensus      (51) NPSWADRDRFVLSNGHGSMLIYSLLHLTGYDLPMEELKNFRQLHSKTPGH
                    101                                                150
EcoliTKpub    (101) PEVGYT-AGVETTTGPLGQGIANAVGMAIAEKTLAAQFNRPGHDIVDHYT
EcoliTKaa     (101) PESGVTPLGVETTTGPLGQGIANAVGMAIAEKTLAAQFNRPGHDIVDHYT
Consensus     (101) PE G T  GVETTTGPLGQGIANAVGMAIAEKTLAAQFNRPGHDIVDHYT
                    151                                                200
EcoliTKpub    (150) YAFMGDGCMMEGISHEVCSLAGTLKLGKLIAFYDDNGISIDGHVEGWFTD
EcoliTKaa     (151) YAFMGDGCMMEGISHEVCSLAGTLKLGKLIAFYDDNGISIDGHVEGWFTD
Consensus     (151) YAFMGDGCMMEGISHEVCSLAGTLKLGKLIAFYDDNGISIDGHVEGWFTD
                    201                                                250
EcoliTKpub    (200) DTAMRFEAYGWHVIRDIDGHDAASIKRAVEEARAVTDKPSLLMCKTIIGF
EcoliTKaa     (201) DTAMRFEAYGWHVIRDIDGHDAASIKRAVEEARAVTDKPSLLMCKTIIGF
Consensus     (201) DTAMRFEAYGWHVIRDIDGHDAASIKRAVEEARAVTDKPSLLMCKTIIGF
                    251                                                300
EcoliTKpub    (250) GSPNKAGTHDSHGAPLGDAEIALTREQLGWKYAPFEIPSEIYAQWDAKEA
EcoliTKaa     (251) GSPNKAGTHDSHGAPLGDAEIALTREQLGWKYAPFEIPSEIYAQWDAKEA
Consensus     (251) GSPNKAGTHDSHGAPLGDAEIALTREQLGWKYAPFEIPSEIYAQWDAKEA
                    301                                                350
EcoliTKpub    (300) GQAKESAWNEKFAAYAKAYPQEAAEFTRRMKGEMPSDFDAKAKEFIAKLQ
EcoliTKaa     (301) GQAKESAWNEKFAAYAKAYPQEAAEFTRRMKGEMPSDFDAKAKEFIAKLQ
Consensus     (301) GQAKESAWNEKFAAYAKAYPQEAAEFTRRMKGEMPSDFDAKAKEFIAKLQ
                    351                                                400
EcoliTKpub    (350) ANPAKIASRKASQNAIEAFGPLLPEFLGGSADLAPSNLTLWSGSKAINED
EcoliTKaa     (351) ANPAKIASRKASQNAIEAFGPLLPEFLGGSADLAPSNLTLWSGSKAINED
Consensus     (351) ANPAKIASRKASQNAIEAFGPLLPEFLGGSADLAPSNLTLWSGSKAINED
                    401                                                450
EcoliTKpub    (400) AAGNYIHYGVREFGMTAIANGISLHGGFLPYTSTFLMFVEYARNAVRMAA
EcoliTKaa     (401) AAGNYIHYGVREFGMTAIANGISLHGGFLPYTSTFLMFVEYARNAVRMAA
Consensus     (401) AAGNYIHYGVREFGMTAIANGISLHGGFLPYTSTFLMFVEYARNAVRMAA
                    451                                                500
EcoliTKpub    (450) LMKQRQVMVYTHDSIGLGEDGPTHQPVEQVASLRVTPNMSTWRPCDQVES
EcoliTKaa     (451) LMKQRQVMVYTHDSIGLGEDGPTHQPVEQVASLRVTPNMSTWRPCDQVES
Consensus     (451) LMKQRQVMVYTHDSIGLGEDGPTHQPVEQVASLRVTPNMSTWRPCDQVES
                    501                                                550
EcoliTKpub    (500) AVAWKYGVERQDGPTALILSRQNLAQQERTEEQLANIARGGYVLKDCAGQ
EcoliTKaa     (501) AVAWKYGVERQDGPTALILSRQNLAQQERTEEQLANIARGGYVLKDCAGQ
Consensus     (501) AVAWKYGVERQDGPTALILSRQNLAQQERTEEQLANIARGGYVLKDCAGQ
                    551                                                600
EcoliTKpub    (550) PELIFIATGSEVELAVAAYEKLTAEGVKARVVSMPSTDAFDKQDAAYRES
EcoliTKaa     (551) PELIFIATGSEVELAVAAYEKLTAEGVKARVVSMSSTDAFDKQDAAYRES
Consensus     (551) PELIFIATGSEVELAVAAYEKLTAEGVKARVVSM STDAFDKQDAAYRES
                    601                                                650
EcoliTKpub    (600) VLPKAVTARVAVEAGIADYWYKYVGLNGAIVGMSTFGESAPAELLFEEFG
EcoliTKaa     (601) VLPKAVTARVAVEAGIADYWYKYVGLNGAIVGMTTFGESAPAELLFEEFG
Consensus     (601) VLPKAVTARVAVEAGIADYWYKYVGLNGAIVGMSTFGESAPAELLFEEFG
                    651        664
EcoliTKpub    (650) FTVDNVVAKAKELL
EcoliTKaa     (651) FTVDNVVAKAKELL
Consensus     (651) FTVDNVVAKAKELL
```

Figure 3.4: Alignment of the cloned compared with published *E. coli* TKA protein sequence. The alignment was done using Align X (Vector NTI). The published sequence is indicated by EcoliTKpub and the cloned sequence is indicated by the name EcoliTKaa. The consensus sequence is illustrated in black, with blue letters indicating a block of similar residues. The green amino acids represent a non-similar alignment of the residues.

Location of amino acid change (*E. coli* TKA)	**DNA change Published→ cloned**	**Amino acid change Published→ cloned**
103	GTG → AGT	V → S
105	TAC → GTT	Y → V
107	GAP → CCG	GAP* → P
108	GCT → CTG	A → L
588	ACC → AGC	P → S
634	AGC → ACC	S → T

Table 3.2: Location of the amino acid changes between the published *E. coli* TKA and the cloned *E. coli* TKA. The location refers to the amino acid number of the aligned sequences in Figure 3.4. The change of DNA and resulting amino acid change are shown comparing cloned to the published. The GAP* represents a gap in the sequence when aligned, an insertion of the amino acid P was present.

3.3. Cloning of the transketolase gene from *Leishmania mexicana*

3.3.1. Cloning of a partial *TKT* gene sequence using conserved motifs

To identify a partial *TKT* gene sequence from *L. mexicana*, two conserved domains from TKT amino acid sequences were used as a template to design oligonucleotides to amplify the gene using PCR. These conserved motifs were initially recognised by Schenk (Schenk *et al.*, 1997) following the alignment of 20 transketolase amino acid sequences. The two most conserved regions are a thiamine pyrophosphate motif, common to all thiamine pyrophosphate dependent enzymes, and a transketolase motif (Figure 3.5).

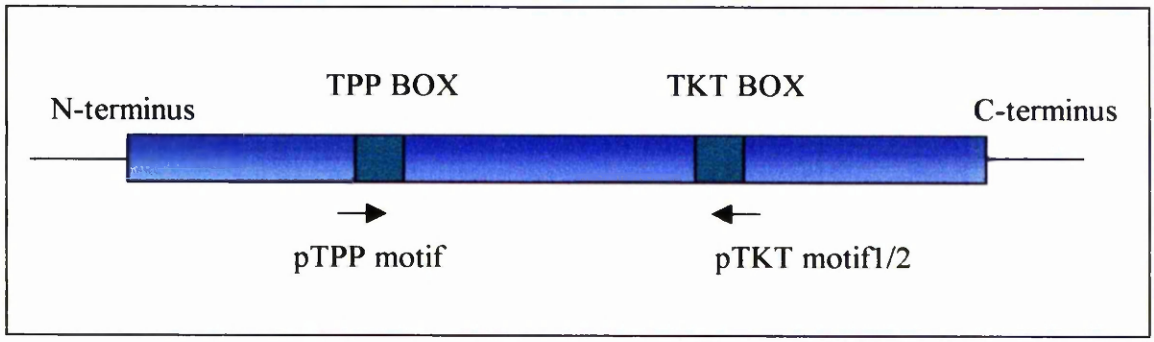


Figure 3.5: Schematic representation of the *L. mexicana* TKT protein highlighting two conserved areas. Transketolase contains a thiamine pyrophosphate motif (TPP BOX) and a transketolase specific motif (TKT BOX), both of which are highly conserved. These motifs were used to design primers to amplify a 1 kb region of the *TKT* gene from *L. mexicana*. The direction in which the primers, pTPP motif and pTKT motif1/2, amplified this region are arrowed.

The conserved amino acids used to design these primers are highlighted in Figure 3.6. The degenerate primers pTPP motif, pTKT motif 1 and 2, were designed to amplify 1 kb of DNA sequence. The sequence of these primers can be seen below in Table 3.3.

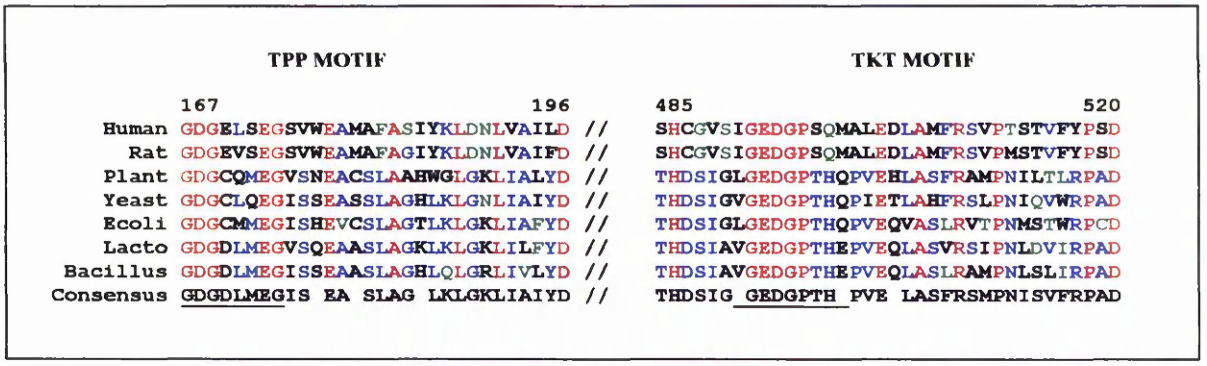


Figure 3.6: Alignment of transketolase amino acid sequences comparing the TPP and TKT motifs. Align X (Vector NTI) was used to create the alignment of transketolases belonging to *Homo sapiens* (Human), *Rattus norvegicus* (Rat), *Craterostigma plantagineum*, tkt7 (Plant), *Saccharomyces cerevisiae* TKT1 (Yeast), *Escherichia coli* TKA (Ecoli), *Lactococcus lactis* (Lacto) and *Bacillus subtilis* (Bacillus). Identical amino acid residues are in red, non-similar residues are in black (except in the consensus), conservative residues are blue and weakly similar residues are in green. The sequence underlined indicates the conserved regions used as a template in the primer design to amplify the *L. mexicana TKT* gene.

Primer name	**Degenerate primer sequence 5' → 3'** **Consensus amino acid sequence**
pTPP motif	GGN GAY GGN TGY YWN ATG GAR GG G D G C L M E G Q
pTKT motif 1	TG RTG NGT NGG NCC RTC YTC NCC Q H T P G D E G
pTKT motif 2	TG RTG NGT NGG NCC RTC YTC NCC NAG Q H T P G D E G L

Table 3.3: Degenerate primers used to amplify a 1 kb region of the *L. mexicana TKT* gene. The primers were designed to complement the TPP motif and the TKT box, which are both conserved amino acid sequences in transketolase. The consensus amino acid sequence is highlighted in blue and corresponds to the degenerate primer sequence (N = ATGC, Y = TC, W = TA, R = AG).

PCR was used to amplify the 1 kb of *TKT* DNA between these conserved motifs from 100 ng of *L. mexicana* genomic DNA. The PCR cycle used was as follows:

Pre-heat reaction:	60 seconds - 95 °C	
Melt and denature:	30 seconds - 95 °C	x 25 cycles
Anneal:	30 seconds - 52 °C	
Extension:	60 seconds - 72 °C	
Final hold:	10 minutes - 72 °C	

Initially a single faint product was amplified and this band was eluted from the gel, used as template DNA in a PCR and two bands were subsequently amplified (Figure 3.7). Both bands A and B were cloned into vector pUC18 using blunt-ended ligation, and sequenced. The DNA sequence amplified from the plasmid-containing band A (pUCTP/TK) (Figure 3.8) was similar to other transketolase sequences (Figure 3.9). Band B had low homology to various hypothetical proteins in the SWISS-PROT database. This was determined by database searching for sequence similarity to other proteins using the computer programme BLAST (Basic Local Alignment Search Tool) (Altschul *et al.*, 1997). More specifically, BLASTX was used that compared a DNA sequence translated into 6 possible reading frames, with a protein sequence database. The E value is the number of matches expected to be found by chance, therefore, the lower the value, the more stringent the match. The score relates to the number of high-scoring segment pairs, thus the higher the values, the more stringent the search. The transketolase sequence from *Leishmania* had the highest similarity to a transketolase from *Saccharomyces cerevisiae* and *Pichia stipitis*. This 1 kb *TKT* DNA sequence was used to design gene specific primers to amplify the remaining 5' and 3' regions of the *TKT* gene.

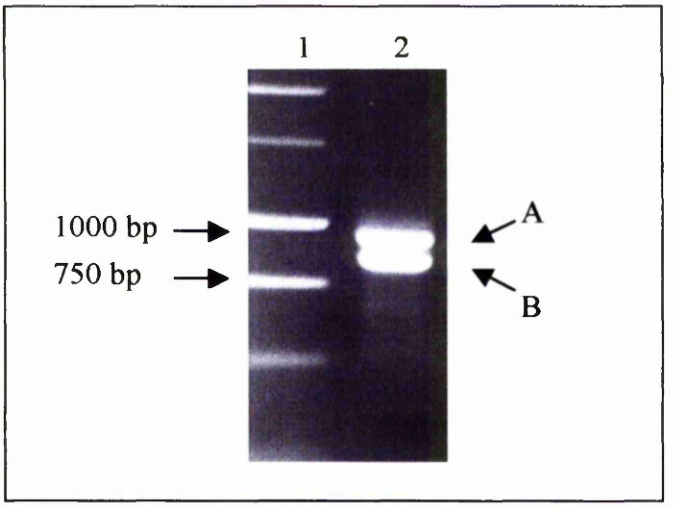


Figure 3.7: PCR product of 1 kb central region of a *TKT* gene from *L. mexicana*. The two PCR products in Lane 2 are approximately 950 bp (band A) and 850 bp (band B) in size according to the 1 kb molecular markers (Lane 1).

pTPPmotif

```
  1 GGCGACGGTT GCTAGATGGA AGGTGTGTGC CAGGAGGCGC TCTCCCTCGC CGGCCACCTC
 61 GCCCTGGAGA AGCTCATCGT CATCTATGAC AGCAACTACA TCTCCATCGA CGGCTCGACA
121 AGCCTCTCCT TCACGGAACA GTGCCACCAG AAGTACGTGG CCATGGGTTT CCACGTGATC
181 GAGGTCAAAA ACGGTGACAC TGACTACGAG GGCCTGCGCA AGGCACTGGC GGAAGCCAAG
241 GCCACGAAGG GCAAGCCGAA GATGATTGTG CAAACCACAA CGATTGGGTT CGGGTCTTCG
301 AAGCAGGGAA CGGAGAAGGT GCACGGCGCG CCGCTGGGTG AAGAGGATAT CGCCAACATC
361 AAGACAAAAT TTGGCCGCGA CCCGCAGAAG AAGTACGACG TCGACGACGA CGTCCGCGCT
421 GTGTTCAGGA TGCACATCGA TAAGTGTTCC GCGGAACAGA AGGCGTGGGA GGAACTCTTG
481 GCGAAGTACA CAGCCGCGTT CCCGGCCGAG GGTGCCGCCT TTGTGGCGCA GATGAGGGGC
541 GAGCTGCCGT CTGGGTGGGA GGCGAAGCTC CCGACGAACT CCTCGGCCAT CGCGACGCGC
601 AAGGCGAGCG AGAACTGCCT GGCTGTGCTC TTCCCGGCCA TCCCGGCTCT CATGGGCGGA
661 TCGGCTGACC TCACGCCGAG CAACCTGACG CGCCCCGCGT CGGCAAACTT GGTGGACTTC
721 TCGTCGAGCA GCAAGGAGGG TCGCTACATT CGCTTCGGTG TCCGTGAACA TGCCATGTGC
781 GCCATCCTCA ACGGTCTCGA CGCCCATGAT GGTATCATCC CGTTCGGCGG CACCTTCCTC
841 AACTTCATCG GCTACGCCCT TGGTGCAGTG CGCCTCGCCG CGATTTCTCA CCACCGCGTC
901 ATCTACGTGG CGACACACGA CAGCATCGGC CTAGGCGAAG ACGGTCCTAC TCATCA
```

pTKTmotif2

Figure 3.8: The central region of a *L. mexicana TKT* gene sequence amplified using the primers pTPP motif and pTKT motif 2 (highlighted in blue). 956 bp of sequence from two conserved regions in transketolase was amplified by PCR.

Sequences producing significant alignments:	Score (bits)	E Value
P23254|TKT1_YEAST TRANSKETOLASE 1 (TK 1)	312	5e-85
P34736|TKT_PICST TRANSKETOLASE (TK)	306	4e-83
Q12630|TKT1_KLULA TRANSKETOLASE (TK)	300	3e-81
P33315|TKT2_YEAST TRANSKETOLASE 2 (TK 2)	296	5e-80
Q43848|TKTC_SOLTU TRANSKETOLASE CHLOROPLAST PRECURSOR (TK)	274	2e-73
Q42676|TKTC_CRAPL TRANSKETOLASE CHLOROPLAST (TK)	271	2e-72
Q42675|TKTA_CRAPL TRANSKETOLASE 10 (TK)	259	8e-69
Q42677|TKT7_CRAPL TRANSKETOLASE 7 (TK)	254	3e-67
P33570|TKT2_ECOLI TRANSKETOLASE 2 (TK 2)	245	1e-64

Figure 3.9: Database search results using BLASTX algorithm to identify sequence similarity of 1 kb of *L. mexicana TKT*. The sequence (Figure 3.8) was searched for similarity to protein sequences from SWISS-PROT in May 1999. The initial numbers refer to the NCBI accession numbers, and the Score and E Value are referred to in the text.

3.3.2. Cloning the 5'-region of the transketolase gene

The 5'-region of the *TKT* gene was identified by using a reverse-transcriptase PCR approach (RT-PCR). Nested reverse primers were designed to DNA specific for *L. mexicana TKT*, and primers designed to the spliced leader sequence found on mRNAs of *L. mexicana amazonensis* (Agami and Shapira, 1992) (Figure 3.10). This approach had been previously used to amplify transcripts of *L. mexicana* glucose transporters (Burchmore and Landfear, 1998). The oligonucleotide sequences are detailed in Table 3.4.

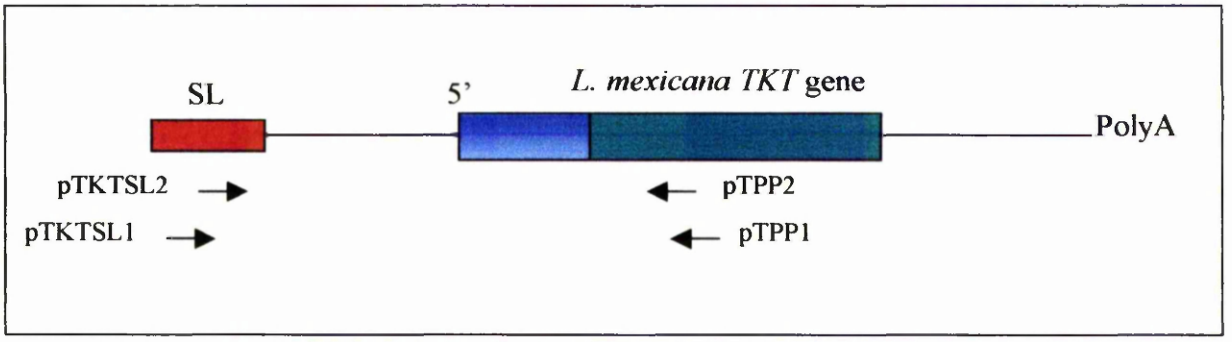


Figure 3.10: Graphic representation of the structure of the mRNA of the *TKT* gene from *L. mexicana*. The spliced leader (SL) and polyadenylated tail (polyA) are at the 5' and 3' of the mRNA transcript, respectively. The primers pTKTSL1 and pTKTSL2 are based on the spliced leader sequence, and pTPP1 and pTPP2 are antisense primers designed to previously determined *L. mexicana TKT* gene sequence.

Primer name	**Primer Sequence 5' → 3'**
pTKTSL1	TAA CGC TAT ATA AGT ATC AGT TTC
pTKTSL2	AGT ATC AGT TTC TGT ACT TTA TTG
pTPP1	CGA TCA CGT GGA AAC CCA TGG CC
pTPP2	CCG TGA AGG AGA GGC TTG TCG

Table 3.4: Primer sequences used to amplify the 5' region of *TKT* from *L. mexicana* cDNA. The spliced leader sequence present on mRNA was used to design primers to amplify *TKT* from cDNA using nested PCR.

100 ng of *L. mexicana* cDNA was used as the template for the initial PCR with the primers pTKTSL1 and pTPP1. 1 µl of the first 50 µl PCR reaction was used as a template for the nested PCR using primers pTKTSL2 and pTPP2. The PCR cycle used was as follows:

Pre-heat reaction:	60 seconds - 95 °C	
Melt and denature:	30 seconds - 95 °C	x 25 cycles
Anneal:	30 seconds - 52 °C	
Extension:	60 seconds - 72 °C	
Final hold:	10 minutes - 72 °C	

The nested PCR yielded a faint 750 bp product that was extracted from the gel and re-amplified using primers pTKTSL2 and pTPP2 (Figure 3.11). The amplified DNA was sub-cloned into pUC-18 for sequencing and was named pUC5'SL. The 5' sequence (Figure 3.12) was identified as having high similarity to transketolase proteins from other organisms (Figure 3.13) over part of its length. The start codon, ATG, was identified by translating the DNA sequence using Vector NTI and aligning the sequence with other transketolase proteins (BLASTX and Align X). The untranslated region between the spliced leader binding site and the start codon was 168 bp, with the start codon positioned at base-pair 169. Homology was seen to other transketolase sequences after this start codon.

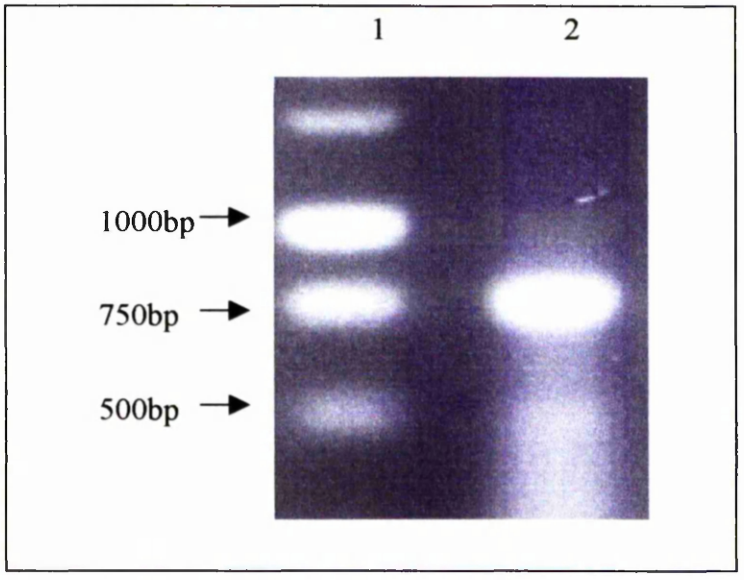


Figure 3.11: A PCR product from the amplification of the 5' *L. mexicana TKT* using an RT-PCR approach. The nested PCR product was re-amplified (Lane 2) so it could be cloned into pUC18 to be sequenced. The product was 750 bp in size according to the 1 kb marker (Lane 1)

```
          pTKTSL2
   1 AGTATCAGTT TCTGTACTTT ATTGCGCAAC ACCAGTAGCC ACCCTACGCT CTCACCAGTA
  61 TACCAAAGGA AAGCCTCCCT CTTCCTCTCG TTTCTCGAGC CCGTCATCAC ACCGTAGCAC
 121 TCGCCCCCGC GCCTTCGCCT CTCTCCGTCG CCCTCACACA CAAGCACAAT GGCCTCCATT
 181 GAGAAGGTGG CAAACTGCAT CCGCTGCCTC GCGGCGGACA TTGTCCAGGG CGGCAAGAGT
 241 GGCCACCCAG GCACGCCGAT GGGCATGGCG CCGATGTCAG CGGTCCTGTG GACGGAAGTG
 301 ATGAAGTACA ACAGCCAGGA TCCTGACTGG GTCGACCGCG ACCGCTTCGT CATGTCGAAC
 361 GGGCACGGCT GCGCACTGCA GTACGCCCTG CTGCACATGG CGGGCTACAA CCTCACCATG
 421 GACGACCTGA AGGGATTCCG CCAAGATGGC TCCCGCACCC CTGGCCACCC CGAGCGTTTC
 481 GTGACGCCCG GGGTGGAGGT GACGACCGGG CCACTTGGTC AGGGTATTGC AAACGCGGTC
 541 GGACTGGCGA TTGCCGAGGC GCACCTTGCC GCCACGTTCA ACCGCCCGGG ATACAACATC
 601 GTTGATCACT ACACTTACGT GTACTGTGGT GACGGTTGTC TGATGGAGGG TGTGTGCCAG
 661 GAGGCGCTCT CCCTCGCCGG CCACCTCGCC CTGAAGAAGC TCGTCGTCAT CTATGACAG
 721 AACTACATCT CCATCGACGG CTCGACAAGC CTCTCCTTCA CGG
                                        pTPP2
```

Figure 3.12: cDNA sequence of the 5' region of the *L. mexicana TKT* gene, including the putative start codon ATG. The primers pTKTSL2 and pTPP2 used to amplify this sequence are highlighted in blue. The putative ATG is underlined within the sequence.

Sequences producing significant alignments:	Score (bits)	E Value
P23254|TKT1_YEAST TRANSKETOLASE 1 (TK 1)	506	e-143
Q12630|TKT1_KLULA TRANSKETOLASE (TK)	488	e-137
P34736|TKT_PICST TRANSKETOLASE (TK)	483	e-136
Q43848|TKTC_SOLTU TRANSKETOLASE, CHL	483	e-136
P33315|TKT2_YEAST TRANSKETOLASE 2 (TK 2)	459	e-129
Q42677|TKT7_CRAPL TRANSKETOLASE 7 (TK)	454	e-127
P27302|TKT1_ECOLI TRANSKETOLASE 1 (TK 1)	449	e-126
Q42675|TKTA_CRAPL TRANSKETOLASE 10 (TK)	445	e-125
P21725|TKTC_ALCEU TRANSKETOLASE, CHR	443	e-124
P43757|TKT_HAEIN TRANSKETOLASE (TK)	442	e-123

Figure 3.13: BLASTX search outcome to identify sequence similarity to cloned 5' *L. mexicana* PCR product. The SWISS-PROT database was searched for proteins homologous to the DNA sequence (Figure 3.12) in August 2000.

3.3.3. Identification of the 3'-region of the transketolase gene

To identify the remaining section of the *TKT* gene, the *Leishmania major* DNA database at the Sanger Centre, Cambridge was used. The cloned 950 bp of *TKT* sequence initially identified was sent to Dr. Al Ivens at the Sanger Centre, where it was used to probe a cosmid library of *L. major* genomic DNA. 15 clones that were positive for *TKT* were picked from the master plate of the library. Individual clones were sent to Glasgow as glycerol stocks, grown in LBamp and the cosmid extracted. Each clone was checked for hybridisation to the *TKT* probe by streaking the culture onto a plate, transferring to a filter membrane and probing with ^{32}P dCTP labelled partial *TKT* sequence. 7/15 clones were confirmed to contain DNA with homology to the *TKT* probe.

In August 2000, a gene sequence with homology to *TKT* was identified in *L. major*. Chromosomes 12 and 24 of *L. major* co-migrate on pulsed field gels and are currently being shotgun sequenced by Beowulf Genomics. Preliminary analysis of unfinished shotgun sequence (Genbank accession number AL390114) located *TKT* to an assembled shotgun read of 4805 bp on chromosome 24. 1 to 3 kb of the *L. major* sequence showed good homology to *TKT* from other organisms with a BLASTX search. However, the sequence was unfinished, and therefore contained many unassigned bases, labelled 'n'. Nested primers were designed to this *L. major* sequence downstream of the putative stop codon, and *L. mexicana TKT* sequence within the region previously determined, to amplify the 3' region of the gene (Figure 3.14).

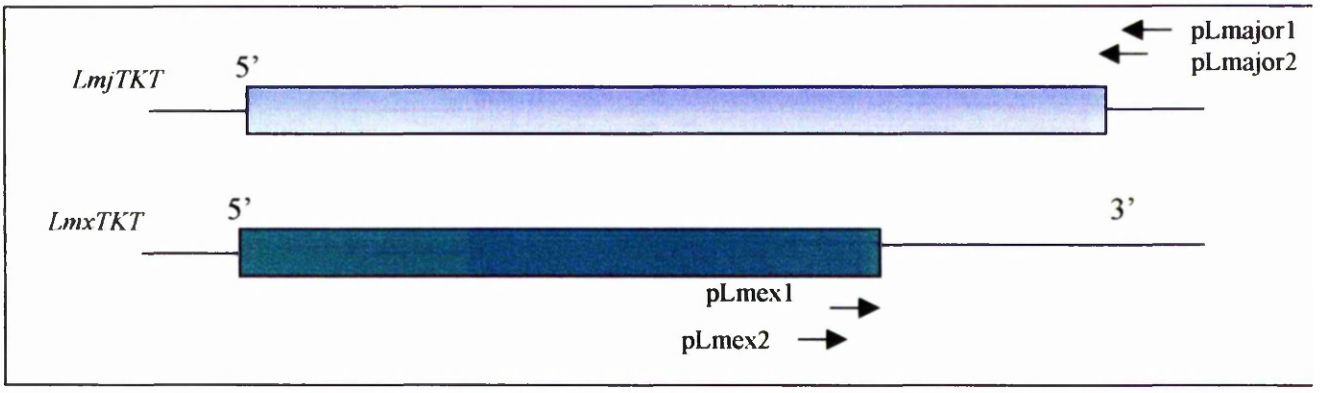


Figure 3.14: Schematic representation of the *LmjTKT* and the *LmxTKT* genes. The primers pLmex1 and pLmex2 complement the *L. mexicana TKT* gene sequence and pLmajor1 and pLmajor2 are antisense primers designed to the 3' flanking region of the *L. mexicana TKT* gene.

Primer name	Primer sequence 5' → 3'
pLmajor1	CGC GTT GCT GTG TGA GCG TA
pLmajor2	ATA AGA GGA GAG AGA GGT CAG G
pLmex1	GAC GAC GAC GTC CGC GCT GTG TT
pLmex2	GAA GCT CCC GAC GAA CTC C

Table 3.5: Oligonucleotides used to amplify the 3' region of the *TKT* gene from *L. mexicana*. The antisense primers were derived from *L. major* sequence, with similarity to transketolase, present in the *L. major* sequencing project database, the Sanger Centre.

L. mexicana genomic DNA and *L. major* cosmid DNA were used as templates for individual PCRs using the primers listed in Table 3.5. The PCR conditions used were as follows:

Pre-heat reaction:	60 seconds - 95 °C	
Melt and denature:	30 seconds - 95 °C	
Anneal:	30 seconds - 52 °C	x 25 cycles
Extension:	60 seconds - 72 °C	
Final hold:	10 minutes - 72 °C	

The first round PCR using primers pLmajor1 and pLmex1, amplified a product of 1.2 kb, which was gel eluted and used as template for a second round PCR with primers pLmajor2 and pLmex2. Two bands were amplified (Figure 3.15), one 250 bp and the other 1 kb. The 1 kb product was gel eluted, cloned into the vector pGEM-T (named pG3'Lmex) and transformed into *E. coli* strain JM109. The insert was sequenced (Figure 3.16) and its predicted translation product had good homology to the 3' region of other TKT protein sequences (Figure 3.17).

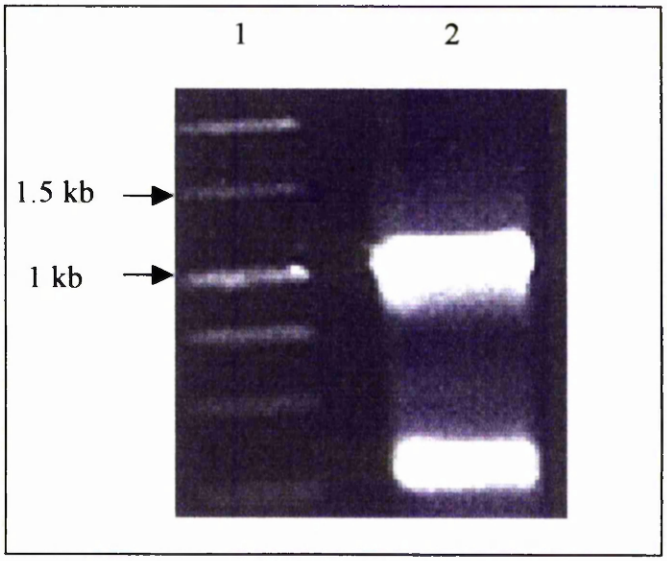


Figure 3.15: PCR product from amplification of 3' region of *L. mexicana TKT*. Two products were amplified (Lane 2) of 1 kb and 250 bp (Lane 1). The 1 kb product was cloned into pGEM-T, as it was the predicted size of the PCR.

pLmex2

```
   1 GAAGCTCCCG ACGAACTCCT CGGCCATCGC GACGCGCAAG GCGAGCGAGA ACTGCCTGGC
  61 TGTGCTCTTC CCGGCCATCC CGGCTCTCAT GGGCGGATCG GCTGACCTCA CGCCGAGCAA
 121 CCTGACGCGC CCCGCGTCGG CAAACTTGGT GGACTTCTCG TCGAGCAGCA AGGAGGGTCG
 181 CTACATTCGC TTCGGTGTCC GTGAACATGC CATGTGCGCC ATCCTCAACG GTCTCGACGC
 241 CCATGATGGT ATCATCCCGT TCGGCGGCAC CTTCCTCAAC TTCATCGGCT ACGCCCTTGG
 301 TGCAGTGCGC CTCGCCGCGA TTTCTCACCA CCGCGTCATC TACGTGGCGA CACACGACAG
 361 CATCGGCGTT GGCGAGGACG GGCCAACCCA CCAGCCTGTC GAGTTGGTGG CTGCCCTGCG
 421 TGCYATGCCA AACCTGCAGG TGATACGTCC TAGCGACCAG ACAGAGACGA GCGGTGCGTG
 481 GGCTGTTGCA CTGTCTAGTA TTCACACTCC AACGGTTCTG TGTCTGAGCC GCCAGAACAC
 541 CGAGCCGCAG TCGGGGTCGA GCATCGAGGG TGTGAGGCAC GGCGCCTACT CGGTGGTGGA
 601 TGTGCCCGAC CTGCAGCTCG TGATCGTGGC GAGCGGCTCG GAGGTGTCGC TGGCGGTGGA
 661 TGCTGCCAAA GCGCTCTCGG GTGAGCTGCG CGTAAGGGTC GTGTCGATGC CGTGCCAGGA
 721 GCTCTTCGAC GCACAACCGG ATACGTACCG CCAGGCTGTG CTCCCCGCGG GTGTGCCGGT
 781 GGTGTCGGTG GAGGCGTACG TCAGTTTTGG CTGGGAAAAA TACTCCCATG CGCACGTGGG
 841 CATGTCCGGC TTCGGTGCCT CGGCCCCGGC GGGTGTGCTA TACAAGAAGT TTGAAATTAC
 901 CGTCGAGGAA GTGGTGAGGA CGGGCCGTGA GTTGGCCAAG CGCTTCCCCG ATGGCACGGC
 961 GCCGCTCAAG AACTCTTCAT TCAGCAAGAT GTAAGGTGAG CCGGTGCTCC GAAACGTGAG
1021 GAATGACACC GGAAGAGGCG GGCAGGGAAG AGGGCGGATG GGGGAGGAAG ATGGTTGCTG
1081 GCGATGGCGG CCCCGCTTGC CCCTGACCTC TCTCTCCTCT TAT
```

pLmajor2

Figure 3.16: DNA sequence of the 3' region of the *L. mexicana TKT* gene. The primers pLmex2 and pLmajor2 are highlighted in blue, with the putative stop codon TAA being underlined.

Sequences producing significant alignments:	Score (bits)	E Value
P23254|TKT1_YEAST TRANSKETOLASE 1 (TK 1)	310	5e-84
Q12630|TKT1_KLULA TRANSKETOLASE (TK)	307	4e-83
P33315|TKT2_YEAST TRANSKETOLASE 2 (TK 2)	293	9e-79
P34736|TKT_PICST TRANSKETOLASE (TK)	290	5e-78
P29277|TKT_RHOSH TRANSKETOLASE (TK)	280	5e-75
Q52723|TKT_RHOCA TRANSKETOLASE (TK)	275	1e-73
P21726|TKTP_ALCEU TRANSKETOLASE, PLASMID (TK)	273	6e-73
Q43848|TKTC_SOLTU TRANSKETOLASE, CHL	270	4e-72
P43757|TKT_HAEIN TRANSKETOLASE (TK)	268	2e-71
Q42676|TKTC_CRAPL TRANSKETOLASE, CHL	268	2e-71

Figure 3.17: Database search outcome using BLASTX to search for similarity of the amplified 3' *L. mexicana TKT* using SWISS-PROT database (September 2000).

The stop codon was identified, as it is the first in frame stop codon in the sequence and by alignment of the predicted translation product with other transketolase amino acid sequences. It is located at 992 bp in Figure 3.18. The 3' sequence overlaps with the initial 1 kb sequenced (Figure 3.8) up to position 367 bp in Figure 3.16. Therefore an additional 625 bp of 3' sequence was identified. This was the final step in cloning the transketolase gene and assembly of the full-length sequence followed. At this point 144 bp and 112 bp of upstream and downstream sequence of the *TKT* gene was obtained.

3.3.4. Amplification of the full *L. mexicana TKT* ORF

To clone the full-length *L. mexicana TKT* gene, primers were designed to the 5' and 3' domains of the gene sequence containing the start and stop codons (Figure 3.18). These primers were designed to incorporate restriction sites (Table 3.6), and create an ORF when cloned into pET-16b for expression of the recombinant protein with a hexa-histidine tag.

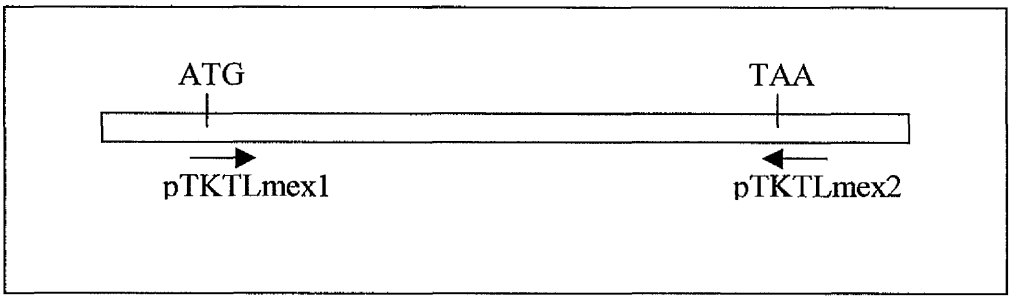


Figure 3.18: Diagrammatic representation of the location of the primers to amplify the full-length *L. mexicana TKT* gene. The primers pTKTLmex1 and pTKTLmex2 correspond to the putative start and stop codons amplified previously.

Primer Name	Primer Sequence 5' → 3'	Restriction site
pTKTLmex1	TCA CAC ACA AGC CAT ATG GCC TCC	*Nde*I
pTKTLmex2	GCG CCC TCG AGC CT TAC ATC TTG C	*Xho*I

Table 3.6: Primers to amplify the full-length *TKT* gene from *L. mexicana*. The restriction sites *Nde*I and *Xho*I (highlighted in red) are present in the oligonucleotides to directionally clone the PCR product into the expression vector pET-16b.

Various conditions were used during PCR in order to amplify a product. The final PCR cycle used was as follows:

Pre-heat reaction:	60 seconds - 95 °C	
Melt and denature:	30 seconds - 95 °C	} x 30 cycles
Anneal:	30 seconds - 52 °C	
Extension:	3 minute - 72 °C	
Final hold:	10 minutes - 72 °C	

The increased extension time was used to preferentially amplify larger products. A 2 kb product was amplified from the PCR mix and can be seen in Figure 3.19. These products were confirmed to be *TKT* by diagnostic PCR using two sets of internal primers.

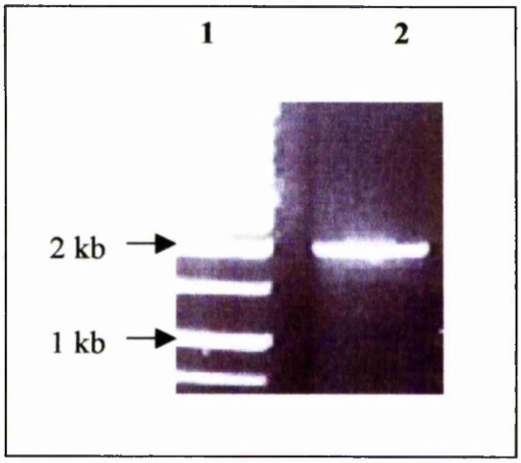


Figure 3.19: PCR products from the amplification of the full-length *L. mexicana TKT*. The PCR product (Lane 2) was the predicted size of 2 kb (Lane1).

The 2 kb PCR product was initially cloned into pGEM-T and cut from the vector using *Nde*I and *Xho*I restriction endonucleases. This was to ensure that the insert contained cut restriction sites for the next ligation. The cut insert was ligated with pET-16b (previously double digested with *Nde*I and *Xho*I) overnight at 4 °C. A control ligation was also set up containing no insert DNA, to control for the self-ligation of the vector. The ligation mix was transformed into *E. coli* JM109 using a heat shock method. The resulting plasmids were checked with PCR using primers for the *TKT* gene, and were digested with *Nde*I and *Xho*I. The plasmid (pET16bLmtkt) was transformed into competent *E. coli* BL21(DE3) cells and also sent for sequencing. A representation of the plasmid containing the insert and the DNA sequence of the *TKT* gene are in Figures 3.20 and 3.21, respectively. This plasmid was used for the expression of recombinant *L. mexicana TKT* (Chapter IV).

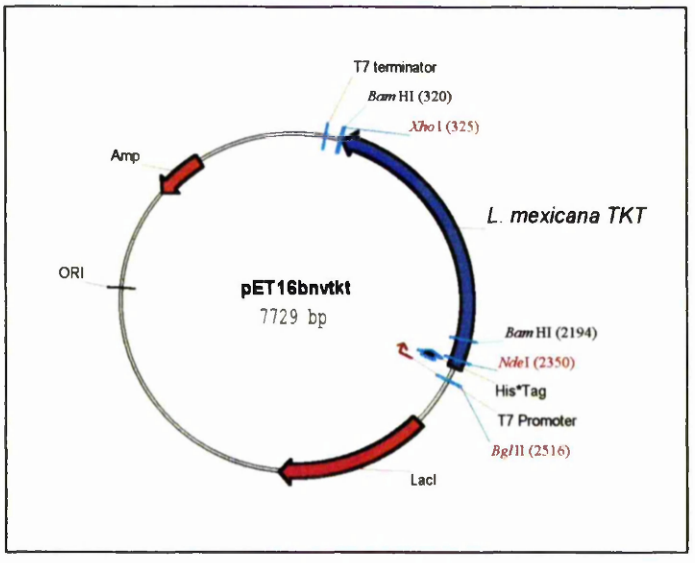


Figure 3.20: Diagrammatic representation of the expression plasmid pET-16b containing the *L. mexicana TKT* gene. The gene was cloned into the *Nde*I/*Bam*HI restriction sites and is represented by a block blue arrow in the direction of transcription. The other features of the plasmid are detailed in Figure 3.2

```
      M   A   S   I   E   K   V   A   N   C   I   R   C   L   A   A   D   I   V   Q
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1     ATG GCC TCC ATT GAG AAG GTG GCA AAC TGC ATC CGC TGC CTC GCG GCG GAC ATT GTC CAG
       G   G   K   S   G   H   P   G   T   P   M   G   M   A   P   M   S   A   V   L
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
61    GGC GGC AAG AGT GGC CAC CCA GGC ACG CCG ATG GGC ATG GCG CCG ATG TCA GCG GTC CTG
       W   T   E   V   M   K   Y   N   S   Q   D   P   D   W   V   D   R   D   R   F
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
121   TGG ACG GAA GTG ATG AAG TAC AAC AGC CAG GAT CCT GAC TGG GTC GAC CGC GAC CGC TTC
       V   M   S   N   G   H   G   C   A   L   Q   Y   A   L   L   H   M   A   G   Y
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
181   GTC ATG TCG AAC GGG CAC GGC TGC GCA CTG CAG TAC GCC CTG CTG CAC ATG GCA GGC TAC
       N   L   T   M   D   D   L   K   G   F   R   Q   D   G   S   R   T   P   G   H
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
241   AAC CTC ACC ATG GAC GAC CTG AAG GGA TTC CGC CAA GAT GGC TCC CGC ACC CCT GGC CAC
       P   E   R   F   V   T   P   G   V   E   V   T   T   G   P   L   G   Q   G   I
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
301   CCC GAG CGT TTC GTG ACG CCC GGG GTG GAG GTG ACG ACC GGG CCA CTT GGC CAG GGT ATT
       A   N   A   V   G   L   A   I   A   E   A   H   L   A   A   T   F   N   R   P
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
361   GCA AAC GCG GTC GGA CTG GCG ATT GCC GAG GCG CAC CTT GCC GCC ACG TTC AAC CGC CCG
       G   Y   N   I   V   D   H   Y   T   Y   V   Y   C   G   D   G   C   L   M   E
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
421   GGA TAC AAC ATC GTT GAT CAC TAC ACT TAC GTG TAC TGT GGT GAC GGT TGT CTG ATG GAG
       G   V   C   Q   E   A   L   S   L   A   G   H   L   A   L   E   K   L   I   V
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
481   GGT GTG TGC CAG GAG GCG CTC TCC CTC GCC GGC CAC CTC GCC CTG GAG AAG CTC ATC GTC
       I   Y   D   S   N   Y   I   S   I   D   G   S   T   S   L   S   F   T   E   Q
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
541   ATC TAT GAC AGC AAC TAC ATC TCC ATC GAC GGC TCG ACA AGC CTC TCC TTC ACG GAA CAG
       C   H   Q   K   Y   V   A   M   G   F   H   V   I   E   V   K   N   G   D   T
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
601   TGC CAC CAG AAG TAC GTG GCC ATG GGT TTC CAC GTG ATC GAG GTC AAA AAC GGT GAC ACT
       D   Y   E   G   L   R   K   A   L   A   E   A   K   A   T   K   G   K   P   K
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
661   GAC TAC GAG GGC CTG CGC AAG GCA CTG GCG GAG GCC AAG GCC ACG AAG GGC AAG CCG AAG
       M   I   V   Q   T   T   T   I   G   F   G   S   S   K   Q   G   T   E   K   V
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
721   ATG ATT GTG CAA ACC ACA ACG ATT GGG TTC GGG TCT TCG AAG CAG GGA ACG GAG AAG GTG
       H   G   A   P   L   G   E   E   D   I   A   N   I   K   A   K   F   G   R   D
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
781   CAC GGC GCG CCG CTG GGT GAA GAG GAT ATC GCC AAC ATC AAG GCA AAA TTT GGC CGC GAC
       P   Q   K   K   Y   D   V   D   D   D   V   R   A   V   F   R   M   H   I   D
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
841   CCG CAG AAG AAG TAC GAC GTC GAC GAC GAC GTC CGC GCT GTG TTC AGG ATG CAC ATC GAT
       K   C   S   A   E   Q   K   A   W   E   E   L   L   A   K   Y   T   A   A   F
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
901   AAG TGT TCC GCG GAA CAG AAG GCG TGG GAG GAA CTC TTG GCG AAG TAC ACA GCC GCG TTC
       P   A   E   G   A   A   F   V   A   Q   M   R   G   E   L   P   S   G   W   E
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
961   CCG GCC GAG GGT GCC GCC TTT GTG GCG CAG ATG AGG GGC GAG CTG CCG TCT GGG TGG GAG
       A   K   L   P   T   N   S   S   A   I   A   T   R   K   A   S   E   N   C   L
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1021  GCG AAG CTC CCG ACG AAC TCC TCG GCC ATC GCG ACG CGC AAG GCG AGC GAG AAC TGC CTG
       A   V   L   F   P   A   I   P   A   L   M   G   G   S   A   D   L   T   P   S
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1081  GCT GTG CTC TTC CCG GCC ATC CCG GCT CTC ATG GGC GGA TCG GCT GAC CTC ACG CCG AGC
       N   L   T   R   P   A   S   A   N   L   V   D   F   S   S   S   S   K   E   G
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1141  AAC CTG ACG CGC CCC GCG TCG GCA AAC TTG GTG GAC TTC TCG TCG AGC AGC AAG GAG GGT
       R   Y   I   R   F   G   V   R   E   H   A   M   C   A   I   L   N   G   L   D
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1201  CGC TAC ATT CGC TTC GGT GTC CGT GAA CAT GCC ATG TGC GCC ATC CTC AAC GGT CTC GAC
       A   H   D   G   I   I   P   F   G   G   T   F   L   N   F   I   G   Y   A   L
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1261  GCC CAT GAT GGT ATC ATC CCG TTC GGC GGC ACC TTC CTC AAC TTC ATC GGC TAC GCC CTT
```

Figure 3.21: *L. mexicana TKT* gene sequence and translated amino acid sequence as predicted by Vector NTI. The DNA sequence is presented in sets of 3, and the corresponding amino acid sequence is shown above. The number to the left indicates the position of the DNA.

```
        G   A   V   R   L   A   A   I   S   H   H   R   V   I   Y   V   A   T   H   D
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1321    GGT GCA GTG CGC CTC GCC GCG ATT TCT CAC CAC CGC GTC ATC TAC GTG GCG ACA CAC GAC
         S   I   G   V   G   E   D   G   P   T   H   Q   P   V   E   L   V   A   A   L
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1381    AGC ATC GGC GTT GGC GAG GAC GGG CCA ACC CAC CAG CCT GTC GAG TTG GTG GCT GCC CTG
         R   A   M   P   N   L   Q   V   I   R   P   S   D   Q   T   E   T   S   G   A
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1441    CGT GCC ATG CCA AAC CTG CAG GTG ATA CGT CCT AGC GAC CAG ACA GAG ACG AGC GGT GCG
         W   A   V   A   L   S   S   I   H   T   P   T   V   L   C   L   S   R   Q   N
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1501    TGG GCT GTT GCA CTG TCT AGT ATT CAC ACT CCA ACG GTT CTG TGT CTG AGC CGC CAG AAC
         T   E   P   Q   S   G   S   S   I   E   G   V   R   H   G   A   Y   S   V   V
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1561    ACC GAG CCG CAG TCG GGG TCG AGC ATC GAG GGT GTG AGG CAC GGC GCC TAC TCG GTG GTG
         D   V   P   D   L   Q   L   V   I   V   A   S   G   S   E   V   S   L   A   V
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1621    GAT GTG CCC GAC CTG CAG CTC GTG ATC GTG GCG AGC GGC TCG GAG GTG TCG CTG GCG GTG
         D   A   A   K   A   L   S   G   E   L   R   V   R   V   V   S   M   P   C   Q
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1681    GAT GCT GCC AAA GCG CTC TCG GGT GAG CTG CGC GTA AGG GTC GTG TCG ATG CCG TGC CAG
         E   L   F   D   A   Q   P   D   T   Y   R   Q   A   V   L   P   A   G   V   P
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1741    GAG CTC TTC GAC GCA CAA CCG GAT ACG TAC CGC CAG GCT GTG CTC CCC GCG GGT GTG CCG
         V   V   S   V   E   A   Y   V   S   F   G   W   E   K   Y   S   H   A   H   V
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1801    GTG GTG TCG GTG GAG GCG TAC GTC AGT TTT GGC TGG GAG AAA TAC TCC CAT GCG CAC GTG
         G   M   S   G   F   G   A   S   A   P   A   G   V   L   Y   K   K   F   G   I
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1861    GGC ATG TCC GGC TTC GGT GCC TCG GCC CCG GCG GGT GTG CTA TAC AAG AAG TTT GGA ATT
         T   V   E   E   V   V   R   T   G   R   E   L   A   K   R   F   P   D   G   T
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1921    ACC GTC GAG GAA GTG GTG AGG ACG GGC CGT GAG TTG GCC AAG CGC TTC CCC GAT GGC ACG
         A   P   L   K   N   S   S   F   S   K   M
        ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
1981    GCG CCG CTC AAG AAC TCT TCA TTC AGC AAG ATG TAA
```

Figure 3.21 continued: *L. mexicana TKT* gene sequence and translated amino acid sequence

The *L. mexicana TKT* gene was 2016 bp in length, encoding a predicted protein of 672 amino acids. In addition to containing the well conserved TPP-domain and TKT motif, an interesting feature of the protein encoded by this gene was a putative C-terminal peroxisome-targeting signal (PTS). This amino acid sequence, SKM, indicated that the gene might be targeted to the glycosome, a peroxisome-like organelle found in trypanosomatids.

The full-length *L. mexicana TKT* gene was amplified using Taq DNA polymerase as the Pfu amplified product could not be successfully cloned. To ensure that the PCR product was the correct sequence, the full-length *TKT* and the three individual partial *TKT* sequences were aligned (Figure 3.22). Six base pair changes were observed between the two sequences, encoding for two amino acid changes (Table 3.7). The glycine residue at the amino acid position 638 (1916 bp) is also present in the 3' *L. major TKT* partial gene sequence and predicted amino acid sequence, confirming that gene encodes for the correct amino acid at this position. The alanine residue at the amino acid position 274 (823 bp) in the full-length sequences changes to a threonine residue in the assembled sequence. This amino acid change occurs in a region of low homology between residues and it is not clear which is correct. This residue is a glutamic acid residue in a *T. brucei* fragment with similarity to this region of transketolase (see Chapter V for an alignment of the fragments). Since the alteration occurs in a poorly conserved part of the protein it was considered unlikely to have much impact on the enzymes activity.

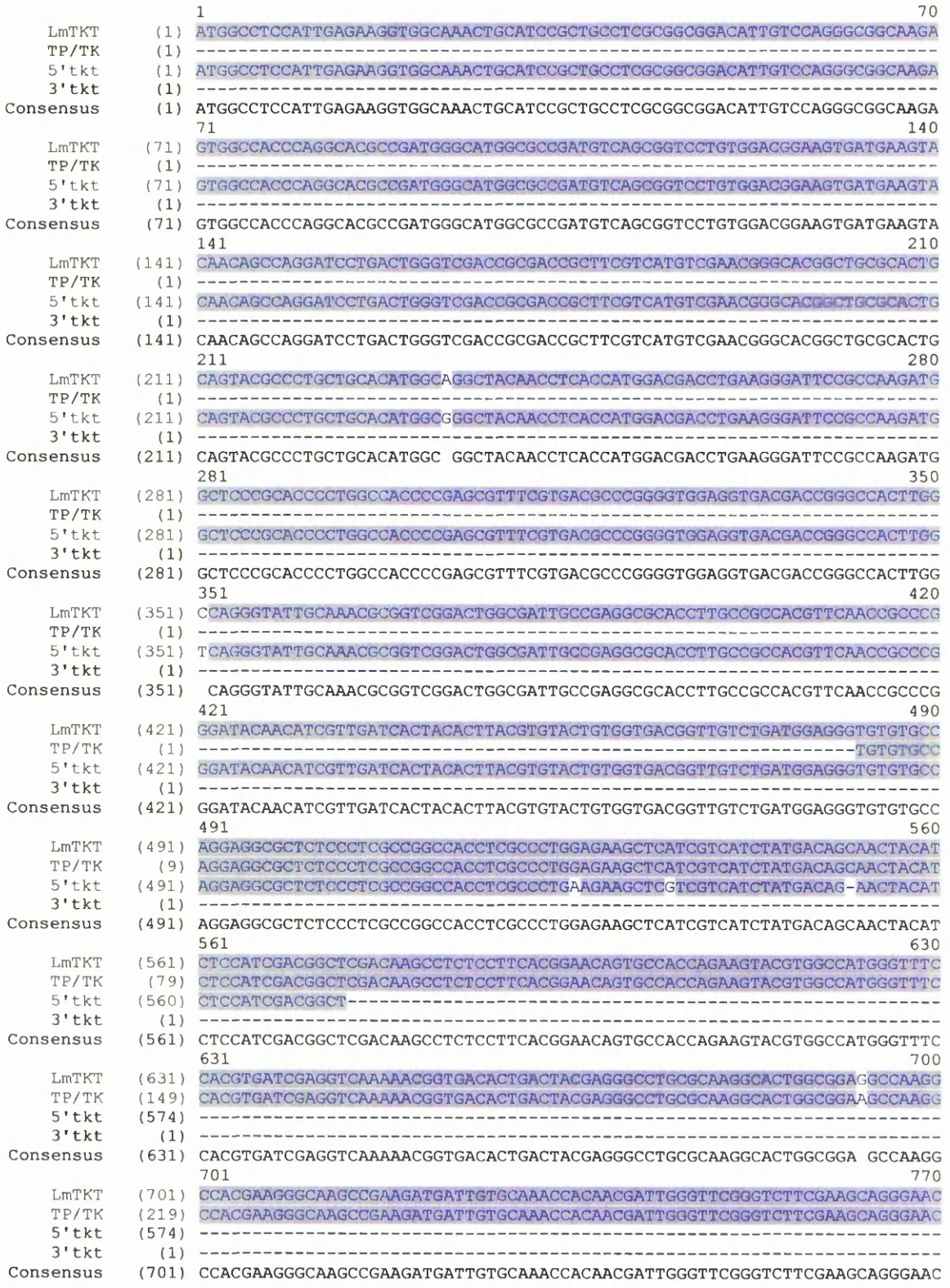


Figure 3.22: Alignment of the *L. mexicana TKT* gene sequence with the three individually amplified sections of the *TKT* gene. The full-length DNA sequence (LmTKT) is aligned with the 1 kb central region of the gene (TP/TK), the 5' region amplified from cDNA (5'tkt) and the 3' region amplified using the *L. major* template (3'tkt). The blocks of identity are highlighted by blue text with a grey background.

```
                771                                                                 840
     LmTKT     (771) GGAGAAGGTGCACGGCGCGCCGCTGGGTGAAGAGGATATCGCCAACATCAAGGCAAAATTTGGCCGCGAC
     TP/TK     (289) GGAGAAGGTGCACGGCGCGCCGCTGGGTGAAGAGGATATCGCCAACATCAAGACAAAATTTGGCCGCGAC
     5'tkt     (574) ----------------------------------------------------------------------
     3'tkt       (1) ----------------------------------------------------------------------
 Consensus     (771) GGAGAAGGTGCACGGCGCGCCGCTGGGTGAAGAGGATATCGCCAACATCAAG CAAAATTTGGCCGCGAC
                841                                                                 910
     LmTKT     (841) CCGCAGAAGAAGTACGACGTCGACGACGACGTCCGCGCTGTGTTCAGGATGCACATCGATAAGTGTTCCG
     TP/TK     (359) CCGCAGAAGAAGTACGACGTCGACGACGACGTCCGCGCTGTGTTCAGGATGCACATCGATAAGTGTTCCG
     5'tkt     (574) ----------------------------------------------------------------------
     3'tkt       (1) ----------------------------------------------------------------------
 Consensus     (841) CCGCAGAAGAAGTACGACGTCGACGACGACGTCCGCGCTGTGTTCAGGATGCACATCGATAAGTGTTCCG
                911                                                                 980
     LmTKT     (911) CGGAACAGAAGGCGTGGGAGGAACTCTTGGCGAAGTACACAGCCGCGTTCCCGGCCGAGGGTGCCGCCTT
     TP/TK     (429) CGGAACAGAAGGCGTGGGAGGAACTCTTGGCGAAGTACACAGCCGCGTTCCCGGCCGAGGGTGCCGCCTT
     5'tkt     (574) ----------------------------------------------------------------------
     3'tkt       (1) ----------------------------------------------------------------------
 Consensus     (911) CGGAACAGAAGGCGTGGGAGGAACTCTTGGCGAAGTACACAGCCGCGTTCCCGGCCGAGGGTGCCGCCTT
                981                                                                1050
     LmTKT     (981) TGTGGCGCAGATGAGGGGCGAGCTGCCGTCTGGGTGGGAGGCGAAGCTCCCGACGAACTCCTCGGCCATC
     TP/TK     (499) TGTGGCGCAGATGAGGGGCGAGCTGCCGTCTGGGTGGGAGGCGAAGCTCCCGACGAACTCCTCGGCCATC
     5'tkt     (574) ----------------------------------------------------------------------
     3'tkt       (1) -------------------------------------------------------------TCGGCCATC
 Consensus     (981) TGTGGCGCAGATGAGGGGCGAGCTGCCGTCTGGGTGGGAGGCGAAGCTCCCGACGAACTCCTCGGCCATC
                1051                                                               1120
     LmTKT    (1051) GCGACGCGCAAGGCGAGCGAGAACTGCCTGGCTGTGCTCTTCCCGGCCATCCCGGCTCTCATGGGCGGAT
     TP/TK     (569) GCGACGCGCAAGGCGAGCGAGAACTGCCTGGCTGTGCTCTTCCCGGCCATCCCGGCTCTCATGGGCGGAT
     5'tkt     (574) ----------------------------------------------------------------------
     3'tkt      (10) GCGACGCGCAAGGCGAGCGAGAACTGCCTGGCTGTGCTCTTCCCGGCCATCCCGGCTCTCATGGGCGGAT
 Consensus    (1051) GCGACGCGCAAGGCGAGCGAGAACTGCCTGGCTGTGCTCTTCCCGGCCATCCCGGCTCTCATGGGCGGAT
                1121                                                               1190
     LmTKT    (1121) CGGCTGACCTCACGCCGAGCAACCTGACGCGCCCCGCGTCGGCAAACTTGGTGGACTTCTCGTCGAGCAG
     TP/TK     (639) CGGCTGACCTCACGCCGAGCAACCTGACGCGCCCCGCGTCGGCAAACTTGGTGGACTTCTCGTCGAGCAG
     5'tkt     (574) ----------------------------------------------------------------------
     3'tkt      (80) CGGCTGACCTCACGCCGAGCAACCTGACGCGCCCCGCGTCGGCAAACTTGGTGGACTTCTCGTCGAGCAG
 Consensus    (1121) CGGCTGACCTCACGCCGAGCAACCTGACGCGCCCCGCGTCGGCAAACTTGGTGGACTTCTCGTCGAGCAG
                1191                                                               1260
     LmTKT    (1191) CAAGGAGGGTCGCTACATTCGCTTCGGTGTCCGTGAACATGCCATGTGCGCCATCCTCAACGGTCTCGAC
     TP/TK     (709) CAAGGAGGGTCGCTACATTCGCTTCGGTGTCCGTGAACATGCCATGTGCGCCATCCTCAACGGTCTCGAC
     5'tkt     (574) ----------------------------------------------------------------------
     3'tkt     (150) CAAGGAGGGTCGCTACATTCGCTTCGGTGTCCGTGAACATGCCATGTGCGCCATCCTCAACGGTCTCGAC
 Consensus    (1191) CAAGGAGGGTCGCTACATTCGCTTCGGTGTCCGTGAACATGCCATGTGCGCCATCCTCAACGGTCTCGAC
                1261                                                               1330
     LmTKT    (1261) GCCCATGATGGTATCATCCCGTTCGGCGGCACCTTCCTCAACTTCATCGGCTACGCCCTTGGTGCAGTGC
     TP/TK     (779) GCCCATGATGGTATCATCCCGTTCGGCGGCACCTTCCTCAACTTCATCGGCTACGCCCTTGGTGCAGTGC
     5'tkt     (574) ----------------------------------------------------------------------
     3'tkt     (220) GCCCATGATGGTATCATCCCGTTCGGCGGCACCTTCCTCAACTTCATCGGCTACGCCCTTGGTGCAGTGC
 Consensus    (1261) GCCCATGATGGTATCATCCCGTTCGGCGGCACCTTCCTCAACTTCATCGGCTACGCCCTTGGTGCAGTGC
                1331                                                               1400
     LmTKT    (1331) GCCTCGCCGCGATTTCTCACCACCGCGTCATCTACGTGGCGACACACGACAGCATCGGCGTTGGCGAGGA
     TP/TK     (849) GCCTCGCCGCGATTTCTCACCACCGCGTCATCTACGTGGCGACACACGACAGCATCGGC-----------
     5'tkt     (574) ----------------------------------------------------------------------
     3'tkt     (290) GCCTCGCCGCGATTTCTCACCACCGCGTCATCTACGTGGCGACACACGACAGCATCGGCGTTGGCGAGGA
 Consensus    (1331) GCCTCGCCGCGATTTCTCACCACCGCGTCATCTACGTGGCGACACACGACAGCATCGGCGTTGGCGAGGA
                1401                                                               1470
     LmTKT    (1401) CGGGCCAACCCACCAGCCTGTCGAGTTGGTGGCTGCCCTGCGTGCCATGCCAAACCTGCAGGTGATACGT
     TP/TK     (908) ----------------------------------------------------------------------
     5'tkt     (574) ----------------------------------------------------------------------
     3'tkt     (360) CGGGCCAACCCACCAGCCTGTCGAGTTGGTGGCTGCCCTGCGTGCCATGCCAAACCTGCAGGTGATACGT
 Consensus    (1401) CGGGCCAACCCACCAGCCTGTCGAGTTGGTGGCTGCCCTGCGTGCCATGCCAAACCTGCAGGTGATACGT
                1471                                                               1540
     LmTKT    (1471) CCTAGCGACCAGACAGAGACGAGCGGTGCGTGGGCTGTTGCACTGTCTAGTATTCACACTCCAACGGTTC
     TP/TK     (908) ----------------------------------------------------------------------
     5'tkt     (574) ----------------------------------------------------------------------
     3'tkt     (430) CCTAGCGACCAGACAGAGACGAGCGGTGCGTGGGCTGTTGCACTGTCTAGTATTCACACTCCAACGGTTC
 Consensus    (1471) CCTAGCGACCAGACAGAGACGAGCGGTGCGTGGGCTGTTGCACTGTCTAGTATTCACACTCCAACGGTTC
```

Figure 3.22 continued: Alignment of the *L. mexicana TKT* gene sequence with the three individually amplified sections of the *TKT* gene.

```
                    1541                                                                1610
    LmTKT   (1541)  TGTGTCTGAGCCGCCAGAACACCGAGCCGCAGTCGGGGTCGAGCATCGAGGGTGTGAGGCACGGCGCCTA
    TP/TK    (908)  ------------------------------------------------------------------------
    5'tkt    (574)  ------------------------------------------------------------------------
    3'tkt    (500)  TGTGTCTGAGCCGCCAGAACACCGAGCCGCAGTCGGGGTCGAGCATCGAGGGTGTGAGGCACGGCGCCTA
Consensus   (1541)  TGTGTCTGAGCCGCCAGAACACCGAGCCGCAGTCGGGGTCGAGCATCGAGGGTGTGAGGCACGGCGCCTA
                    1611                                                                1680
    LmTKT   (1611)  CTCGGTGGTGGATGTGCCCGACCTGCAGCTCGTGATCGTGGCGAGCGGCTCGGAGGTGTCGCTGGCGGTG
    TP/TK    (908)  ------------------------------------------------------------------------
    5'tkt    (574)  ------------------------------------------------------------------------
    3'tkt    (570)  CTCGGTGGTGGATGTGCCCGACCTGCAGCTCGTGATCGTGGCGAGCGGCTCGGAGGTGTCGCTGGCGGTG
Consensus   (1611)  CTCGGTGGTGGATGTGCCCGACCTGCAGCTCGTGATCGTGGCGAGCGGCTCGGAGGTGTCGCTGGCGGTG
                    1681                                                                1750
    LmTKT   (1681)  GATGCTGCCAAAGCGCTCTCGGGTGAGCTGCGCGTAAGGGTCGTGTCGATGCCGTGCCAGGAGCTCTTCG
    TP/TK    (908)  ------------------------------------------------------------------------
    5'tkt    (574)  ------------------------------------------------------------------------
    3'tkt    (640)  GATGCTGCCAAAGCGCTCTCGGGTGAGCTGCGCGTAAGGGTCGTGTCGATGCCGTGCCAGGAGCTCTTCG
Consensus   (1681)  GATGCTGCCAAAGCGCTCTCGGGTGAGCTGCGCGTAAGGGTCGTGTCGATGCCGTGCCAGGAGCTCTTCG
                    1751                                                                1820
    LmTKT   (1751)  ACGCACAACCGGATACGTACCGCCAGGCTGTGCTCCCCGCGGGTGTGCCGGTGGTGTCGGTGGAGGCGTA
    TP/TK    (908)  ------------------------------------------------------------------------
    5'tkt    (574)  ------------------------------------------------------------------------
    3'tkt    (710)  ACGCACAACCGGATACGTACCGCCAGGCTGTGCTCCCCGCGGGTGTGCCGGTGGTGTCGGTGGAGGCGTA
Consensus   (1751)  ACGCACAACCGGATACGTACCGCCAGGCTGTGCTCCCCGCGGGTGTGCCGGTGGTGTCGGTGGAGGCGTA
                    1821                                                                1890
    LmTKT   (1821)  CGTCAGTTTTGGCTGGGAGAAATACTCCCATGCGCACGTGGGCATGTCCGGCTTCGGTGCCTCGGCCCCG
    TP/TK    (908)  ------------------------------------------------------------------------
    5'tkt    (574)  ------------------------------------------------------------------------
    3'tkt    (780)  CGTCAGTTTTGGCTGGGAAAAATACTCCCATGCGCACGTGGGCATGTCCGGCTTCGGTGCCTCGGCCCCG
Consensus   (1821)  CGTCAGTTTTGGCTGGGA AAATACTCCCATGCGCACGTGGGCATGTCCGGCTTCGGTGCCTCGGCCCCG
                    1891                                                                1960
    LmTKT   (1891)  GCGGGTGTGCTATACAAGAAGTTTGGAATTACCGTCGAGGAAGTGGTGAGGACGGGCCGTGAGTTGGCCA
    TP/TK    (908)  ------------------------------------------------------------------------
    5'tkt    (574)  ------------------------------------------------------------------------
    3'tkt    (850)  GCGGGTGTGCTATACAAGAAGTTTGAAATTACCGTCGAGGAAGTGGTGAGGACGGGCCGTGAGTTGGCCA
Consensus   (1891)  GCGGGTGTGCTATACAAGAAGTTTG AATTACCGTCGAGGAAGTGGTGAGGACGGGCCGTGAGTTGGCCA
                    1961                                                        2016
    LmTKT   (1961)  AGCGCTTCCCCGATGGCACGGCGCCGCTCAAGAACTCTTCATTCAGCAAGATGTAA
    TP/TK    (908)  ------------------------------------------------------
    5'tkt    (574)  ------------------------------------------------------
    3'tkt    (920)  AGCGCTTCCCCGATGGCACGGCGCCGCTCAAGAACTCTTCATTCAGCAAGATGTAA
Consensus   (1961)  AGCGCTTCCCCGATGGCACGGCGCCGCTCAAGAACTCTTCATTCAGCAAGATGTAA
```

Figure 3.22 continued: Alignment of the *L. mexicana TKT* gene sequence with the three individually amplified sections of the *TKT* gene.

Location (LmxTKT sequence)	DNA change LmxTKT→ LmxTKT assembled	Amino acid change LmxTKT→ LmxTKT assembled
213	A → G	-
351	C → T	-
693	G → A	-
823	G → A	A → T
1839	G → A	-
1916	G → A	G* → E

Table 3.7: DNA and corresponding amino acid changes of the full-length *L. mexicana TKT* gene compared to the assembled *TKT* gene. The location of the changes corresponds to the DNA sequence alignment (Figure 3.22). Out of six base pair changes, only two code for amino acid changes. The glycine (G*) residue is present in the *L. major* TKT amino acid sequence.

3.4. Southern Blotting

To determine the gene copy number of *TKT* present in the *Leishmania* genome, the technique of Southern blotting was used. Knowledge of the gene copy number is essential if experiments to attempt to produce null mutant parasites using gene replacement vectors are to be carried out. Figure 3.23 shows a restriction map of the *TKT* gene and the 5' and 3' flanking regions. Figures 3.24a and 3.24b show Southern blots probed with the 1 kb TPP-TKT gene fragment and with the full-length *TKT* gene (2 kb), respectively. The control lane contained 100 pg of the plasmid derived DNA containing *TKT* used to make the probe. This confirmed successful transfer of DNA to the filter and the specificity of the probe. The data suggests that *L. mexicana TKT* is a single copy gene. DNA digested with *Eco*RI, an enzyme not cutting the *TKT* gene, displayed a single band indicating that the *TKT* probe bound to only one digested fragment. *Pst*I cut the gene at three sites, 915 bp, 2163 bp and 2340 bp in Figure 3.23. The gene fragment seen was 1.3 kb that corresponded to the central fragment of the gene, smaller fragments are probably not visible under the conditions used. In Southern Blot 1 *Hind*III shows 2 bands at ~3 kb and ~6 kb, however *Hind*III does not cut this gene. The blot was therefore repeated using a *Hind*III digest of the *L. mexicana* DNA (Blot 2) and shows a single band of 8 kb indicating that the larger fragment in Blot 1 was the product of a labelling error in setting up the digestion. *Eco*RV digests *TKT* almost directly in half therefore two bands were clearly seen on Blot 1, representing the two separate fragments to which the probe bound.

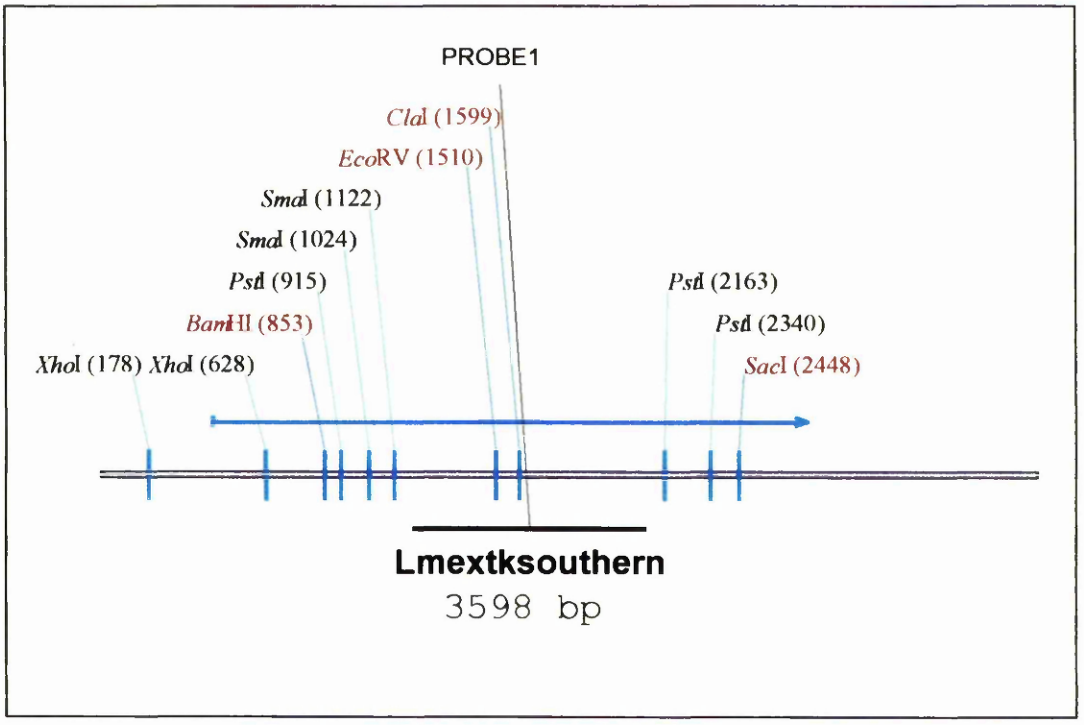


Figure 3.23: Location of the restriction sites present in the *L. mexicana TKT* gene ORF and flanking regions. The blue arrow indicates the *TKT* ORF, the grey line indicates the *TKT* gene and the flanking regions and the black line indicates the DNA used to probe southern blot 1 (PROBE 1). The numbers beside the restriction sites indicate the position of the restriction site in relation to the first known base of the 5' *TKT* flanking region. The restriction enzymes highlighted are those used to cut DNA for Southern blotting.

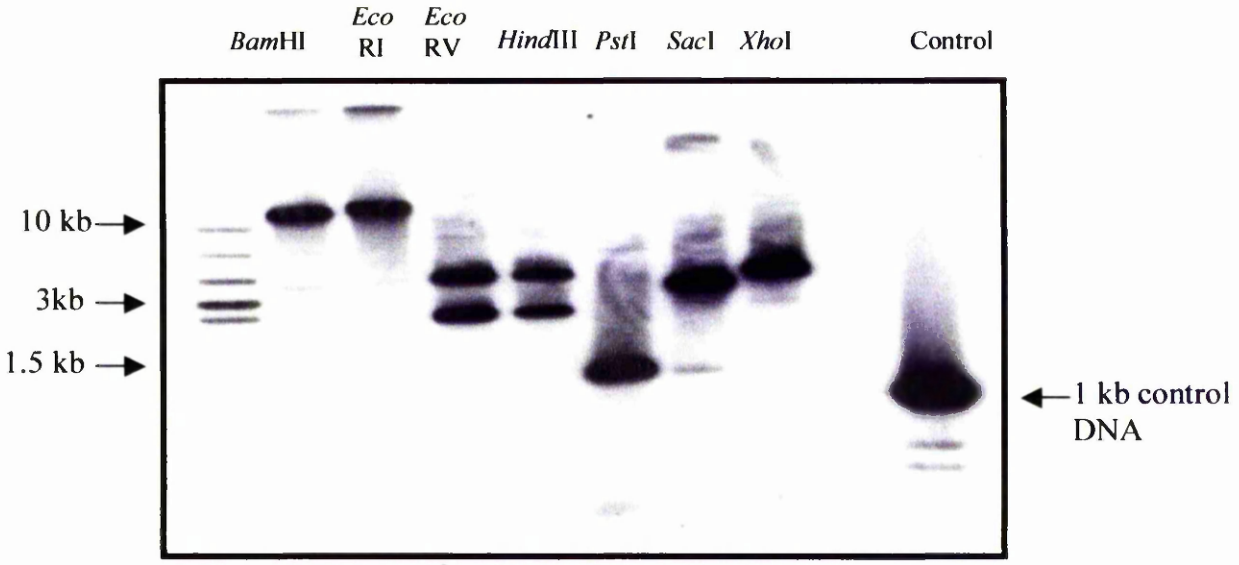


Figure 3.24a: Southern blot 1. The restriction enzymes used to cut the DNA are highlighted above the lanes. 100 pg of control DNA was used to ensure transfer and probe synthesis (Control lane). The 1 kb from *L. mexicana TKT* initially amplified was used to probe the blot. The 1 kb marker in the first lane displays bands of sizes indicated.

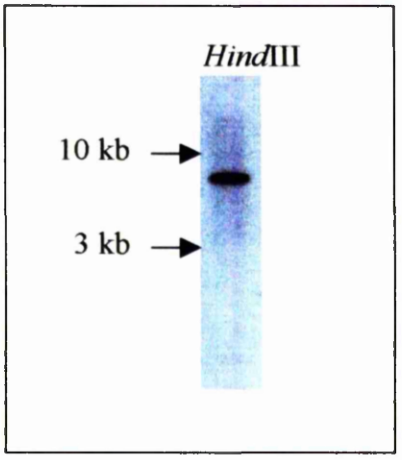


Figure 3.24b: Southern blot 2. *HindIII* was used to digest the DNA, as highlighted above the lane. The probe binds to one fragment of 8 kb, as indicated by the 1 kb marker ladder.

3.5. Cloning the 3' flanking region of the transketolase gene

The identification of the 3' flanking region of *TKT* provided valuable sequence from which constructs could be designed for gene replacement studies in *L. mexicana.* The cloning of the 5' flanking region is detailed in Section 3.8.

To clone and sequence the 3' flanking region of the *TKT* gene, the 4 kb sequence with homology to transketolase present in the *L. major* database, was used as template to design oligonucleotides (Figure 3.25 and Table 3.8). A low annealing temperature was used (52 °C), as it could not be presumed that *L. major* and *L. mexicana* DNA sequences would be identical. The sense primers were designed to the previously sequenced *TKT* DNA from *L. mexicana.*

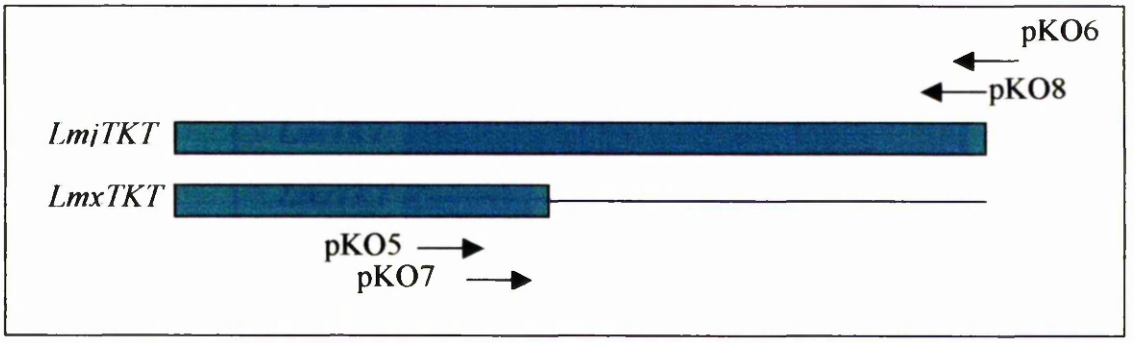


Figure 3.25: Schematic representation of the location of the primers used to amplify the 3' flanking region of the *L. mexicana TKT* gene. The primers pKO5 and pKO7 complement the *L. mexicana TKT* (*LmxTKT*) gene and pKO6 and pKO8 are antisense primers to the downstream region of the *L. major TKT* (*LmjTKT*) gene.

Primer name	Primer sequence 5' → 3'
pKO5	GAT CGT GTG TGC AGG GTG TGC
pKO6	CAT GCG CAC GTG GGC ATG TCC GGC
pKO7	AGG GTC TGC GCA GCG GGT CGC GT
pKO8	TCG GCC CCG GCG GGT GTG CTA TAC

Table 3.8: Primer sequences used to amplify the 3' flanking region of the *L. mexicana TKT* gene. The antisense primers were designed to sequence downstream to the putative *L. major TKT*, obtained from the *L. major* DNA sequence database (Sanger Centre).

To amplify the 3'-flanking region, primers pKO5 and 6 were used in the initial PCR, followed by a nested PCR reaction with primers pKO7 and 8. The PCR conditions used were identical to those used when amplifying the *TKT* gene. *L. major* DNA (obtained from Roderick Williams) was also used as a template in PCR as a control reaction. The nested PCR yielded a 700 bp and a 1.1 kb product (Figure 3.26, lane 2). Lane 3 in Figure 3.26 shows the product of the same primer combination, using *L. major* DNA. Both reactions amplified a product of 1.1 kb and the product from *L. mexicana* was isolated by gel elution.

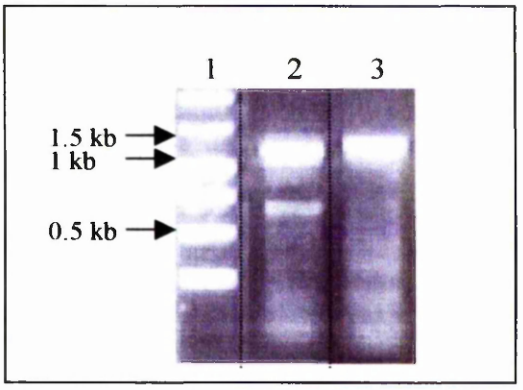


Figure 3.26: PCR products from the amplification of the *L. mexicana TKT* gene 3' flanking region. The 3' flank from *L. mexicana TKT* (Lane 2) is of similar size to the product amplified using *L. major* (Lane 3) using the same primer set. Both products are approximately 1.1 kb (Lane 1).

The 3' product was cloned into pGEM-T (pG3'FL) and sequenced (Figure 3.27). The sequence overlapped with the 3' region of the gene and contained the sequences of both primers used for nested PCR. This provided 903 bp of 3' flanking region. No ORF could be identified within this region.

pKO7

```
   1  TCGGCCCCGG CGGGTGTGCT ATACAAGAAG TTTGGAATTA CCGTCGAGGA AGTGGTGAGG
  61  ACGGGCCGTG AGTTGGCCAA GCGCTTCCCC GATGGCACGG CGCCGCTCAA GAACTCTTCA
 121  TTCAGCAAGA TGTAAGGTGA GCCGGTGCTC CGAAACGTGA GGAATGACAC CGGAAGAGGC
 181  GGGCAGGGAA GAGGGCGGAT GGGGGAGGAA GATGGTTGCT GGCGATGGCG GCCCCGCTTG
 241  CCCCTGACCT CTCTCTCCCC TTATAGAAAC TCTCTACCGT GGGCGCCATG GTCGACTACA
 301  CGCACAAAAC AACGCGTACA TGTTATGCAT GCCGATGCAA CGTAGTATGT CGGTGCCCGT
 361  TCGCACACAC GTGTGCGTGT GTTGGTCTAT CGTCGTCCCC ACCCCTACCC CCTCCTCCTT
 421  TGTCGCTCGC CGGTCACGGC TTGTTGCTAT GTCTCCTTGT CTATTTCCCC TTGTTGGGGC
 481  TGGTGCTGGA TTGATTTTTC CTCCTCTGTT ATGATCTTGC CTTGACTTGT GCGTGTCTCC
 541  CTCCCTCTCT GACACTCTAC CCACCCACCC CACTTCTTGC CGCCTCTTCA CTTTCTCTGG
 601  TGTGCTTCGT CCGGGGCGCG TTGGGGGGAT GCATGAGTGA CGTGGACACC TTATTTTATG
 661  CTTCGTCTTT CTTTTCATCA GTGCGGGAGT CGAACTGCAT GATGCATGCT CGCGTTGTGC
 721  GCCGGTATCG GAAGGGCAAG GAAGTCCTAT TGTGCCCGTG GCAGCGCCAT GGGCGTGCGG
 781  CTGCGTGTGC GTGAGCAGGT GTGTCAAGGT CTACTGAGGC GCACACTCAC ACGCGTTCGC
 841  GCACACACAG ACGCACGCCA GCGACGAACC AACGGGTGCC GGCGCGATGG AGTGGGTGAA
 901  GGTGTTTCCT CGCCTTCAGC AGAGAGGCCC TGCTCTCTGC CTGCCTCCCT CCCGCAGCCC
 961  TCCTCTCTCG CCCCTTTTTC CCTTTTATCT GACTTTCTTA TCCCATGTAC TGATGAACGC
1020  GACCCGCTGC GCAGACCCT
```

pKO8

Figure 3.27: DNA sequence of the 3' flanking region of the *L. mexicana TKT* gene. The primers pKO7 and pKO8 are highlighted in blue, and the putative stop codon is underlined. The *TKT* DNA sequence is highlighted in red.

3.6. Preparation of constructs to replace the *TKT* gene in *Leishmania mexicana*

The vectors used are described in Table 2.1. The constructs pGL520 and pGL345 were obtained from Dr. Hubert Denise, University of Glasgow and contained cysteine proteinase A (*cpa*) flanking regions. The constructs have been previously used to replace the *cpa* gene locus in *L. mexicana* (Souza *et al.*, 1994). The *cpa* flanking regions were directly replaced with *TKT* gene flanks to prepare the construct. pGL520 contains a Nourseothricin (SAT) selectable marker and pGL345 a Hygromycin B (HYG) selectable marker. 120 bp of the 5' flanking region and 800 bp of the 3' flanking region of *TKT* were cloned into the constructs. The small size of the 5' flanking region was not thought to be problematic during homologous recombination as only 200 bp of DNA has been shown to be required for integration into the *Leishmania enriettii* genome (Tobin and Wirth, 1992).

The vector for expression of the *TKT* protein in *Leishmania* was named pGL102 and was originally based on pX shuttle vectors described by LeBowitz (LeBowitz *et al.*, 1990). Details of these vectors are in Figure 3.28. The plasmids were ampicillin selective in *E. coli* and contained restriction sites appropriate for the DNA being cloned. These constructs also contained DHFR flanking regions to drive expression of the antibiotic marker.

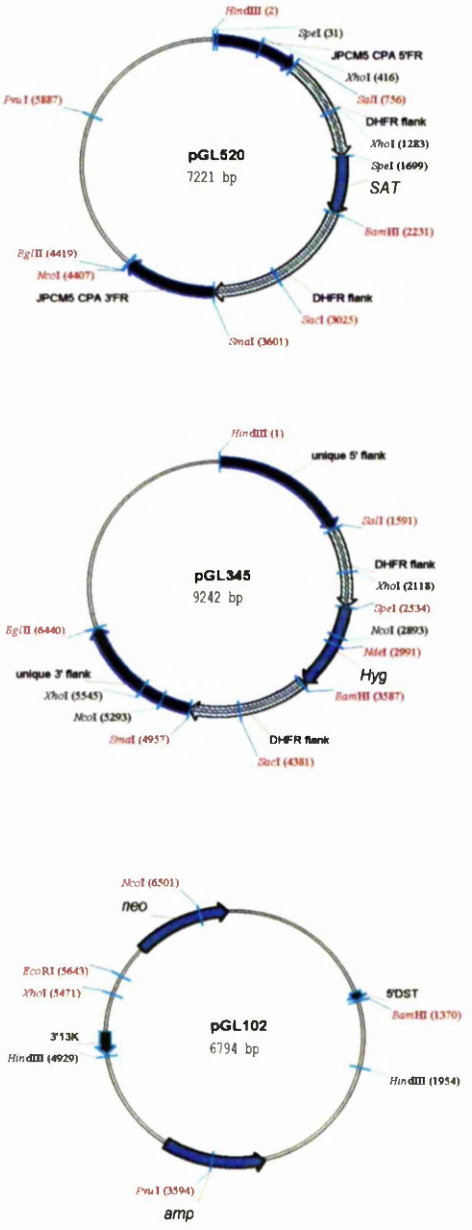


Figure 3.28: Graphic representation of vectors pGL520, pGL345 and pGL102. The constructs pGL520 and pGL345 are *Leishmania* replacement vectors used to knock-out targeted genes and replace them with antibiotic resistant markers. pGL520 has a nourseothricin resistance marker gene (SAT) and contains the 5' and 3' flanking regions of the CPA protein, (JPCM5 PA 5'FR and JPCM5 CPA 3'FR, respectively). pGL345 contains a hygromycin B resistance marker (Hyg) and contains CPA flanking regions (unique 5' flank and unique 3' flank). Both vectors contain multiple restriction sites, the location of these are indicated in brackets after the enzyme name. The DHFR flanking sequence is present directly 5' and 3' of the antibiotic marker in each of these constructs (DHFR flank). pGL102 is a *Leishmania* expression vector and contains a neomycin resistance marker indicated by a block blue arrow (neo). This vector contains a mini-exon addition sequence (5'DST) and a 3'flanking region of DHFR-TS (3'13K) to ensure translation of the gene. pGL102 also contains a gene encoding ampicillin resistance (amp). A variety of restriction sites are present in the plasmid and their location is indicated in brackets after the name of the enzyme.

3.6.1. Preparation of flanking regions from *L. mexicana TKT*

The flanking regions of the *L. mexicana TKT* gene were re-amplified from plasmids pUC5'SL and pG3'FL using primers containing appropriate restriction sites, so that the flanks could be cloned into digested constructs. The sequences of the oligonucleotides are shown in Table 3.9. The primers p5'FK1 and p5'FK2 were used to amplify the 5' flanking region and p3'FL1 and p3'FL2 were for amplification of the 3' flanking region. PCR amplified products were obtained by annealing the primers at various temperatures. The following cycle was used:

Pre-heat reaction:	60 seconds - 95 °C	
Melt and denature:	30 seconds - 95 °C	x 25 cycles
Anneal:	30 seconds – 55-65 °C	
Extension:	30 seconds - 72 °C	
Final hold:	10 minutes - 72 °C	

The PCR products obtained using varying concentrations of template DNA and annealing temperatures. The amplified bands were 120 bp for the 5' flank, and 800 bp for the 3' flank. These products were gel eluted, ligated into pGEM-T and transformed into *E. coli* strain JM109. Due to the small size of the 5' insert, diagnostic PCR was performed on plasmid DNA prepared from colonies following transformation to verify the ligation. These plasmids were digested with the appropriate restriction enzymes. The 5' flank was digested with *Hind*III and *Sal*I, and the 3' flank digested with *Sma*I (25 °C) followed by *Bgl*II. The insert was gel eluted having been cut from the vector, and used in subsequent ligations with digested constructs pGL345 and pGL520.

Primer Name	Primer Sequence 5' → 3'	Restriction site
p5'FK1	CGC AAG CTT AGT AGC CAC CCT	*Hind*III
p5'FK2	AGG TCG ACG GAG AGA GGC GAA	*Sal*I
p3'FK1	GCA CCC GGG CTC CGA AAC	*Sma*I
p3'FK2	AGGAGATCTGCGGGAGGGAGGC	*Bgl*II

Table 3.9: Oligonucleotide sequences to amplify the flanking regions of the *L. mexicana TKT* gene for cloning into replacement constructs. The restriction sites *Hind*III/*Sal*I and *Sma*I/*Bgl*II (highlighted in red) were incorporated into the primers used to amplify the 5' 120 bp flanking region and the 3' 800 bp flanking region for directional cloning into vectors pGL520 and pGL345.

3.6.2. Vector construction

E. coli transformed with plasmids pGL520 and pGL345 were initially streaked onto an LB amp plate, a colony selected and grown in LB amp medium. The vectors were digested with *Hind*III to check that the correct vector had been selected. 50 µg of *Hind*III digested vector was digested with *Sal*I to cut out the 5' flank unique to these constructs. The vector and 5' flank were ligated overnight at 4 °C and transformed into *E. coli* strain JM109. Individual colonies were selected, grown in LB-amp and the plasmid extracted. To determine whether the extracted plasmids contained the 150 bp insert, PCR was performed using the primers p5'FK1 and p5'FK2. Plasmids positive for the 5' insert were digested with *Sma*I and *Bgl*II in preparation for ligation with the 3' flank.

The 3' flanking regions specific to these constructs were excised from the plasmids. The constructs were gel purified and de-phosphorylated. The pre-prepared 3' flank was ligated with the plasmid overnight at 4 °C and the mix transformed into *E. coli* JM109. The plasmids were extracted from the resulting colonies and digested with *Bgl*II to determine the size. The constructed plasmids pGL520tkt and pGL345tkt are illustrated in Figure 3.29.

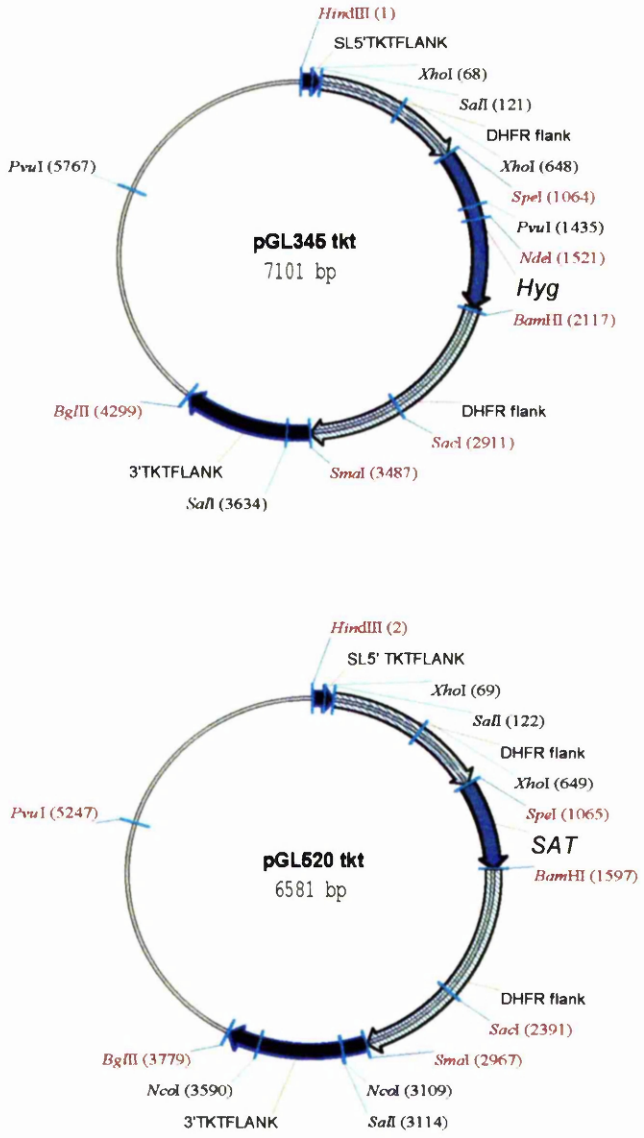


Figure 3.29: Graphic representation of vectors pGL345tkt and pGL520tkt. The plasmids contain the 5' and 3' flanking regions of *L. mexicana TKT* represented by black block arrows (SL5'TKTFLANK and 3'TKTFLANK. respectively). The plasmids are detailed in Figure 3.28.

Using diagnostic PCR and restriction enzyme digestions, the constructed vectors were confirmed to contain the flanking regions. The plasmids were digested with *Hind*III and *Bgl*II to create a linear fragment of DNA for transfection into *Leishmania* cells. Digested pGL520tkt releases a fragment of 3.8 kb that can be separated from the remaining plasmid by agarose gel elution. pGL345 excises an insert of 4.5 kb which can also be separated from the remaining DNA (Figure 3.30). 5 μg of linear DNA was required per transfection, therefore 50-100 μg of DNA was initially prepared, digested and gel eluted.

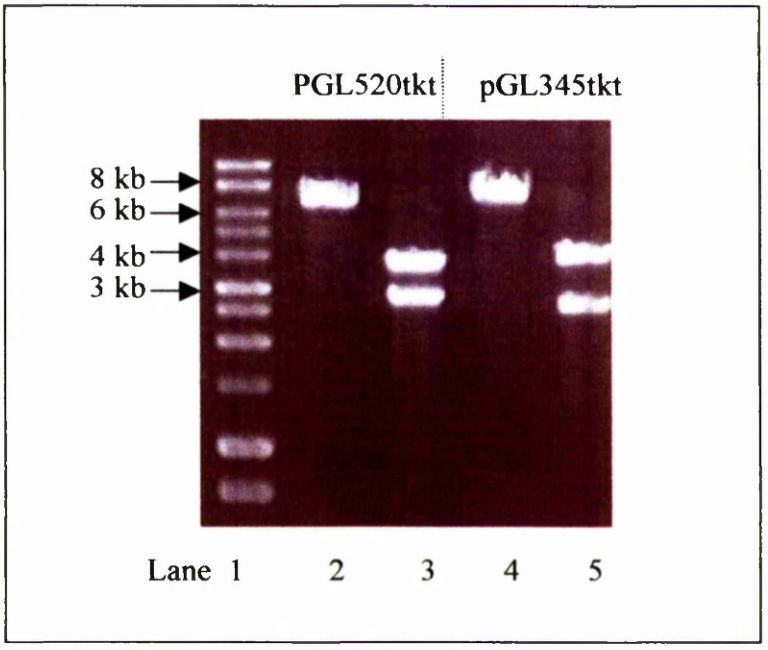


Figure 3.30: Agarose gel of digested constructs pGL520tkt and pGL345tkt. Lanes 2 and 4 represent pGL520tkt and pGL345tkt cut with *Bgl*II to produce linear products of approximately 7 kb (Lane 1 represents the 1 kb marker). Lanes 3 and 5 are the plasmids cut with *Bgl*II and *Hind*III to produce two bands, the upper of the two represents the construct required for integration experiments, and the lower band the discarded plasmid backbone.

3.6.3. Preparation of the re-expression vector

To re-express the *TKT* gene in *L. mexicana*, a pX shuttle vector (LeBowitz *et al.*, 1990) named pGL102 was used (detailed in Figure 3.28). Primers were designed to amplify *L. mexicana TKT* and incorporated *Age*I and *Bgl*II restriction sites (Table 3.10) for cloning into the expression vector pGL102. pGL102 has cloning sites *Bam*HI and *Age*I, however the *L. mexicana TKT* gene contained a *Bam*HI restriction site and therefore *Bgl* II was used.

Primer name	Sequence 5'→ 3'	Restriction site
p5'TktBg	CGA AGG TAG ATC TAT GGC CTC CAT TGA G	*Bgl*II
p3'TktAg	GAT CCT ACC GGT TTA CAT CTT GCT GAA T	*Age*I

Table 3.10: Primer sequences to amplify the *L. mexicana TKT* gene for cloning into the re-expression vector pGL102. Restriction sites *Bgl*II and *Age*I (highlighted in blue) were incorporated into the primer sequences for directional cloning into the vector pGL102.

The gene was PCR amplified from pET-16blmtkt and was ligated into pGEM-T then transformed into *E. coli* strain JM109. The resulting colonies were grown in LB-amp and the plasmid extracted. The plasmids were cut with *Age*I and *Bgl*II, and produced two bands of 3 kb and 2 kb in size. The 2 kb insert was gel eluted cloned into pGL102 pre-digested with *Age*I and *Bgl*II (Figure 3.31).

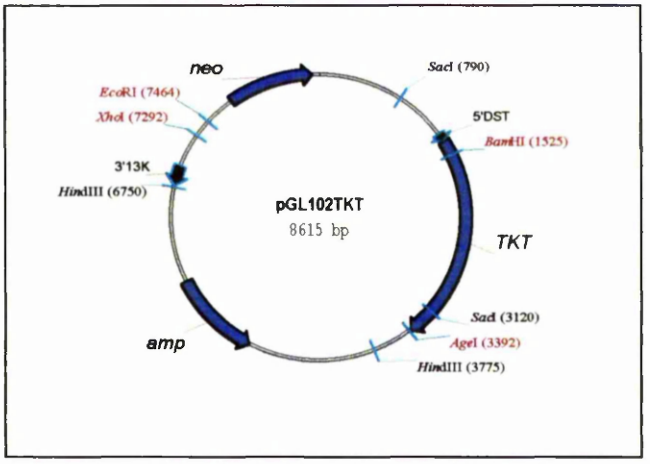


Figure 3.31: Graphic representation of the *L. mexicana TKT* expression vector pGL102tkt. The *TKT* gene is represented by a block blue arrow indicating the direction of expression. The plasmid is detailed in Figure 3.28.

3.7. Transfection and selection of antibiotic resistant *Leishmania*

Two transfections were performed using *L. mexicana* promastigotes with pGL502tkt and pGL345tkt, selecting in medium and on agar plates containing 25 μg/ml of nourseothricin (SAT) and hygromycin B (HYG), respectively. Any colonies seen on the plate were transferred into medium immediately and after 2 sub-passages, transferred into medium containing 50 μg/ml of drug.

The EC_{50} of the drugs used have been previously determined (Cruz *et al.*, 1991; Joshi *et al.*, 1995), however wild type (WT) *L. mexicana* were cultured in medium containing SAT and HYG at 25 μg/ml and 50 μg/ml individually and combined to ensure that the drugs were toxic at these concentrations. After 5-7 days WT parasites were all dead using 25-50 μg/ml of antibiotic in comparison to uninfected cells that continued dividing beyond day 7 (Figures 3.32a and b).

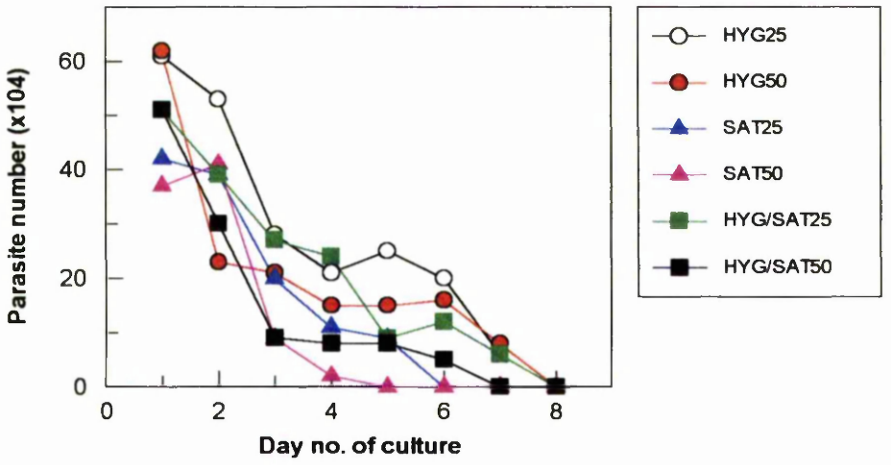
A summary of the number of colonies obtained from each transfection is shown in Table 3.11. *L. mexicana* appeared on plates as transluscent round colonies after

approximately 10 days, and more continued to appear for up to 20 days. Several colonies were able to grow at a normal rate when transferred to medium containing drug. However diagnostic PCR using primers designed to amplify either the knockout construct, or the WT gene, failed to amplify the predicted fragment. This was possibly due to the large size of the region being amplified. To determine the specificity of the replacement, a southern blot could be performed. The culture growing in both SAT and HYG antibiotics did contain a transketolase gene, as confirmed by PCR. PCR to amplify the selectable marker was positive in colonies growing in antibiotic and not in the wild-type. However, the experiments performed were unable to determine if the construct had recombined in the correct place within the genome. Further work would have to be done to determine where the construct had entered the genome to, as it could not be located to the transketolase locus.

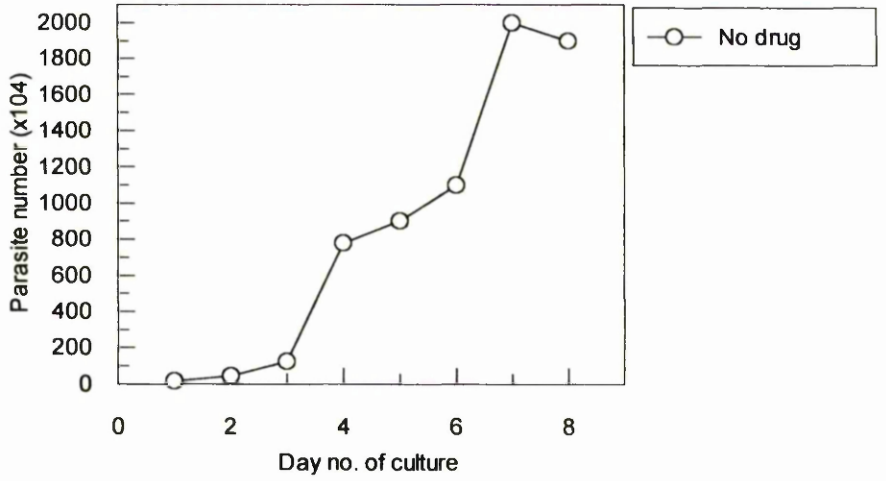
This initial attempt to generate TKT null mutants failed. One possible reason was that the 5' flanking region was too short to enable high efficiency targeting. Therefore more 5' sequence was amplified from the genome (Section 3.8).

Experiment number and selectable marker	Number of colonies on plate	Number of colonies growing in medium + drug (50 μg/ml)
SAT 1	42	2/8
SAT 2	12	5/12
HYG 1	1	1/1
HYG 2	20	6/10
SAT/HYG 1	5	1/5

Table 3.11: Numbers of resulting *L. mexicana* cultures growing in medium containing antibiotic subsequent to transfection with SAT/HYG resistant constructs. The individual *Leishmania* colonies were isolated from agar plates after transfection with the SAT or HYG baring constructs and grown in medium containing the relevant antibiotic. The number of cultures growing in the medium out of the total number put into medium was recorded.



GraphA



GraphB

Figure 3.32: Graph of wild type *L. mexicana* in medium containing antibiotic compared to medium without drug. The parasites were grown in medium containing hygromycin B, nourseothricin, a combination of both drugs and in normal medium. Hygromycin B and nourseothricin were added into the medium at 25 μg/ml and 50 μg/ml (HYG25, HYG50, SAT25 and SAT50 respectively). The growth curves with the combined drugs at 25 μg/ml and 50 μg/ml are labelled HYG/SAT25 and HYG/SAT50, respectively (GraphA). As a control, parasites were grown in medium without antibiotic (GraphB).

3.8. Amplification of the 5' flank of *TKT*

The appearance of the transketolase 5' flank from the *L. major* genome database assisted in the sequencing of this intergenic region. Primers were designed to the upstream 5' region of this sequence in order to amplify the 5' flank from *L. mexicana* DNA. The 5' flanking region was amplified using primers pKO1 and pKO2, nested with pKO3 and pKO4 (Figure 3.33). The sequences of the primers are detailed in Table 3.12. Nested PCR amplified a band of 750 bp in size with an annealing temperature of 60 °C and extension for 2 minutes for 30 cycles. This product was cloned into pGEM-T (pG5'FL) and sequenced (Figure 3.34).

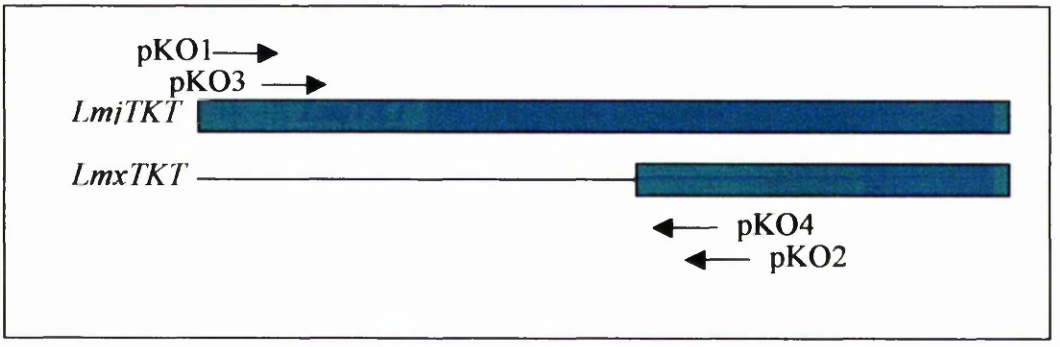


Figure 3.33: Schematic representation of the location of the primers used to amplify the 5' flanking region of the *L. mexicana TKT* gene. The primers pKO1 and pKO3 complemented the *L. major TKT* gene (*LmjTKT*) 5' flanking region and pKO4 and pKO2 are antisense primers designed to the *L. mexicana TKT* gene (LmxTKT).

Primer name	Primer sequence 5' → 3'
pKO 1	ACA CAC GCA CGC ACC TCA TAC AGC
pKO2	CAT GCC CAT CGG CGT GCC TGG
pKO3	TGC CAA TGC CAA ACC ACT TCG CGT
pKO4	CGC CGC GAG GCA GCG GAT GCA

Table 3.12: Primer sequences used to amplify the 5' flanking region of the *L. mexicana TKT* gene. The sense primers were derived from the 5' region of the 4 kb contig containing the putative *TKT* gene from *L. major*.

pKO3

```
  1  TGCCAATGCC AAACCACTTC GCGTGGTGAC AGGGCCAAGC ACCTACAAGG TAGGGGAGGT
 61  CAGAGCGACG CACCACCACT AATGCCGGCG GCCAGGTCCT GGACGGGCGC TGCGTCGGAG
121  CGACCTGCGA CGGCGAGCAC GCCTGTGCGC CGCCCACATG ACCAGGCAGT GTGTCGACGT
181  GACTCGAGCG CGGCTCCGCC GGCCCTCACA CTGCCCACTG GGGCGAGGGG TGCGCATGAG
241  CCGCCGCGAA GCGGGTGGGT GGGTGGGTGG AGGGTTTGAG GCCGACGCCA TGCTCCGATG
301  GCTGAGGCGG CGCATTGCCC TGGCGCGTGT GTCGACGGAT GCTTCGCACC ACGCGATGGG
361  TGGCGTGTGA CCTCACGCTG TGTGTGCGGC AGAGAAAATC GGACACACGT TGAGTAAGCG
421  AAAACGTCGA TGCTCTCGAC CTCGTACGCC TGCCCGTGCC CCCTTCTTGC TCTCGCTCTT
481  CAGCGGCTAG CACTCCTCCC TCTCTCTACC CCCGTCTCTC TGTTCCCGCG AGCCGCACCC
541  CCCCCACCAC CACCACCGCA AAGCGCAACA CCAGTAGCCA CCCTACGCTC TCACCAGTAT
601  ACCAAAGGAA AGCCTCCCTC TTCCTCTCGT TTCTCGAGCC CGTCATCACA CCGTAGCACT
661  CGCCCCCTCC GCCTTCGCCT CTCTTCGTCG CCCTCACACA CAAGCACAAT GGCCTCCATT
721  GAGAAGGTGG CAAACTGCAT CCGCTGCCTC GCGGCG
```

pKO4

Figure 3.34: DNA sequence of *L. mexicana TKT* 5' flanking region. The primers pKO3 and pKO4 are highlighted in blue and the start codon ATG is underlined, with the *TKT* gene sequence in red text. The spliced leader acceptor site, AG is underligned and the polypyrimidine tracts, possibly involved in polyadenylation recognition of an upstream gene, are highlighted with a dotted underline.

```
                    1                                                  50
5'FlankcDNA     (1) AGTATCAGTTTCTGTACTTTATTGCGCAACACCAGTAGCCACCCTACGCT
5'Flank DNA     (1) CCCCCCACCACCACCACCGCAAAGCGCAACACCAGTAGCCACCCTACGCT
  Consensus     (1)      CA     C   AC    A  GCGCAACACCAGTAGCCACCCTACGCT
                   51                                                 100
5'FlankcDNA    (51) CTCACCAGTATACCAAAGGAAAGCCTCCCTCTTCCTCTCGTTTCTCGAGC
5'Flank DNA    (51) CTCACCAGTATACCAAAGGAAAGCCTCCCTCTTCCTCTCGTTTCTCGAGC
  Consensus    (51) CTCACCAGTATACCAAAGGAAAGCCTCCCTCTTCCTCTCGTTTCTCGAGC
```

Figure 3.35: Alignment of the *L. mexicana TKT* 5' flank cDNA and DNA. The primer used to amplify the cDNA (bold text) complemented the 3' region of the conserved spliced leader sequence, therefore the spliced leader recognition site, AG, present on DNA can be identified by alignment (underlined).

709 bp of 5' flanking sequence were obtained from this information. The 5' UTR contained 65.8 % GC residues. The 5' flank contained an AG splice acceptor site (Figure 3.34) 145 bp upstream of the start codon ATG. This site was identified as the spliced leader acceptor site when comparing the cDNA with the DNA from the 5' flank of *TKT* (Figure 3.35). Upstream to this AG, there is a polypyrimidine tract of 30 nucleotides in length (90 % T/C rich) that may potentially be important in accurate polyadenylation of upstream transcripts and spliced leader recognition.

The *L. mexicana* and putative *L. major TKT* flanking regions have been checked for identity with other proteins using BLASTX, and no homology has been found. Therefore it is not yet known which genes lie adjacent to *TKT* in the *Leishmania* genome.

3.9. Discussion

The *E. coli tktA* has been cloned with a histidine tag present on the N-terminus of the over-expressed TK protein. The soluble enzyme was active and used subsequently as a control for transketolase assays and for the production of xylulose 5-phosphate and DAHP.

The *TKT* gene from *L. mexicana* has been cloned and sequenced using a variety of molecular techniques. The gene was cloned in three fragments to obtain the initial sequence. Conserved sequences present in transketolases from other organisms were used as template DNA to amplify the middle region of the gene from *L. mexicana* genomic DNA. The spliced leader, present on mRNA of all transcripts in trypanosomatids, was used to amplify the 5'-coding region of the transketolase gene by designing nested primers and using PCR to amplify the area from cDNA. The final 3'-coding region of the gene was obtained through the *Leishmania major* genome sequencing database.

Translation of the *TKT* DNA sequence led to the identification of the highly conserved thiamine pyrophosphate and transketolase motifs, and the discovery of the putative peroxisome targeting sequence (PTS) at the C-terminal of the protein. This finding led to an investigation of the localisation of transketolase within the *Leishmania* cell (Chapter V). Sequence from chromosome 24 of *L. major* had similarity to transketolase and was used to complete the gene sequence and also to clone the 5'- and 3' regions flanking the gene. The 5' flanking region exhibits an AG spliced leader acceptor site and an upstream polypyrimidine tract that may possibly be involved in accurate polyadenylation of the upstream construct and spliced leader recognition. The AG spliced leader recognition site has recently been shown to be coupled with an

upstream polypyrimidine tract, that is necessary for accurate polyadenylation of the gene 5' to the UTR of DHFR-TS in *L. major* (LeBowitz *et al.*, 1993), procyclin in *T. brucei* (Hug *et al.*, 1994; Schürch *et al.*, 1994) and tubulin in *T. brucei* (Matthews *et al.*, 1994).

Southern blotting suggested that *TKT* is a single copy gene and therefore experiments to knock out *TKT* from *L. mexicana* were initiated. Constructs were prepared to replace the *TKT* gene in the *L. mexicana* genome with antibiotic selection by homologous recombination of 5' and 3' flanking regions. However, time limitations have not permitted the study to be completed. A *TKT* re-expression vector was also prepared, and this may be required for future experiments making null mutants.

Chapter IV

Expression and characterisation of *Escherichia coli* and *Leishmania mexicana* recombinant transketolase

4.1. Introduction

The cloning of the *E. coli* and the *L. mexicana* transketolase genes described in chapter III enabled the production and purification of recombinant proteins. A bacterial expression system using an inducible RNA polymerase T7 promoter was used to produce both proteins. The proteins were efficiently purified using an immobilised metal affinity column, producing a large quantity of recombinant histidine-tagged protein. The specific activity was determined for both transketolase proteins. The native transketolase enzyme activity also was measured from *L. mexicana* promastigotes crude lysates.

The pure *E. coli* TKA was used to synthesise xylulose 5-phosphate (X5P) and 3-deoxy-D-arabino-heptulosonate-7-phosphate (DAHP), and was also used as a control transketolase when developing assay systems. X5P is a substrate for transketolase and is used to measure the enzyme activity of the protein. During this study, commercial production of X5P ceased. The substrate was prepared in the lab using an enzymatic biotransformation, previously determined by Zimmermann (Zimmermann *et al.*, 1999). This allowed for further characterisation of the *L. mexicana* TKT protein, including specific activity, pH optimum, storage conditions and the kinetic analysis of the enzyme in terms of the affinity of binding to the substrate ribose 5-phosphate.

The primary *L. mexicana* TKT protein sequence was analysed with reference to multiple transketolases from other organisms that have been identified and sequenced.

Literature and database searching identified 41 transketolase amino acid sequences. Of these, 22 sequences are putative as they are derived from raw genome sequencing data. The alignment of these sequences with the *L. mexicana* TKT amino acid sequence led to the analysis of the transketolase protein family. To determine if the *L. mexicana* TKT would be a potential drug target, a tertiary structure would be helpful. The recombinant *L. mexicana* TKT protein is currently undergoing crystallisation trials. A direct comparison of *L. mexicana* TKT with *S. cerevisiae* transketolase1 showed that the two proteins have a high sequence identity, and that the *L. mexicana* TKT contains all previously determined invariant residues.

4.2. Transketolase enzyme activity in *Leishmania mexicana*

Transketolase enzyme activity was measured in crude lysates of *L. mexicana* promastigotes using ribose 5-phosphate (R5P) and xylulose 5-phoshate (X5P) as substrates (Datta and Racker, 1961; Smeets *et al.*, 1971). A control measurement was taken without the substrate R5P, as NADH oxidation within the parasite cells depleted the NADH present in the assay. This background activity was subtracted from the rate to obtain a net activity. Using the Bio-Rad kit based on the method of Bradford, 2 $x10^9$ *L. mexicana* promastigotes were found to contain 5 – 8 mg protein. The specific activity of TKT in *L. mexicana* was determined as 58.75 ± 11.75 nmol/ min/ mg cell protein (Table 4.1). The specific activity of transketolase in *Leishmania* species has been previously determined and ranges from 6 – 12.4 nmoles/min/mg of cellular protein (Martin *et al.*, 1976). *L. mexicana* promastigotes had a specific activity of 12.4 nmoles/min/mg of cellular protein in this study (Martin *et al.*, 1976). The specific activity recorded in this study was slightly higher in comparison to the previously recorded value, and could be due to the difference in substrates used to assay the enzyme. The specific activity of transketolase in *Trypanosoma brucei* was 50.1 ± 10.8 nmol/min/mg cell protein (Cronin *et al.*, 1989).

Mean rate of transketolase activity deduced (ΔAU_{340}/min)	**Transketolase activity (nmol.min^{-1}.ml^{-1})**	**Mean specific activity (nmol.min^{-1}.mg^{-1}.cell protein)**
$2.05 \times 10^{-3} \pm 0.45 \times 10^{-3}$	$3.3 \times 10^{-2} \pm 0.74 \times 10^{-2}$	58.75 ± 11.75

Table 4.1. The specific activity of transketolase in *L. mexicana* cell lysates using the substrates R5P and X5P. The mean rate (n=4) without the acceptor substrate R5P was subtracted from the mean rate with R5P to obtain a rate recorded in absorbance units /minute (ΔAU_{340}/min). The mean specific activity was measured in nmol/min/mg of cell protein. Each assay utilised 56 µg of protein from *L. mexicana*.

4.3. Overexpression of *E. coli* transketolase

E. coli BL21(DE3) was used to host the expression of *E. coli* TKA. Three colonies transformed with plasmid pET16bEctkt (see Figure 3.2) were grown in LB amp and half the culture was induced to express TKA protein using 0.4 mM IPTG for four hours at 37 °C. The culture was lysed and the soluble extract was run on an 8 % SDS-PAGE gel (Figure 4.1). Cultures one and three exhibited over-expression of a protein at approximately 75 kDa in the soluble fraction compared to the un-induced cultures that were grown under the same conditions with no IPTG. The published *E. coli* recombinant TKA had an estimated molecular weight of 73 kDa (Sprenger *et al.*, 1995) therefore the over-expressed protein was considered to be the correct molecular mass (Figure 4.1).

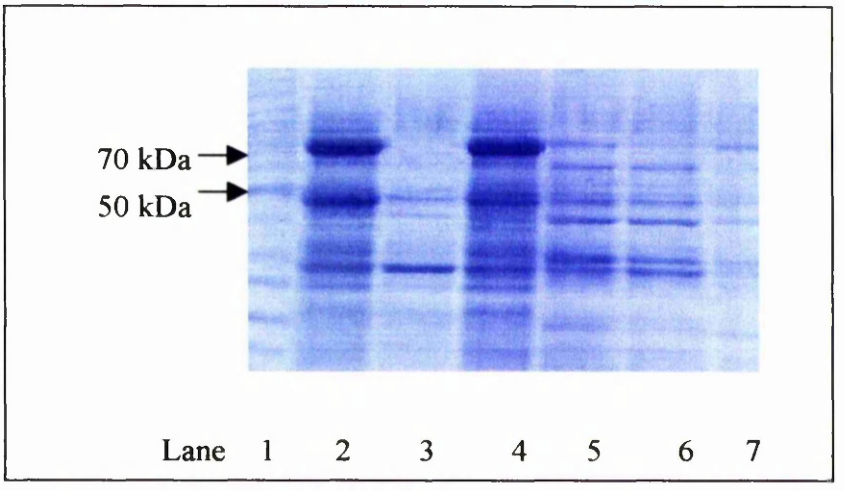


Figure 4.1: SDS-PAGE gel of induced and un-induced *E. coli* BL21(DE3) containing the vector pET-16b *E. coli* TKA. The induced cultures 1-3 are in lanes 2-4 and the un-induced cultures are in lanes 5-7. Lane 1 is the molecular weight marker, with 50kDa and 70 kDa highlighted to show the approximate size of the protein over-expressed at 75 kDa.

4.3.1. Purification of *E. coli* transketolase

The *E. coli* recombinant TKA was expressed as an N-terminal 6-histidine tagged protein. Cloning of the gene encoding this enzyme is described in Section 3.2.1. The protein was purified using a nickel affinity column on a BioCAD 700E Workstation (PE Biosystems) with the assistance of Alan Scott, University of Glasgow. The cell pellet from 100 ml of a culture induced at 37 °C for four hours was sonicated and the soluble fraction extracted and suspended in 5 ml of lysis buffer (0.5 M NaCl, 0.1 M Tris-HCl, pH 8.0) (see Section 2.3.6.1. for details). 10 µl of each fraction of the purification was run on an 8 % SDS-PAGE gel so that the purification could be assessed (Figures 4.2a and b). The purification profile for the nickel-ion affinity column is shown in Figure 4.3. The protein was eluted from the column at ~ 400-500 mM imidazole according to the A_{280} trace and the SDS-PAGE gel, as fractions 10-12 show increased protein concentrations. Fractions 8-14 were pooled and dialysed overnight into 1 L of Tris buffer (100 mM, pH 7.5) yielding ~ 3.2 mg of pure *E. coli* TKA from 100 ml of culture.

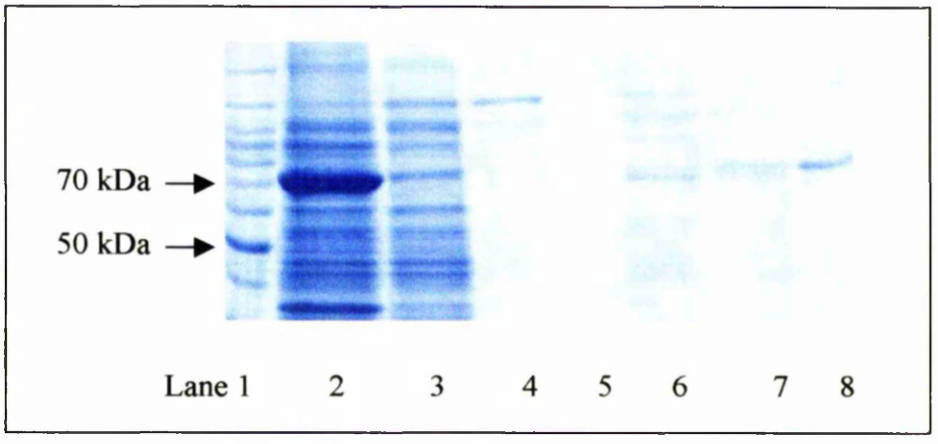


Figure 4.2a: SDS-PAGE gel of *E. coli* TKA purification using a nickel affinity column. The soluble fraction loaded onto the column is in Lane 2, the flow through (10 µl/ 18 ml) is in Lane 3, the 20 mM imidazole wash (10µl/ 15 ml) is in Lane 4 and fractions 4 – 7 (10 µl/ 1.7 ml) are in Lanes 5 – 8 and correspond to Figure 4.2c of increasing imidazole concentration.

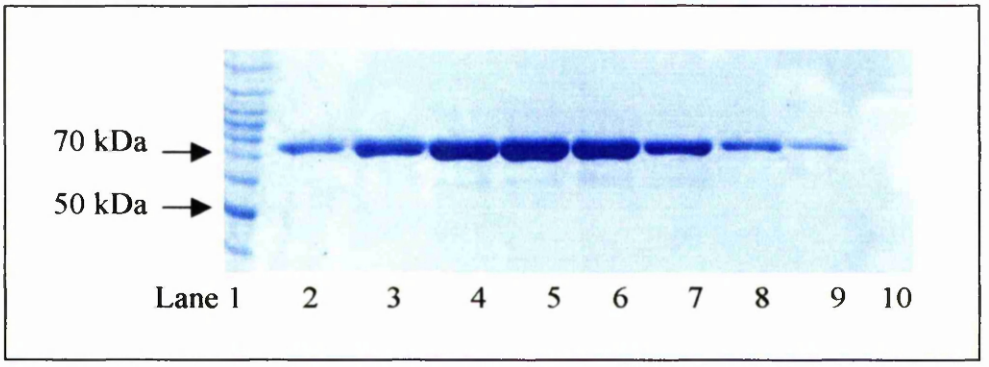


Figure 4.2b: The elution of *E. coli* TKA with increasing concentrations of imidazole. The fractions 8-16 (according to Figure 4.2c) are in Lanes 2-10 (10 µl/1.7 ml), with the molecular weight marker (Lane 1) showing the over-expressed protein to be approximately 75 kDa.

4.3

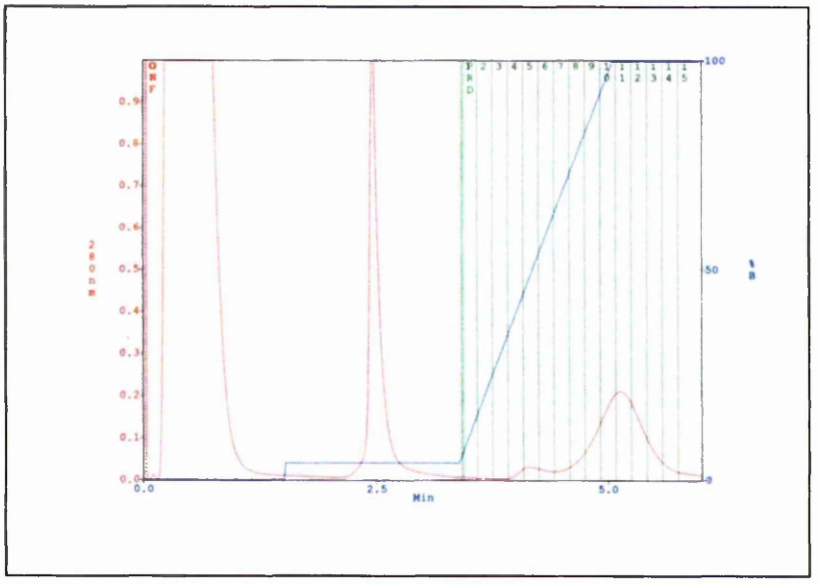


Figure 4.3: The purification profile of *E. coli* TKA protein. The blue line represents the imidazole concentration and the gradient used to elute the protein is from 20 mM to 500 mM (shown as percentage B, with 500 mM = 100 %). The red line indicates the absorbance at 280 nm and was a measurement of the amount of protein present. The fractions collected during the imidazole gradient and elution of the protein are represented in green.

The specific activity measured using R5P and X5P as substrates was 214 nmoles/min/mg of protein (0.214 U/mg). The *E. coli* TKA expressed without the histidine tag had a specific activity of 50 U/mg (Sprenger *et al.*, 1995). The reduction in specific activity may be due to the N-terminal histidine tag or due to incorrect folding in this particular system. The crystal structure of the transketolase from *S. cerevisiae* has shown that the N-terminal domain interacts with the cofactor TPP and substrates in the active site (Lindquist *et al.*, 1992; Nikkola *et al.*, 1994). The human transketolase has been expressed using an N-terminal histidine tag protein and had a specific activity of 13 U/mg subsequent to pre-incubation with the cofactors TPP and $MgCl_2$ (Schenk *et al.*, 1998 (b)). The specific activity of the *E. coli* TKA preparation may have a low specific activity due to the assay system or the preparation may be partially inactive. Approximately 30 mg of pure enzyme was obtained from one litre of culture. *E. coli* TKA was used during this study to prepare X5P and DAHP, and used as a positive control for enzyme assays.

4.4. Expression of *L. mexicana* recombinant transketolase

E. coli BL21(DE3) containing the expression plasmid for *L. mexicana* TKT (pET16blmtkt) was grown in 100 ml LB amp and half the culture induced with 1 mM IPTG at 37 °C for 4 hours. The cells were harvested and analysed on an SDS-PAGE gel. No over-expression of the TKT protein was seen under these conditions when compared to the un-induced culture. This experiment was repeated using various concentrations of IPTG (0.1 mM-1 mM), however no expression was seen. Using the same range of IPTG concentrations, induction was then attempted using a lower temperature of 25 °C (Figure 4.4). An over-expressed protein of ~ 72 kDa was induced

which corresponded to the approximate size of a transketolase. No expressed protein was evident in the control culture, without IPTG. The *L. mexicana* TKT protein was expressed most abundantly at 0.4 mM IPTG for 16 hours.

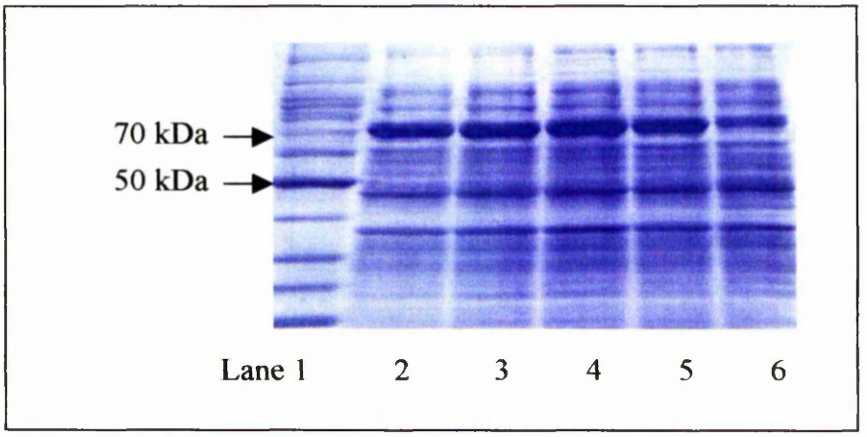


Figure 4.4: SDS-PAGE gel of the recombinant *L. mexicana* TKT protein induced at 25 °C using various concentrations of IPTG. The IPTG concentration used are 1 mM (lane 2), 0.5 mM (lane 3), 0.4 mM (lane 4) and 0.1 mM (lane 5). An un-induced control is shown in lane 6. The molecular weight markers (lane 1) show the over-expressed protein to be of approximately 72 kDa in size.

A 200 ml culture was grown in LB amp and half of the culture was induced at 25 °C with 0.4 mM IPTG. The cells were harvested, resuspended in 5 ml of buffer (Tris 0.1 M, pH 8) and lysed by sonication (see Section 2.3.6.1. for details). The soluble and insoluble fractions were separated by centrifugation. The over-expressed protein was present exclusively in the insoluble fraction of the cells, as seen on SDS-PAGE gel (Figure 4.5). Purification of proteins from inclusion bodies is much more problematic in comparison to soluble proteins due to the re-folding process required, therefore a number of other conditions were attempted to gain soluble protein.

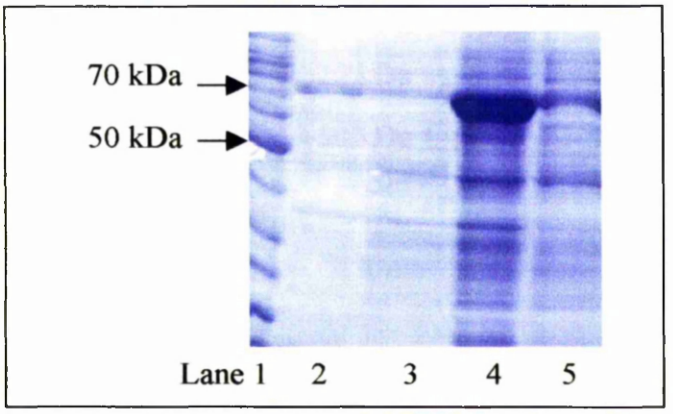


Figure 4.5: SDS-PAGE gel of insoluble and soluble fractions of culture induced at 25 °C. The induced and un-induced cultures are in lanes 2 and 3 respectively. Lanes 4 and 5 represent the insoluble and soluble fractions 20 times concentrated as the 100 ml culture was sonicated in 5 ml of buffer. The over-expressed protein was present in the insoluble fraction of the lysate and was approximately 70 kDa (according to the molecular weight markers in lane 1).

Several different lysis buffers were used during sonication in order to vary buffering conditions, one of which may be more favourable for maintaining the solubility of the protein. The basic iso-osmotic buffer used was Tris (0.1 M) and NaCl (0.5 M), pH 8, and the detergents added were Triton X-100 (1 % (v/v)), Tween-20 (1 % (v/v)) and CHAPS (1 % (v/v)). However, under all buffering conditions the protein remained insoluble, with the protein aggregating.

Another method to invoke conditions to produce a soluble TKT protein was to reduce the temperature of induction to 15 °C. The cells were induced with 0.4 mM IPTG at 15 °C for 16 hours, sonicated and separated into the soluble and insoluble phase under these conditions. Approximately 50 % of the over-expressed protein was present in the soluble fraction (Figure 4.6).

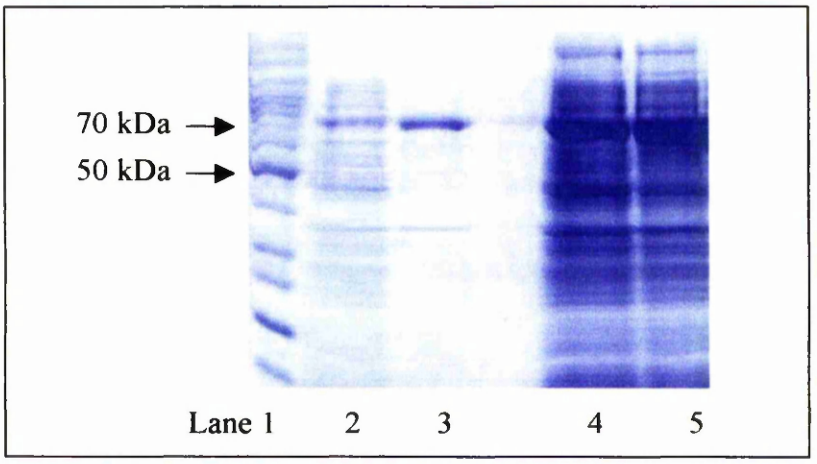


Figure 4.6: SDS-PAGE gel showing induction of the recombinant *L. mexicana* TKT protein at 15 °C in the soluble and insoluble fractions. The un-induced and induced cultures are in Lanes 2 and 3 respectively. Lanes 4 and 5 show that the insoluble and soluble fractions both contain the over-expressed TKT protein.

To determine if more than 50 % of the protein could be soluble, various IPTG concentrations and two incubation times were tested. Figures 4.7a and 4.7b show the expression of transketolase after 4 and 16 hours (respectively) of incubation with 0, 0.1, 0.2 and 0.4 mM IPTG. Substantially more protein was over-expressed after 16 hours of incubation with IPTG. Inducing the cells with 0.1 mM IPTG was just as efficient as induction with 0.4 mM. During the prolonged incubation of 16 hours at 15 °C, the TKT protein was over-expressed when IPTG was not present. This could be due to the prolonged incubation conditions used to express the TKT as soluble protein and may have been caused by a leaky promoter. To prepare the sample for purification, the cells were induced overnight at 15 °C at 0.1 mM IPTG and the soluble fraction extracted.

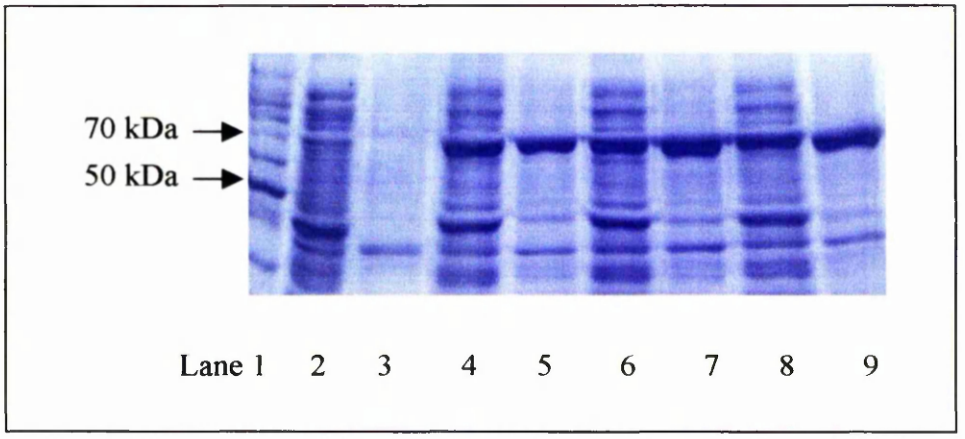


Figure 4.7a: Induction of the *L. mexicana* TKT recombinant protein at 15 °C for 4 hours using various concentrations of IPTG. The un-induced control culture fractionated into insoluble and soluble fractions are shown in Lanes 2 and 3. The varying IPTG concentrations are 0.1 mM (Lane 4 and 5), 0.2 mM (Lanes 6 and 7) and 0.4 mM (Lanes 8 and 9). The first of the lanes show the insoluble fraction and the second the soluble fraction. The over-expressed protein is approximately 70 kDa according to the molecular weight marker (Lane 1).

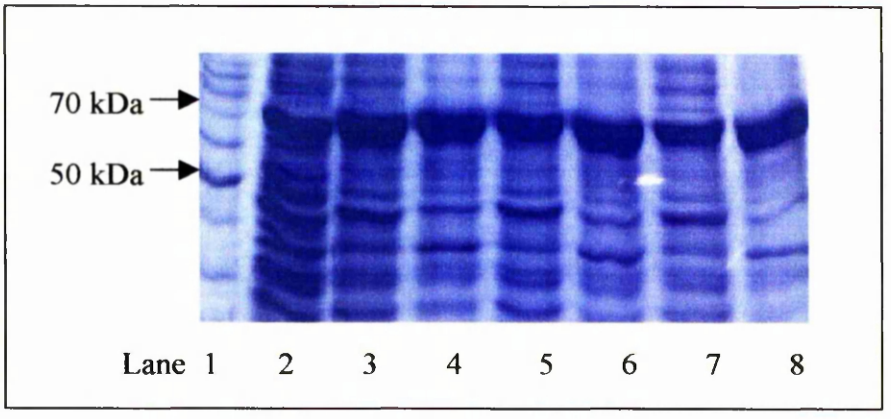


Figure 4.7b: Induction of the *L. mexicana* TKT protein at 15 °C for 16 hours and at various concentrations of IPTG. The soluble un-induced control is shown in Lane 2, and has an over-expression of protein, possibly due to the long incubation time. The varying IPTG concentrations are 0.1 mM (Lane 3 and 4), 0.2 mM (Lanes 5 and 6) and 0.4 mM (Lanes 7 and 8). The first of the lanes show the insoluble fraction and the second the soluble fraction. The over-expressed protein is approximately 70 kDa according to the molecular weight marker (Lane 1).

4.4.1. Purification of *L. mexicana* transketolase

The method derived in the previous section to produce soluble *L. mexicana* TKT protein was used to prepare protein for purification. The *L. mexicana* TKT protein contained an N-terminal 6 x histidine tag and the same protocol for purification was used as for the *E. coli* TKA protein. 500 ml was initially induced for 14-16 hours at 15 °C with 0.1 mM IPTG and sonicated in 10 ml of buffer. The soluble and insoluble fractions were separated and 10 ml of soluble fraction was loaded onto the BioCAD nickel column. The elution profile is seen in Figure 4.8. The flow through is seen to contain little of the histidine-tagged protein and nothing is visible on an SDS-PAGE gel during the initial 20 mM imidazole wash. The pure *L. mexicana* TKT protein was eluted at imidazole concentrations between 400 mM and 500 mM (Figure 4.8b and 4.8c). The fractions 7-13, containing the eluted recombinant TKT, were pooled and dialysed overnight into 100 mM Tris buffer, pH 7.5.

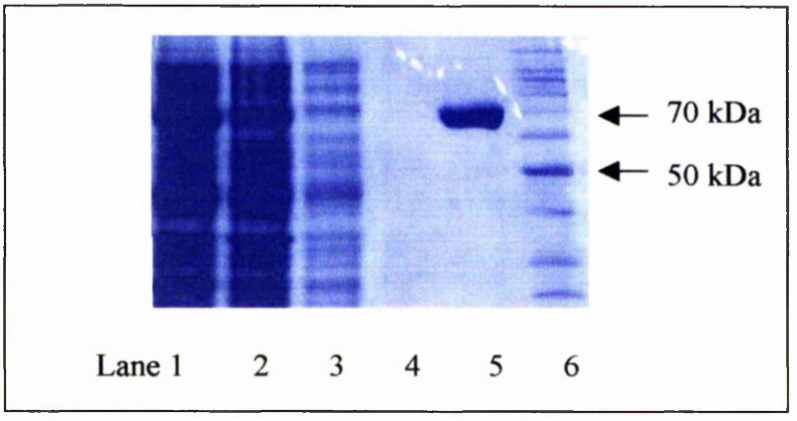


Figure 4.8a: Purification of the *L. mexicana* TKT protein from *E. coli* soluble lysate. The induced insoluble and soluble fractions are in Lanes 1 and 2, and 10 µl /10 ml total volume was run on the gel. The 10 ml soluble fraction was loaded onto the column. The flow through is seen in Lane 3 and 10 µl from a total 36 ml collected was run on the gel. The 20 mM imidazole wash is in Lane 4 and represents 10 µl / 15 ml collected. The pure protein was pooled from fractions 8-13 (Figure 4.8c) and 10 µl / 10 ml was run on the gel (Lane 5). The molecular weight marker (Lane 6) indicates that the purified protein is of approximately 70 kDa.

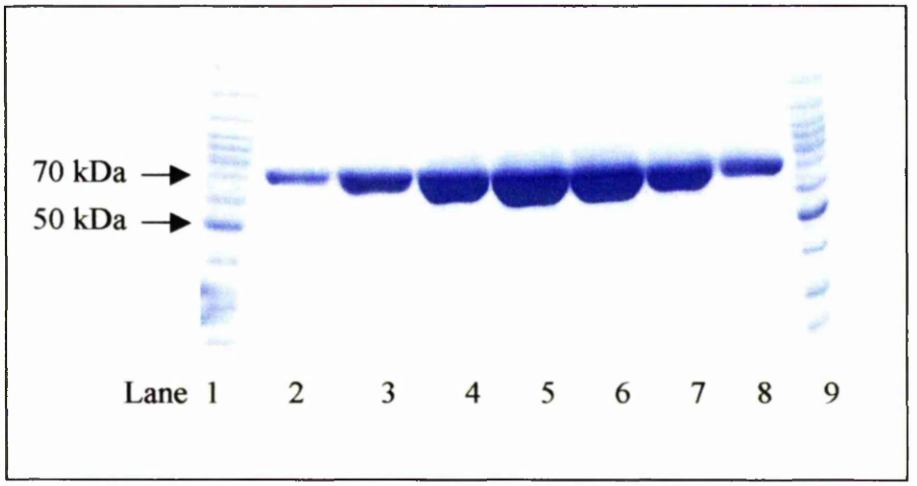


Figure 4.8b: The elution profile of the *L. mexicana* TKT protein over an increasing concentration of imidazole. Lanes 2 -8 represents fractions 7-13 on the trace of the purification in Figure 4.8c. Fraction 5 corresponds to the top of the peak of protein being eluted from the column at approximately 450 mM imidazole. Lanes 1 and 9 are the molecular weight markers indicating that the eluted protein is approximately 70 kDa.

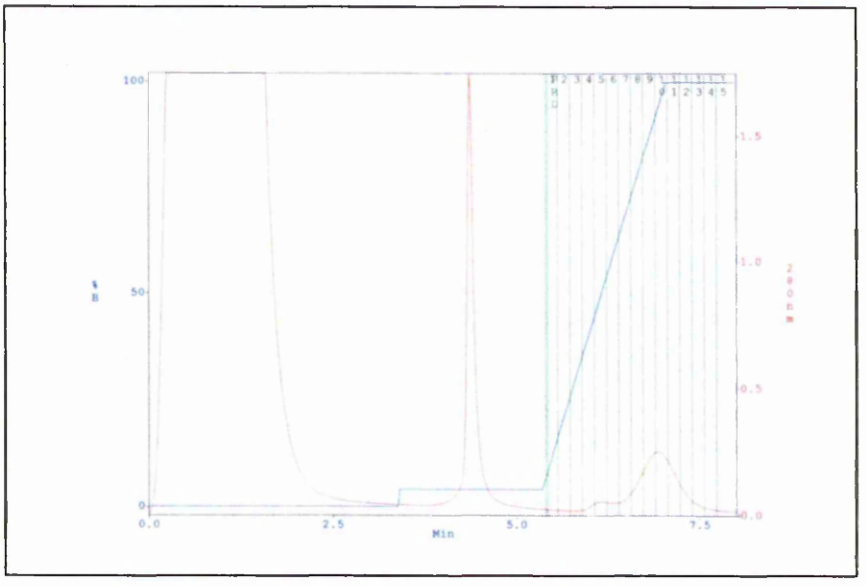


Figure 4.8c: The purification profile of the *L. mexicana* TKT protein. The blue line represents the imidazole concentration and the gradient used to elute the protein is from 20 mM to 500 mM (shown as percentage B, with 500 mM 100 %). The absorbance at 280 nm is a measurement of the amount of protein present (red line/text). The fractions collected during the imidazole gradient and elution of the protein are represented by green lines and text.

The eluted protein was tested for activity using R5P and X5P as substrates (see Section 2.3.1.1.), and had a specific activity of 1.65 U/mg. 5-8 % of the total protein loaded onto the column was the purified recombinant *L. mexicana* TKT, with 10 to 15 mg of pure TKT being obtained from one litre of *E. coli* culture.

4.5. Preparation of xylulose 5-phosphate

Xylulose 5-phosphate was prepared using a biotransformation method described by Zimmermann (Zimmermann *et al.*, 1999) (see Section 2.3.2.). 35 mg of *E. coli* TKA was used to catalyse the reaction. The pH increases during the reaction due to the utilisation of hydroxypyruvate and must be compensated for by the addition of HCl to retain the pH at 7.5. This can be used to follow the course of the reaction (Figure 4.9). During the reaction, fructose 1,6-bisphosphate was measured to determine if the starting material was being utilised (Figure 4.9). The assay system used also measured dihydroxyacetone phosphate (DHAP), an intermediate product made during this reaction. Fructose 1,6-bisphosphate was not present at the end of the reaction, however DHAP was present in the preparation. A noticeable colour change to yellow was also documented after 48 hours.

The X5P was purified using an anion exchange column (detailed in Section 2.3.2.2.) whereby the X5P was eluted using 1.3 M formic acid over 80 fractions of 15 ml each. The fractions were assayed for X5P using the transketolase assay, and the fractions yielding transketolase activity were pooled (Figure 4.10) and freeze-dried. The resulting sample was a yellow crystalline substance that could not be fully dehydrated and was therefore resuspended in 100 mM Tris buffer and freeze-dried for a second time. The sample could not be quantified by weight, however was used in assays in a quantity that was not inhibiting to the assay. The final sample contained no fructose 1,6-bisphosphate, however did contain DHAP. The preparation was subjected to NMR analysis, which confirmed that the X5P had been made (Figure 4.11) (according to the NMR trace presented by Zimmermann *et al.*, 1999). The final preparation is not pure, as other traces can be seen on the NMR, and DHAP activity can be detected.

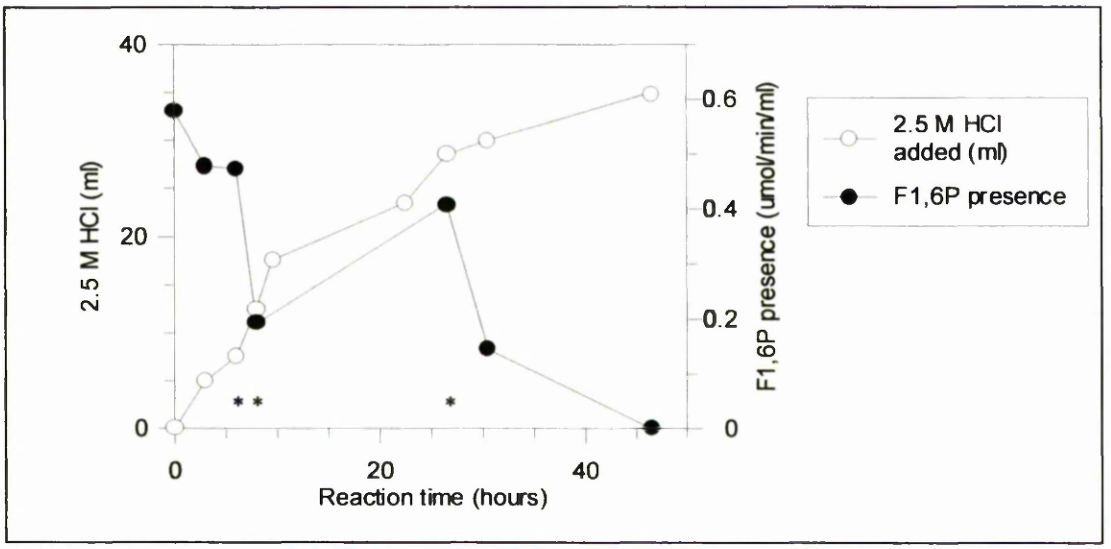


Figure 4.9: The course of the biotransformation as followed by addition of HCl to stabilise the pH at 7.5 and the presence of fructose 1,6-bisphosphate (F1,6P). The reaction proceeded over 48 hours and 2.5 M HCl was added to retain the pH at 7.5, an optimal pH for the reaction. The * indicates when components were added (detailed in Section 2.3.2.).

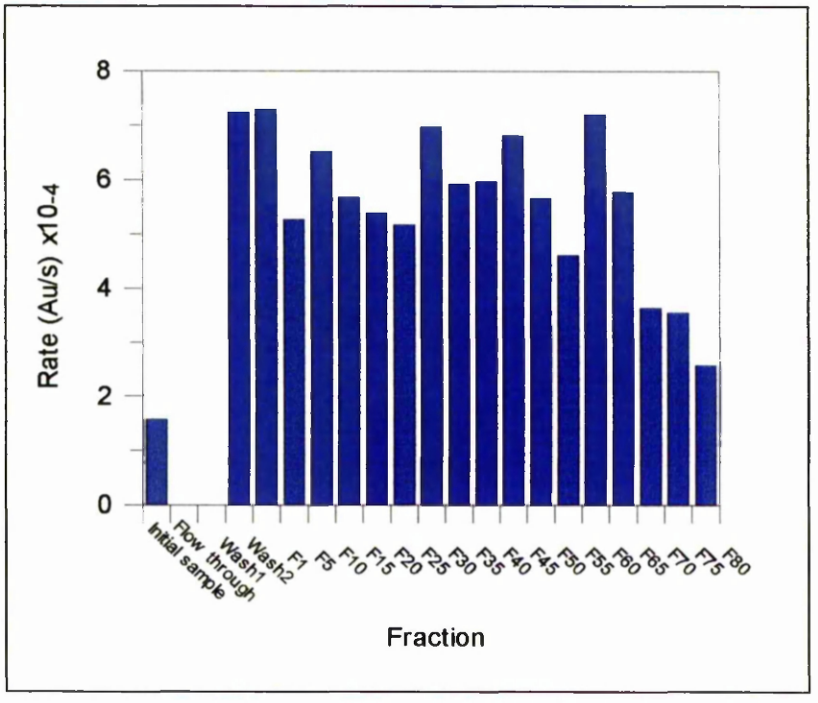
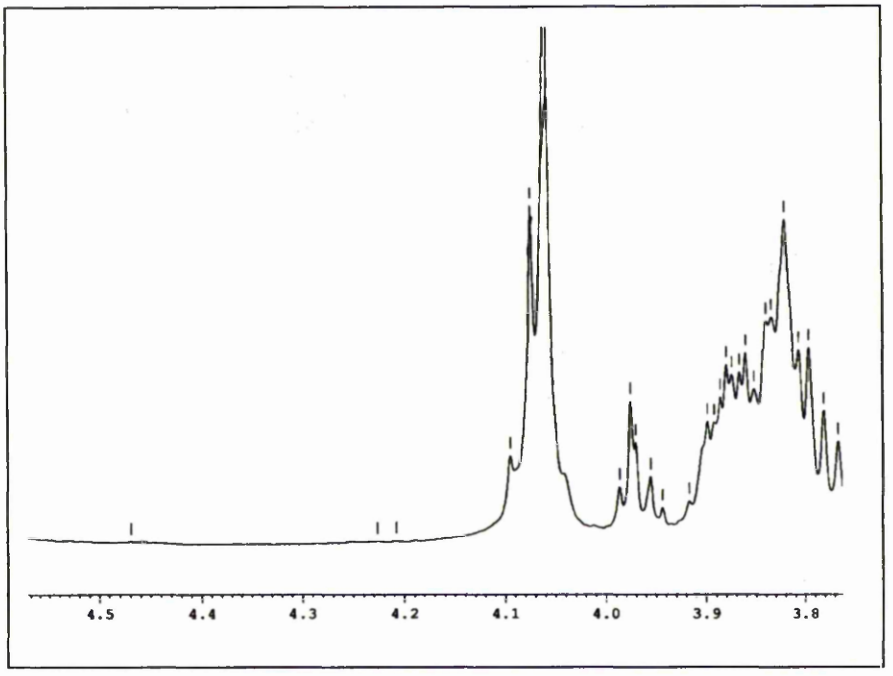


Figure 4.10: The detection of X5P in various fractions of the purification of the biotransformation. The presence of X5P was measured by a transketolase assay using R5P as the acceptor substrate. The rate was measured in absorbance units/second (AU/s) x 10^{-4}. The initial sample was diluted 1/10 and measured for the presence of X5P presence. The flow through and first wash with 1M formic acid, contained no detectable trace of X5P. The second wash, and all subsequent fractions collected, did contain X5P.

A)



B)

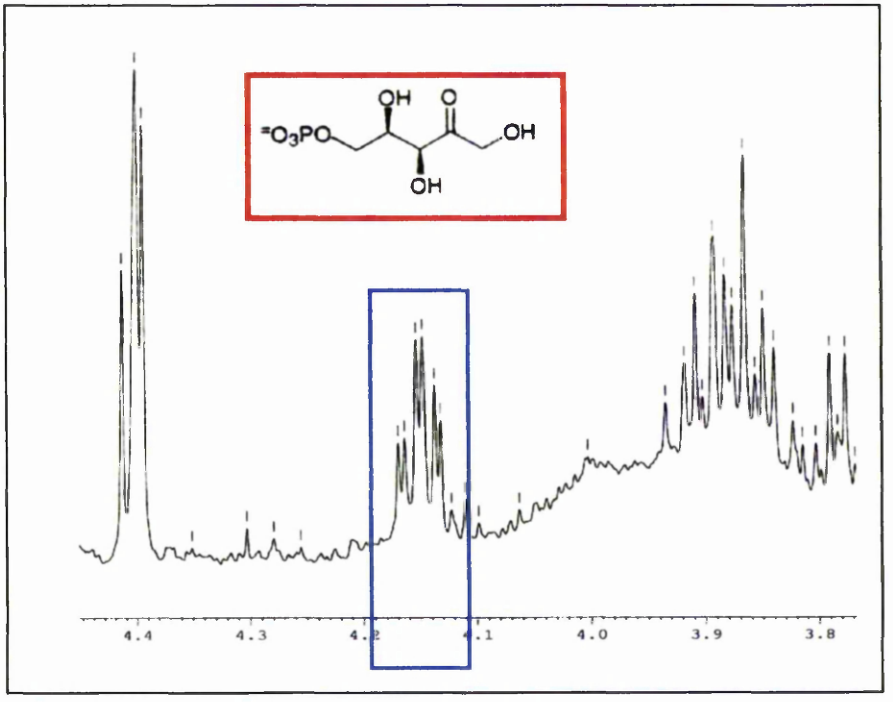


Figure 4.11: ^{1}H NMR of the substrate and product from the xylulose 5-phosphate biotransformation. Figure A) represents the NMR trace of fructose 1, 6-bisphosphate and Figure B) represents the NMR trace of the final product of the biotransformation. The blue box indicates the triplet of doublets characteristic of X5P (molecule in red box), according to Zimmermann *et al.*, 1999. The scale below the NMR trace represents the chemical shift in parts per million (ppm)

This X5P preparation was used to further characterise the *L. mexicana* TKT subsequent to Sigma halting production of commercial X5P. The experiments that used the X5P made in-house are indicated in the text with an asterisk. The preparation did contain DHAP, which is a product produced during the transketolase assay (see Section 2.3.1.1.) and therefore initial DHAP was converted to glycerol 3-phosphate using glycerol-3-phosphate dehydrogenase before the transketolase was assayed. Repetition of this biotranformation was problematic, with the reaction not going to completion as there was a significant amount of substrate left and there was a difficulty in re-using the column to purify the reaction.

4.6. Characterisation of the *L. mexicana* recombinant transketolase protein

4.6.1. Characteristics of the *L. mexicana* TKT protein

A lag phase has been previously documented for transketolases from human erythrocytes, *E. coli* and *S. cerevisiae*, and is thought to be due to the slow binding of the cofactors to the enzyme (Booth and Nixon, 1993; Sprenger *et al.*, 1995) and the slow dimerisation of the subunits (Egan and Sable, 1981). *L. mexicana* TKT exhibits a lag phase of 50-200 seconds prior to reaching steady-state velocity (Figure 4.12).

The *L. mexicana* recombinant TKT enzyme was dialysed into buffer subsequent to the purification procedure to remove imidazole and sodium phosphate from the preparation. The addition of the transketolase cofactors TPP and $MgCl_2$ to the dialysis buffer did not increase the activity of the *L. mexicana* TKT protein, or alter the lag phase observed during the assay (Figure 4.12 and Table 4.2). This was also observed during the reconstitution of the *E. coli* TKA holoenzyme, where a pronounced lag phase occurred due to slow enzyme reconstitution (Sprenger *et al.*, 1995). Incubation of various concentrations of the cofactor TPP with *L. mexicana* TKT resulted in a slight increase of activity after a one-hour incubation at room temperature with 10 mM TPP

(Table 4.3). Incubation with 100 mM TPP caused complete inhibition of activity of the *L. mexicana* TKT enzyme. This is possibly due to excess TPP blocking the active site of the transketolase. The lag phase of the reaction did not alter after incubation of protein with various TPP concentrations. Yeast transketolase displays a decreased lag phase if preincubated with TPP, however this was dependent on the time of incubation and the concentration of the protein (Egan and Stable, 1981). This observation led to a model suggesting that a slow dimerisation of inactive monomer-coenzyme complex occurs and if the TPP concentration is high, most of the enzyme would remain in the monomer form.

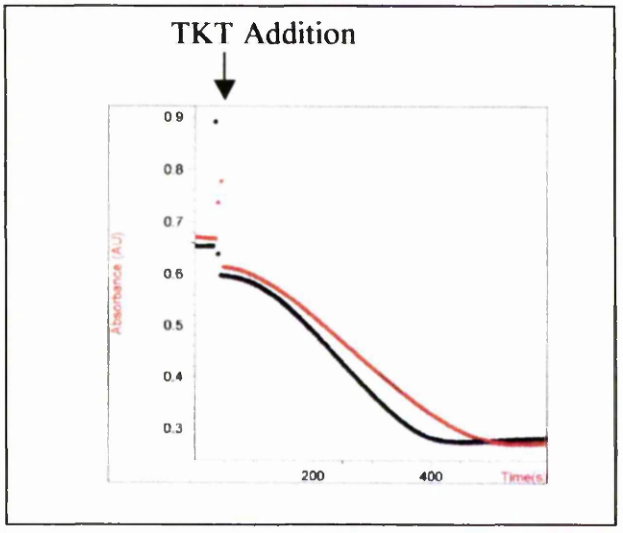


Figure 4.12: A spectrophotometric trace of the *L. mexicana* TKT reaction. This assay utilises R5P and X5P* as substrates. The absorbance (AU) at 340 nm is recorded over 600 seconds (Time(s)). The point at which transketolase is added to the reaction is indicated by an arrow. Two reactions are followed, the black trace follows TKT after dialysis into Tris, and the red trace follows TKT after dialysis into Tris, TPP and $MgCl_2$ (detailed in text). A lag phase from 50 –150 seconds can be seen.

Dialysis buffer	Activity (U/ml)
100 mM Tris, pH 7.5	9.6×10^{-2}
100 mM Tris, 2 mM TPP, 5 mM $MgCl_2$, pH7.5	9.1×10^{-2}

Table 4.2: The activity of *L. mexicana* TKT after incubation in different dialysis buffers. The assay used R5P and X5P* as substrates. The activity was recorded in Units/ml (μmols/min/ml).

TPP concentration (mM)	Activity (U/ml)
100	3.2×10^{-7}
10	5.2×10^{-1}
1	4.2×10^{-1}
0.1	4.2×10^{-1}
0	4.2×10^{-1}

Table 4.3: Resulting activity of *L. mexicana* TKT protein after incubation with the cofactor TPP at various concentrations. Thiamine pyrophosphate (TPP) at various concentrations was incubated for 1 hour with the *L. mexicana* TKT protein and assayed using R5P and X5P as substrates. The activity was recorded in Units/ml (U/ml). $MgCl_2$ was used at a saturating concentration (5mM) and was not varied during this experiment.

The *L. mexicana* TKT protein was assayed for activity in Tris buffer with varying pH. *L. mexicana* TKT has a pH optimum of 7.5 in Tris (100 mM) (Table 4.4). This is similar to other transketolases as *S. cerevisiae* and *E. coli* have an optimum pH of 7.6 and pH 8.0-8.5 respectively (Datta and Racker, 1961; Sprenger *et al.*, 1995) (Table 4.5).

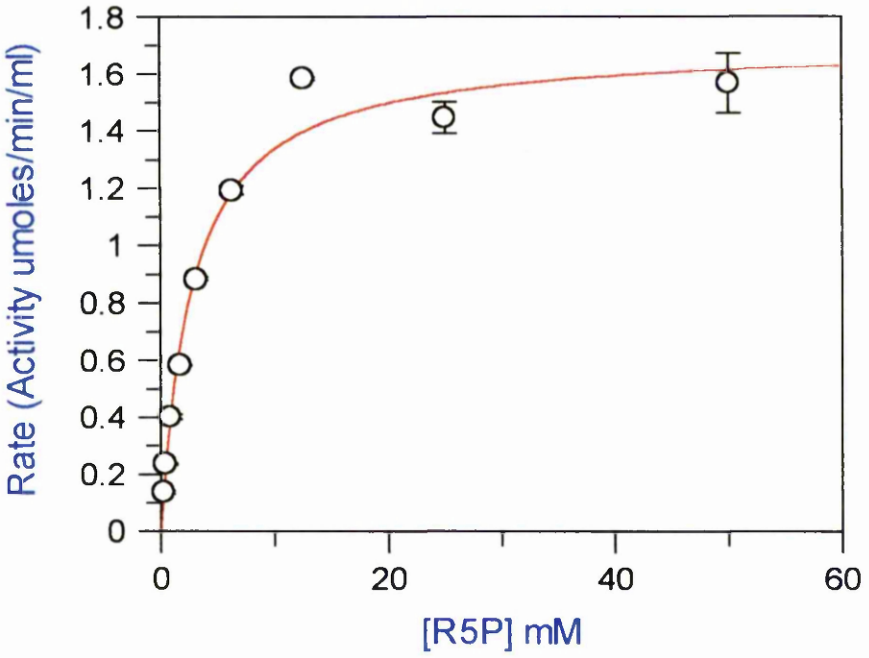
pH	Activity (U/ml)
8.5	3.2×10^{-1}
8	3.8×10^{-1}
7.5	4.5×10^{-1}
7	4.1×10^{-1}

Table 4.4: The effect of pH on the *L. mexicana* TKT protein activity using 100 mM Tris buffer. R5P and X5P* were used as substrates to assay this enzyme in this experiment.

The highest specific activity of the *L. mexicana* TKT is 1.65 U/mg. The K_m and the V_{max} for ribose 5-phosphate are 2.75 mM and 1.7 μmol/min/mg protein, respectively (Figure 4.13). The xylulose 5-phosphate was used at saturating concentrations during the experiment, as were all other components of the assay. No kinetic parameters could be determined for xylulose 5-phosphate, as the preparation could not be quantified. The K_m value for ribose 5-phosphate is similar to the recombinant *E. coli* TKA protein. A table of K_m and specific activities is shown in Table 4.5. The K_m for the *E. coli* transketolase is the most similar to the *L. mexicana* TKT, with yeast, spinach and human transketolases all having a higher affinity of binding.

Substrate inhibition was seen during the kinetic analysis, with >25 mM ribose 5-phosphate inhibiting the reaction (Figure 4.14). This effect has been seen previously in transketolase from *S. cerevisiae* (bakers yeast) (Solov'eva *et al.*, 2000; Kovina *et al.*, 2000). Ribose 5-phosphate inhibits yeast transketolase at a concentration of 3 mM (K_m of ribose 5-phosphate is 0.21 mM). It has been suggested that there is a common binding site in transketolase for both substrates and therefore inhibition will occur if there is an excess of product or substrate (Solov'eva *et al.*, 2000).

A)



B)

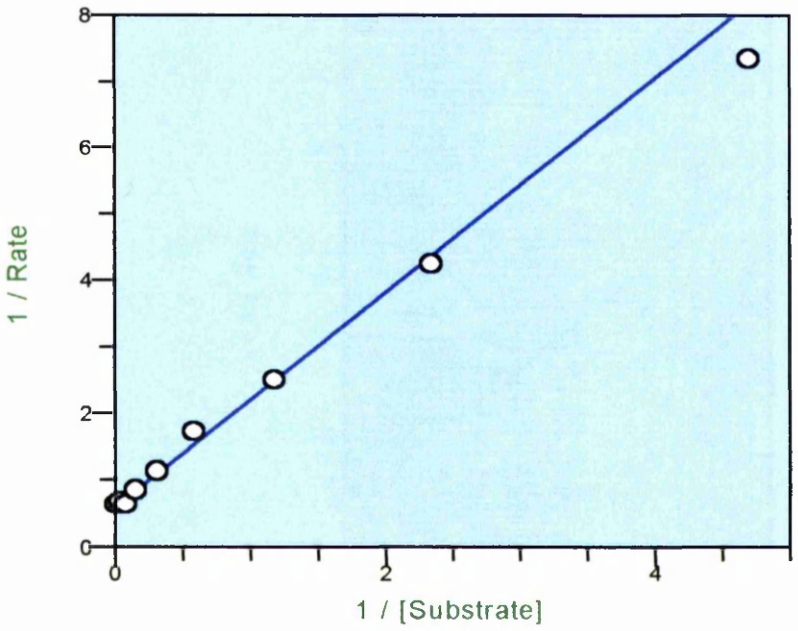


Figure 4.13: Affinity of *L. mexicana* TKT for the substrate ribose 5-phosphate. R5P and X5P* were used as substrates to assay this enzyme in this experiment. Graph A) shows the effect of R5P on the specific activity of the *L. mexicana* TKT and graph B) show the Linweaver –Burk reciprocal plot of A). The standard error of the mean was calculated with n=3.

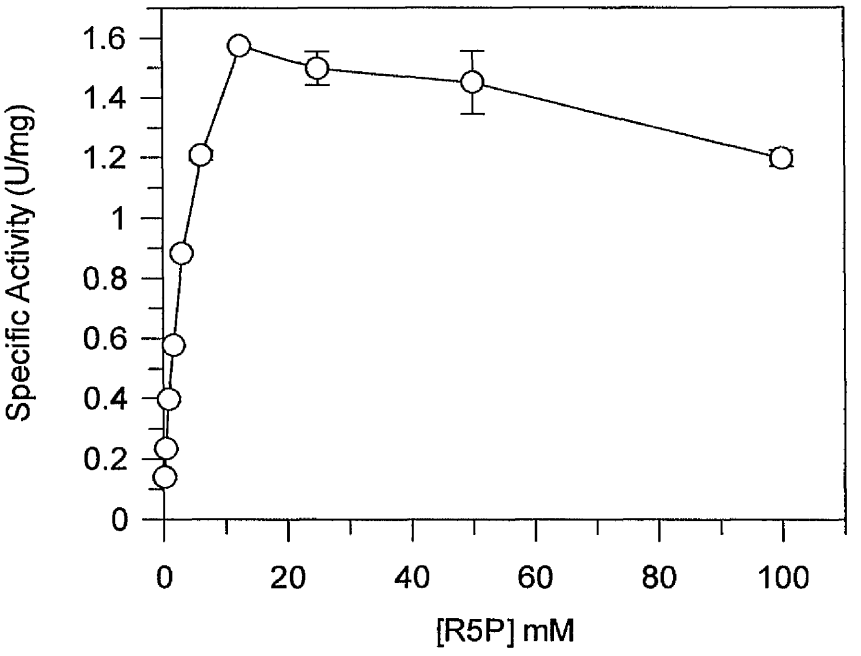


Figure 4.14. The decrease in activity of transketolase with increasing concentrations of acceptor substrate. The data points are the mean of 3 experiments, with the standard error of the mean shown.

Source of transketolase	Specific activity (U/mg)	K_m for R5P (mM)	Optimum pH	Reference
S. cerevisiae	20	0.4	7.6	Data and Racker, 1961; Kochetov, 1982
E. coli (recombinant)	50	1.4	8-8.5	Sprenger *et al.*, 1995
Spinach chloroplasts	12	0.330	8	Teige *et al.*, 1998
Rat liver	2.42	0.066	8	Paolettii, 1983
Human (recombinant)	13.5	0.51	n.d.	Schenk *et al.*, 1998(a)

Table 4.5: Comparison of characteristics of transketolases from various organisms. n.d. means not done.

4.6.2. Storage conditions

To determine the optimal storage conditions for the *L. mexicana* TKT protein, a number of conditions were examined. Subsequent to dialysis with one litre of Tris (100 mM, pH 7.5), the *L. mexicana* TKT was stored at 0.5 mg/ml at 4 °C, -20 °C and –80 °C, with and without 50 % glycerol. The enzyme activity was measured over 2 months and is summarised in Table 4.6. The enzyme had an initial activity of 7.5×10^{-2} U/ml with a concentration of 0.5 mg/ml. 0.1 mg of the purified protein was measured in each 1 ml assay. The enzyme was most efficiently stored at –80 °C. The enzyme was not active subsequent to storage at –80 °C if the preparation was not dialysed.

Storage conditions	**Activity (U/ml)**		
	Day 7	Day 14	Day 28
4 °C	1×10^{-2}	7.7×10^{-3}	0
4 °C + 50 % glycerol	1.8×10^{-2}	5.7×10^{-3}	0
-20 °C	0	0	0
-20 °C + 50 % glycerol	5.6×10^{-2}	2.4×10^{-2}	2.4×10^{-2}
- 80 °C	5.7×10^{-2}	7.4×10^{-2}	7.4×10^{-2}
-80 °C + 50 % glycerol	4.3×10^{-2}	4.2×10^{-2}	4.2×10^{-2}

Table 4.6: The effect of various storage conditions on the enzyme activity of the *L. mexicana* TKT protein. R5P and X5P* were used to measure the activity (U/ml) of the *L. mexicana* TKT purified protein.

A complete biochemical investigation and a kinetic profile of the *L. mexicana* TKT could not be performed due to the lack of X5P available. An alternative assay using hydroxypyruvate and erythrose 4-phosphate as substrates was used. However this assay gave an extremely low specific activity of 0.025 U/mg, which precluded the determination of kinetic parameters.

4.7. *L. mexicana* transketolase protein sequence analysis

Having sequenced the *TKT* gene from *L. mexicana*, the predicted amino acid sequence was obtained and analysed. TKT from *L. mexicana* contains 671 amino acids and each individual monomer has a predicted molecular weight of 71.8 kDa (ProtParam tool, ExPASy). To compare this transketolase with other known transketolase protein sequences, an alignment was made comparing all known full-length transketolase amino acid sequences using the ClustalW algorithm (Thompson *et al.*, 1994) (Vector NTI). A direct comparison of the *L. mexicana* and *S. cerevisiae* transketolase amino acid sequences was performed to highlight the conserved residues and overall identity.

4.7.1. The amino acid alignment of transketolase proteins

19 full-length and two partial transketolase amino acid sequences were aligned and analysed by Schenk *et al.*, 1997. However subsequent to that analysis, other transketolases have been sequenced and functionally characterised, and numerous putative transketolases have been predicted from database similarity searching of sequenced genomes. The *L. mexicana* TKT amino acid sequence was aligned with 41 transketolase sequences from various organisms using the Clustal W algorithm (Vector NTI) (Figure 4.15). The full names of the organisms and accession numbers used in the alignment are listed in Table 4.7. The alignment contains 42 full-length transketolase amino acid sequences, 22 are derived from genome sequencing data and are therefore putative proteins (indicated in Table 4.7). The *Plasmodium falciparum* transketolase protein was obtained from The Plasmodium Genome Resource (PlasmoDB) (accession number PF00456) and *Plasmodium yoelli* was obtained from DNA sequence (chrPhyl_c1155, PlasmoDB) and subsequently translated into a putative amino acid sequence (Vector NTI).

The transketolase protein sequences from *Nicotiana tabacum* (Tob), *Solanum tuberosum* (Sol) and *Spinacia oleracea* (Spinach) all exhibit an N-terminal extension of 54 to 57 amino acids in length. *Solanum tuberosum* and *Spinacia oleracea* both have transketolase sequences localised to the chloroplast, therefore this N-terminal region may be a chloroplast targeting motif. The *Nicotiana tabacum* transketolase sequence has not been confirmed to be located in the chloroplast, however the amino acid sequence was analysed using a protein localisation prediction programme (PSORT), and was recognised as having N-terminal chloroplast targeting signals. The TPP binding motif is highly conserved in transketolase amino acid sequences (Hawkins *et al.*, 1989) and is present in the alignment from 240 to 271 amino acids in bold type. It comprises 32 amino acids with the consensus sequence (G/S)(D/E)GX_3EGX_3EX_4AX_3LX_2LX_4 DXN, where X denotes any amino acid variation. The variation of the aspartic acid to a glutamic acid residue at position two of the TPP motif occurs in *Treponema pallidum* (Trep). This aspartic acid residue has been shown to be essential for catalytic activity and dimer formation in human transketolase (Wang *et al.*, 1997) and is highly conserved in all other species except *Treponema pallidum*. This *Treponema pallidum* sequence is a putative transketolase as it was derived from the annotation of the complete genome sequence (Fraser *et al.*, 1998) and therefore may not be an active enzyme, although the retention of an acidic residue may be sufficient to retain its activity. The TKT motif present in the middle region of the protein sequence is shown in the alignment at positions 567 to 602 in bold type. This well conserved stretch of 36 amino acids has the consensus sequence (T/S/K)(H/K)(D/C*) (S/G*)X_3G(E/G) DGP (T/S*)(H/Q*)X(P/A)X(E/D)X_5(R/P)X_2(P/E)X_5(R/Y*)PXD where X denotes any amino acid residue and * represents the residues that vary only in mammalian sequences. The second residue of the TKT motif is a highly conserved histidine residue known to

interact with the cofactor TPP and is involved in substrate binding (Nilsson *et al.*, 1997). At this position, *Sinorhizobium meliloti* (sin) has a lysine residue, differing from the conserved histidine. This may suggest that the protein interacts differently with the cofactor and substrates, in comparison to other transketolases. The yeast transketolase amino acid sequences (and *Sinorhizobium meliloti*) all have small C-terminal extensions of 12 – 30 amino acids in length. The *L. mexicana* TKT also has a C-terminal extension of 13 amino acids, the final three being a peroxisome-targeting signal SKM. No other transketolase so far has been seen to have a peroxisome-targeting sequence. A *L. major* transketolase partially sequenced during this study did contain the C-terminal motif SKM (see chapter V). 24 residues over the 42-transketolase sequences are completely conserved.

PHYLA	NAME OF ORGANISM	SHORTENED NAME	LENGTH OF SEQUENCE (AA)	GENBANK ACCESSION NUMBER
Mammalia	*Homo sapiens*, tkt	Human	623	P29401
	Mus musculus	Mus	623	P40142
	Rattus norvegicus, tkt	Rat	623	p50137
Plant	*Nicotiana tabacum*, tkt	Tob	743	A52295
	Solanum tuberosum, tkt chloroplast precursor	Sol	741	Q43848
	Spinacia oleracea, tkt chloroplast precursor	Spinach	741	T09015
	Craterostigma plantagineum, tkt7	Crat7	676	Q42677
	Craterostigma plantagineum, tkt10	Crat10	679	Q42675
Yeast	*Saccharomyces cerevisiae*, tkt1	Sacc1	680	P23254
	Saccharomyces cerevisiae, tkt2	Sacc2	681	P33315 (P)
	Schizosaccharomyces pombe, tkt	Sch	685	T40162
	Pichia stipitis, tkt	Pic	695	P34736
	Candida albicans, *tkt*	Cand	677	T18231 (P)
	Kluyveromyces lactis, tkt	Klu	679	Q12630 (P)
Protozoa	*Leishmania mexicana*, tkt	Lmex	671	AJ427448
	Plasmodium falciparum, tkt	P.falc	669	PF00456* (P)
	Plasmodium yoelli, tkt	P.yoel	670	chrPyl_c1155* (P)
Bacteria	*Escherichia coli*, tkt1	Ecoli1	663	P27302
	Escherichia coli, tkt2	Ecoli2	667	P33570
	Candidatus Carsonella ruddii, tkt	C.rud	636	AAK17116 (P)
	Pasteurella multocida, tkt1	Past	668	P57927(P)
	Buchnera aphidicola, tkt	Buch	665	P57195 (P)
	Xylella fastidiosa, tkt	Xyl	666	AAF84738 (P)
	Haemophilus influenza, tkt	Haem	665	P43757 (P)
	Caulobacter crescentus, tkt1	Caulol	653	AAK25582 (P)
	Mesorhizobium loti, tkt	Mesor	683	BAB50573 (P)
	Sinorhizobium meliloti, tkt	sin	695	P56900 (P)
	Rhodobacter sphaeroides, tkt	Rhods	657	P29277
	Rhodobacter capsulatus, tkt	Rhod	672	Q52723
	Ralstonia eutropha, tkt plasmid	Ralp	670	P21726
	Ralstonia eutropha, tkt chromosomal	Ralc	670	P21725
	Bacillus subtilis, tkt	Bac	667	P45694 (P)
	Staphylococcus aureus, tkt	S.aureus	662	BAB42435 (P)
	Mycobacterium tuberculosis, tkt	Mycobt	700	O06811
	Mycobacterium leprae, tkt	Mycobl	699	P46708 (P)
	Mycoplasma pulmonis, tkt	Mycop	615	CAC13684 (P)
	Mycoplasma pneumoniae, tkt	Myco	648	P75611 (P)
	Mycoplasma genitalium, tk	Mycopg	648	P47312
	Ureaplasma urealyticum, tkt1	Ureal	653	NP_078425 (P)
	Treponema pallidum, tkt	Trep	661	O83571 (P)
	Streptococcus pneumoniae, probable tkt	Strep	656	P22976
	Corynebacterium glutamicum, tkt	Coryn	676	ATCC31833

Table 4.7: The transketolase amino acid sequences used in this study. The organisms are divided into Phyla of mammals, plants, yeast, protozoa and bacteria. The gram-negative bacteria are further divided into subdivisions by colour, with the γ-subdivision in blue, the α-subdivision in green and the β-subdivision in pink. The name of the organism is followed by the transketolase (tkt) enzyme details. The shortened name is used in the alignment figures. The length of the sequence in amino acids (AA) and the EMBL accession numbers are listed. The *Plasmodium* sequences were obtained from The Plasmodium Genome Resource (PlasmoDB) and the relevant accession numbers are specific to this database and are annotated with a *. The (P) after the accession number represents if the sequence is putative (derived from sequencing of genomes and sequence similarity and therefore not biologically characterised).

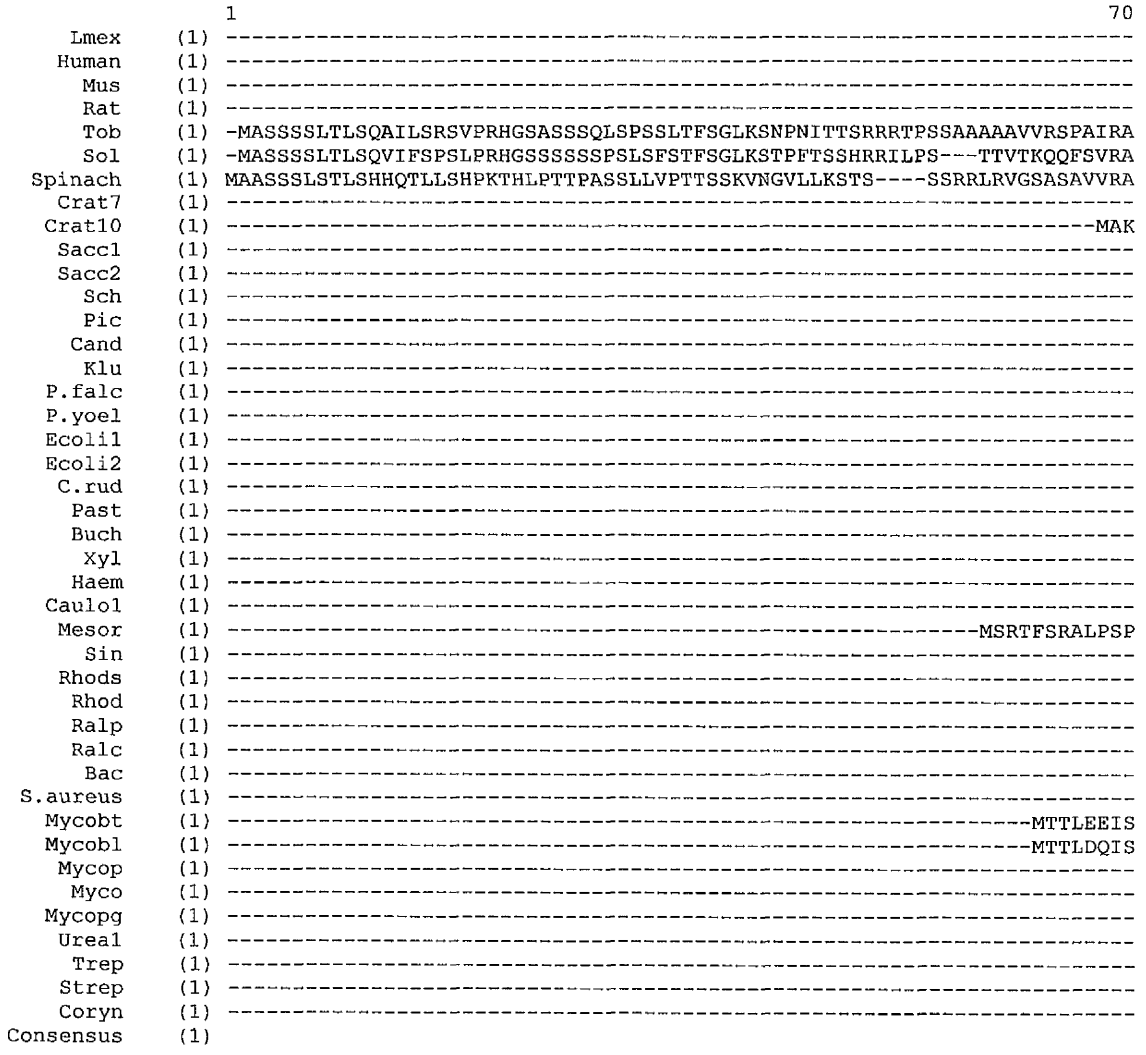


Figure 4.15: Alignment of transketolase amino acid sequences from various organisms (detailed in Table 4.6). The sequences were aligned using Clustal W algorithm (Vector NTI). The numbers in brackets refer to the amino acid number of the individual sequence, and the numbers at the top of the alignment refer to the overall alignment number. The amino acid residues that are identical in the alignment are highlighted in red type with a grey background. Blocks of amino acids with similar physiochemical properties are in black type with grey background. Weakly similar residues have green text and conservative residues have blue text. Non-similar residues are represented by black text. The lines above the alignment indicate the sections of sequence removed when preparing the phylogenetic trees in chapter V. The TPP and TKT motifs are indicated with underlined bold text in the consensus sequence.

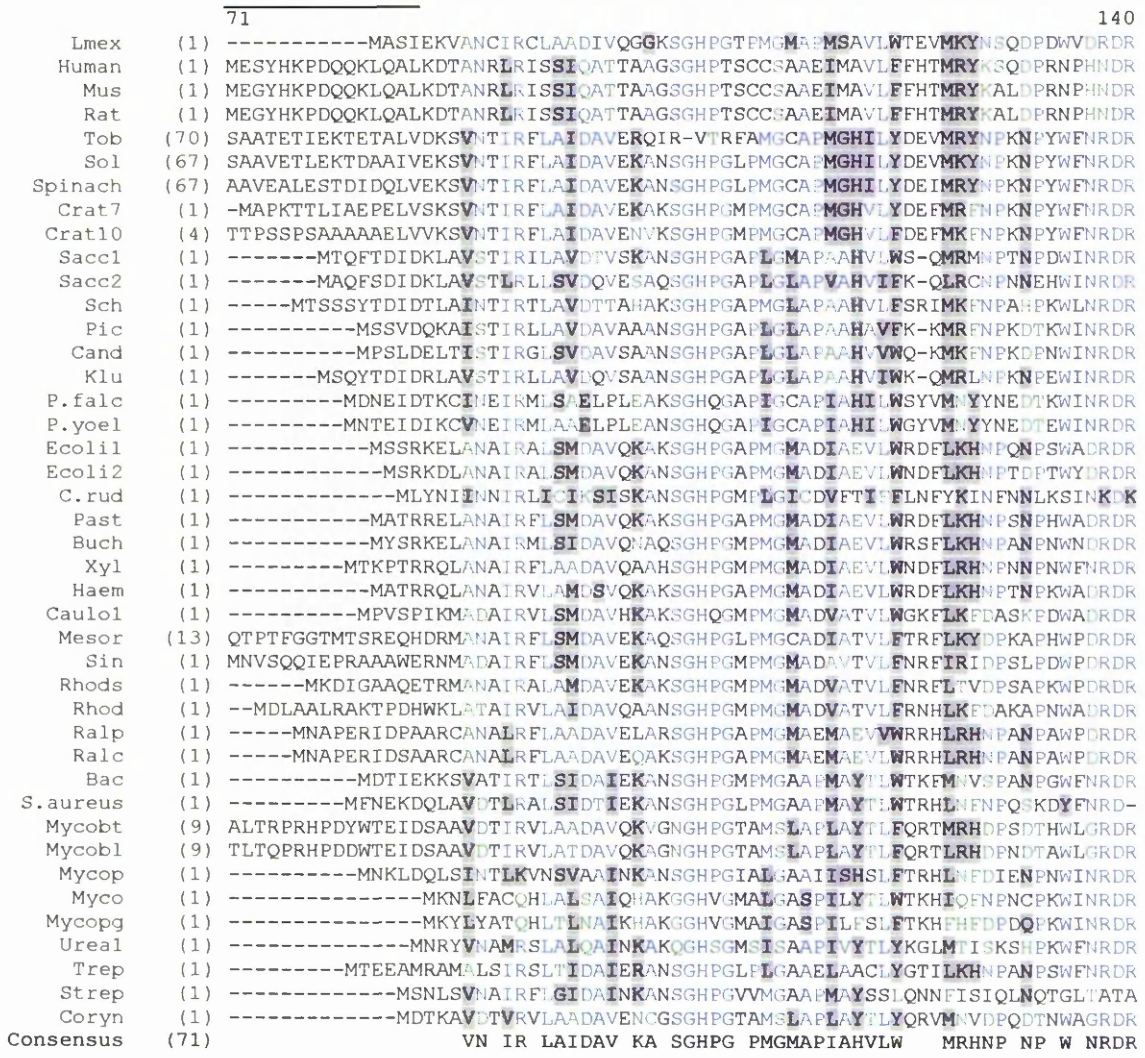


Figure 4.15 cont.: Alignment of transketolase amino acid sequences

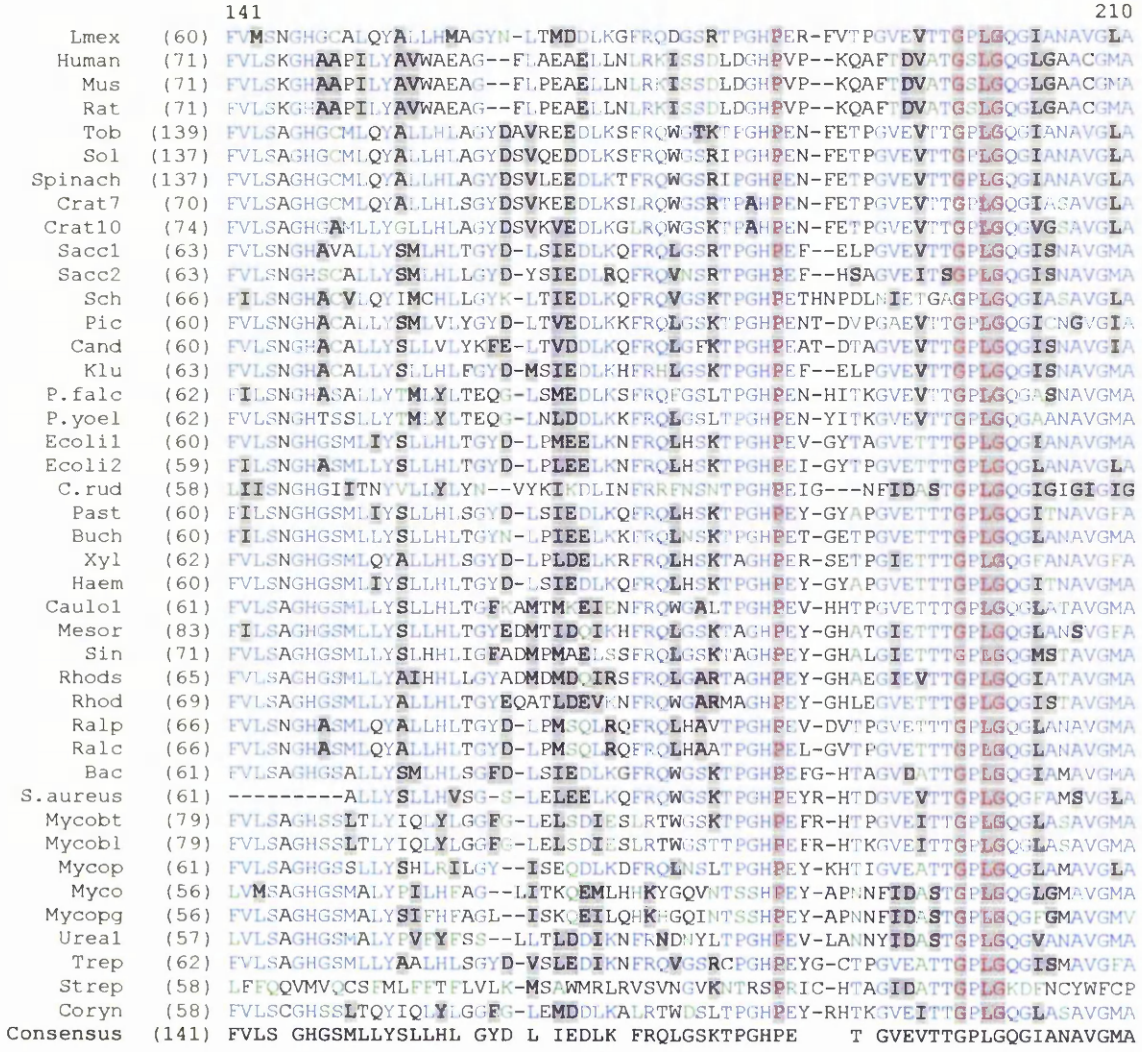


Figure 4.15 cont.: Alignment of transketolase amino acid sequences

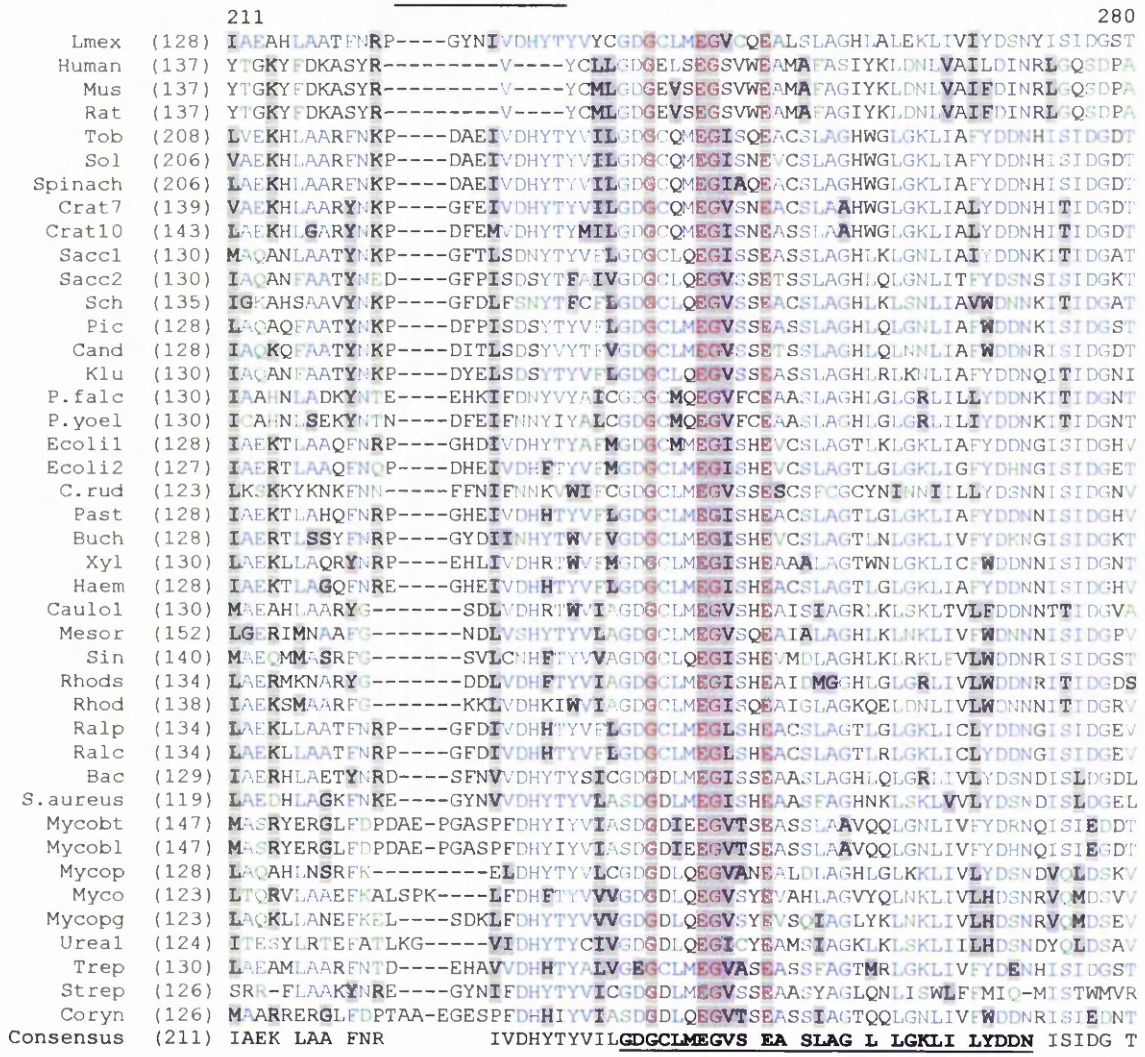


Figure 4.15 cont.: Alignment of transketolase amino acid sequences

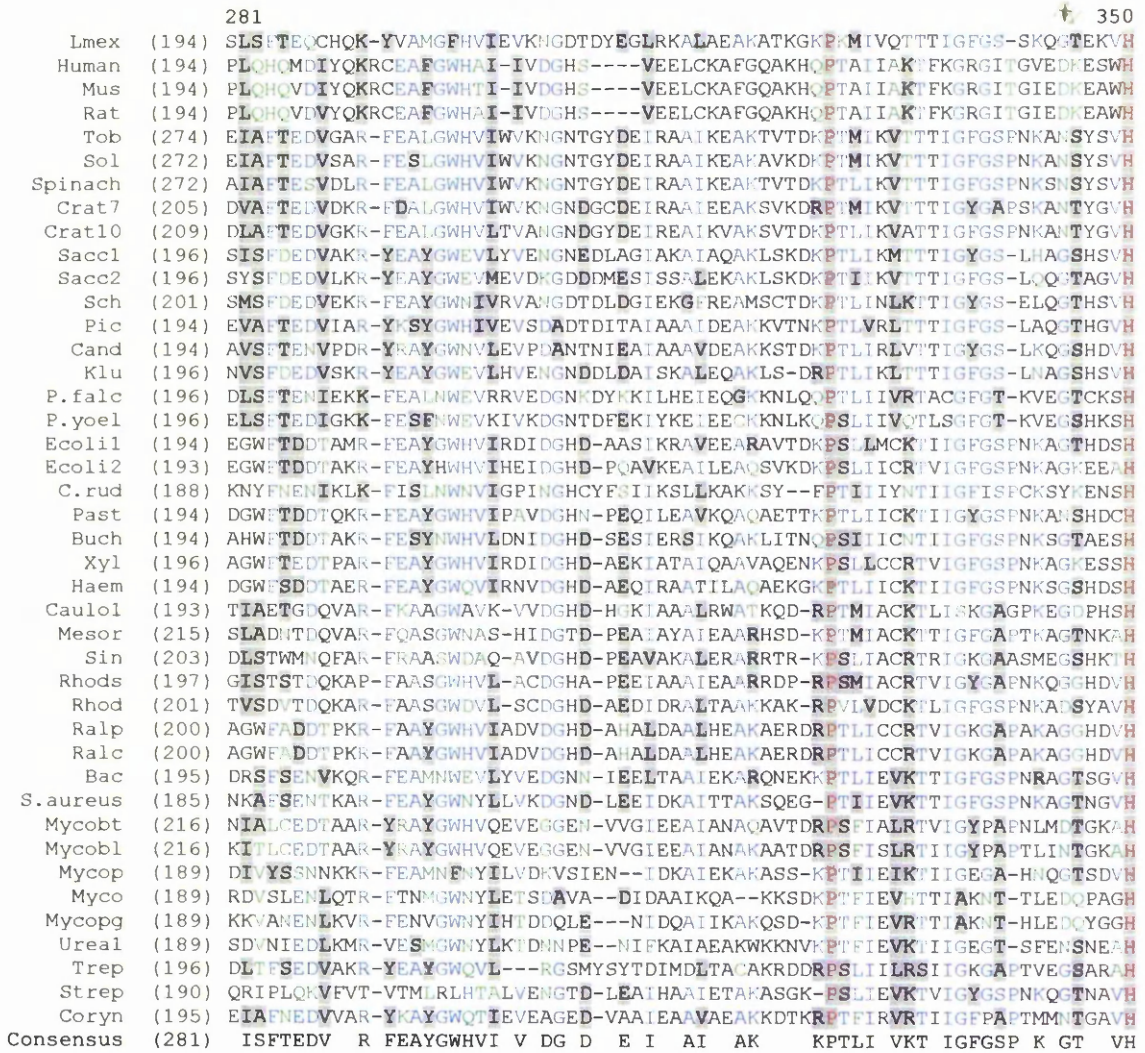


Figure 4.15 cont.: Alignment of transketolase amino acid sequences

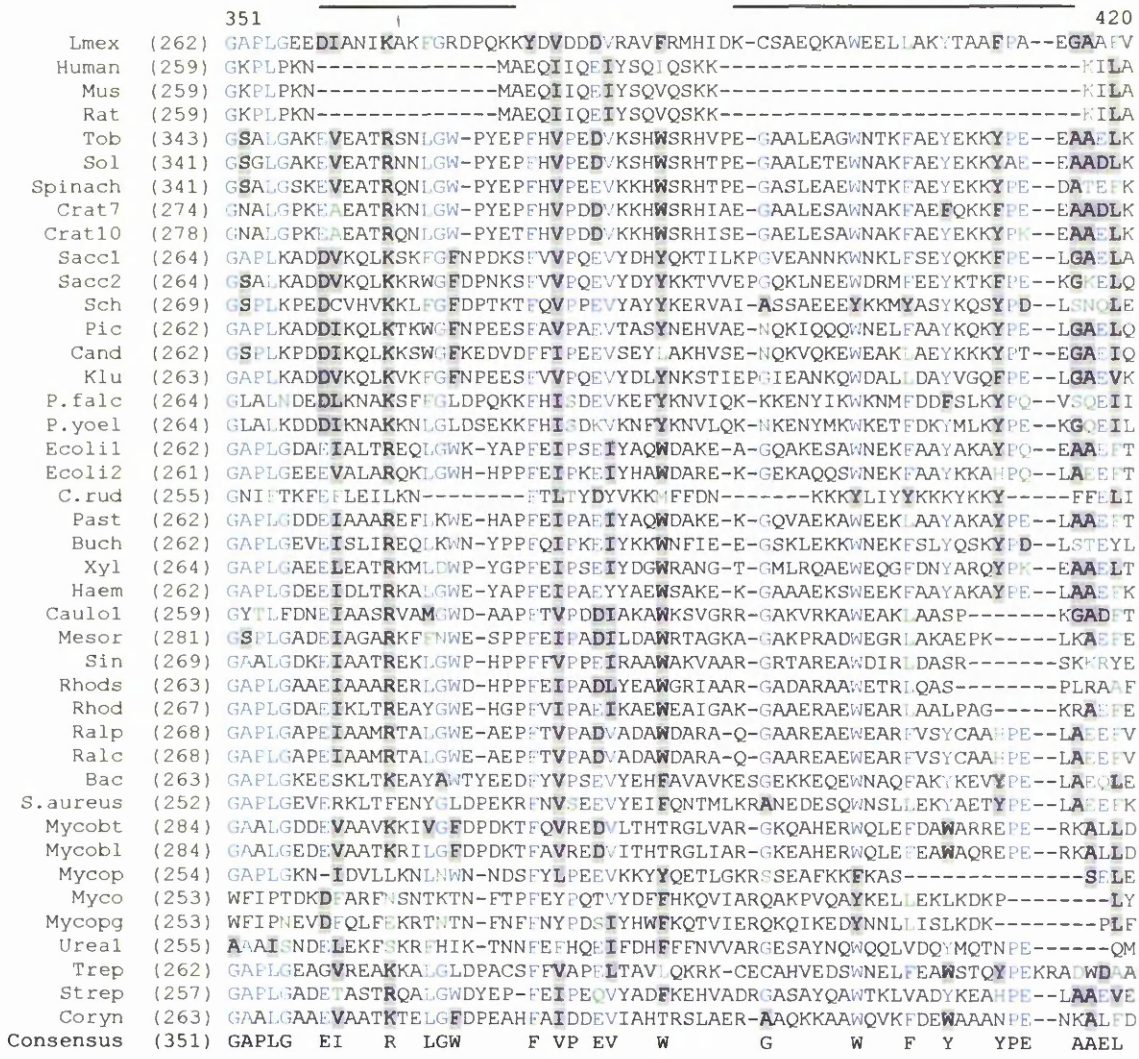


Figure 4.15 cont.: Alignment of transketolase amino acid sequences

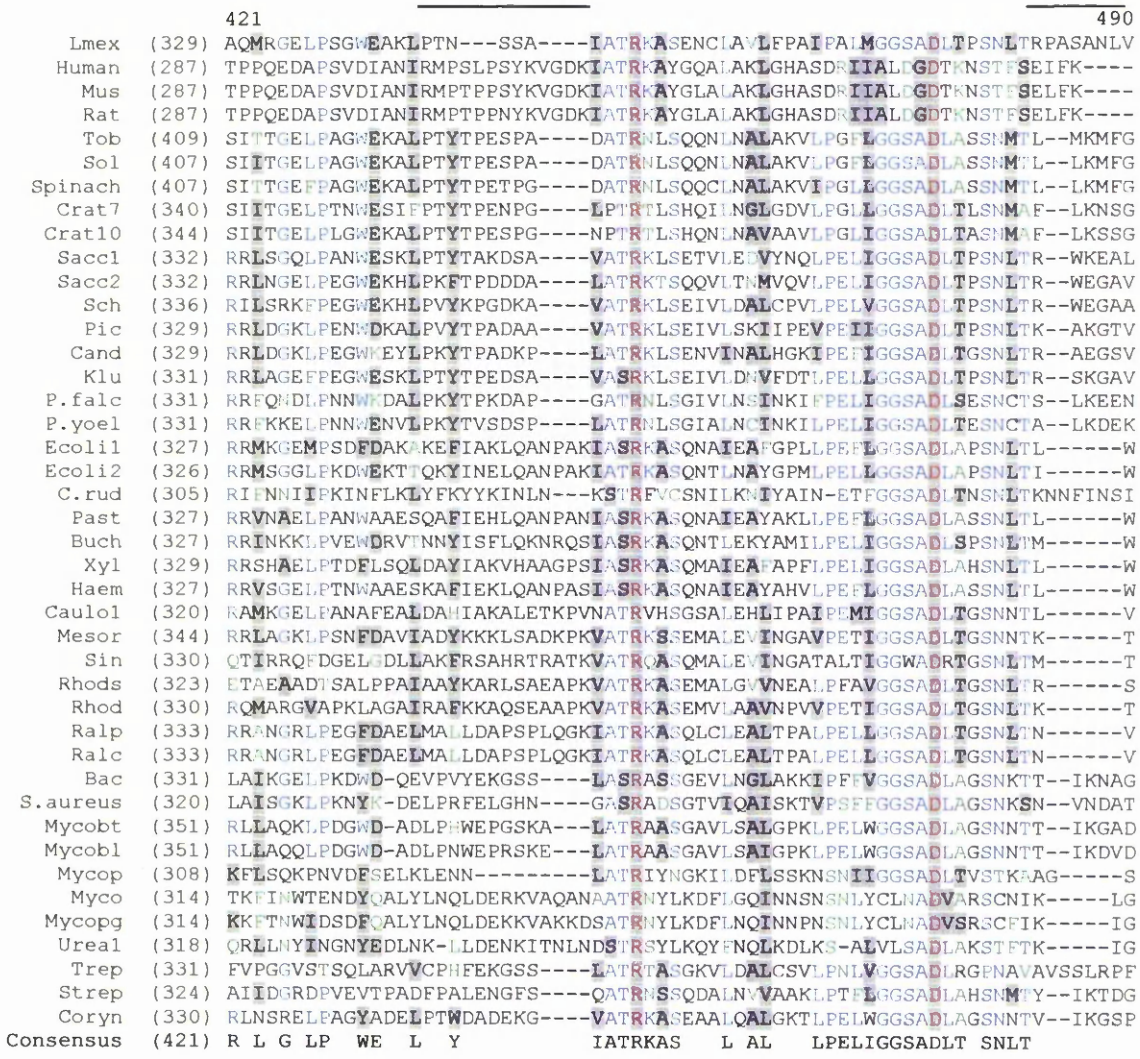


Figure 4.15 cont.: Alignment of transketolase amino acid sequences

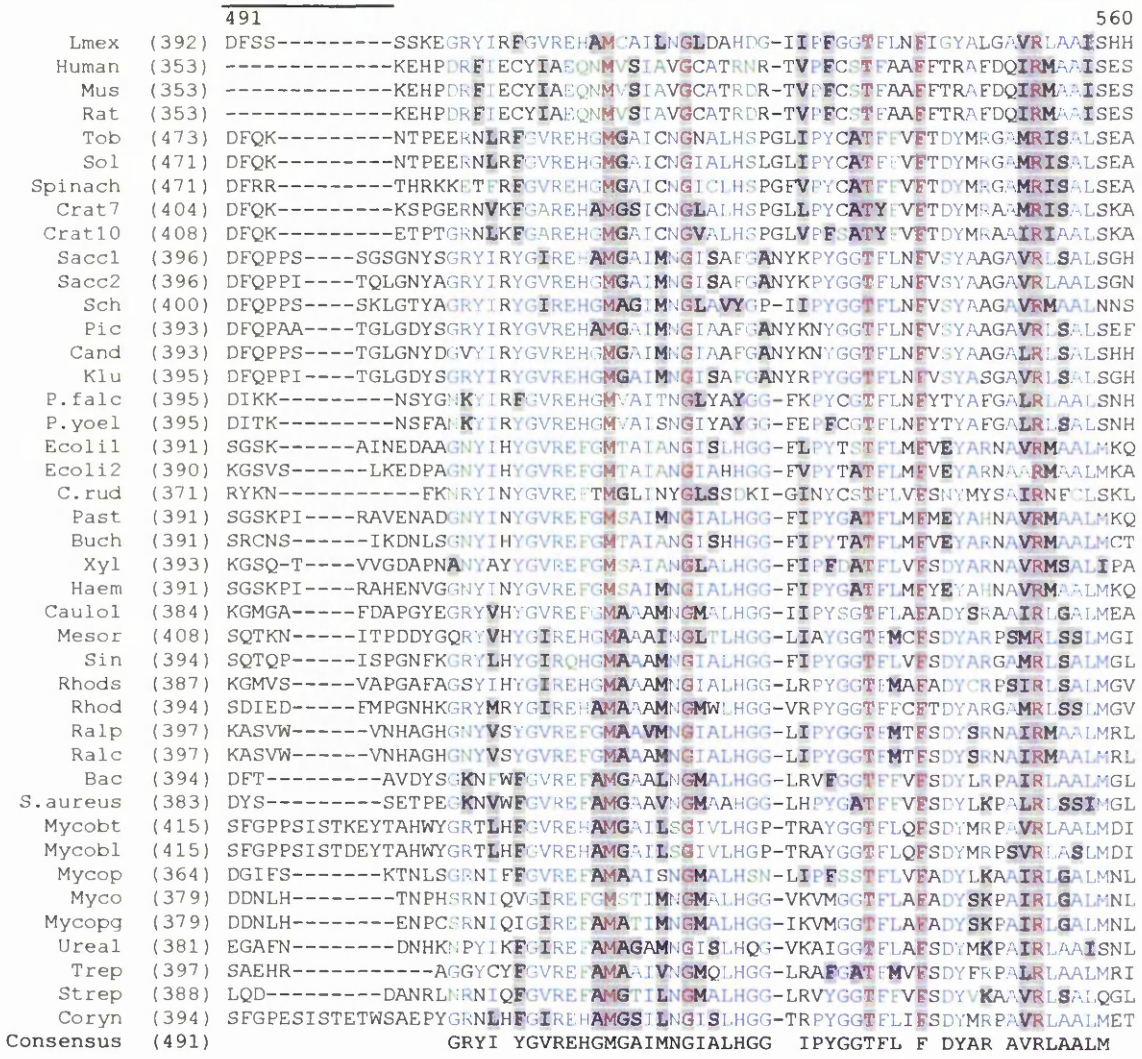


Figure 4.15 cont.: Alignment of transketolase amino acid sequences

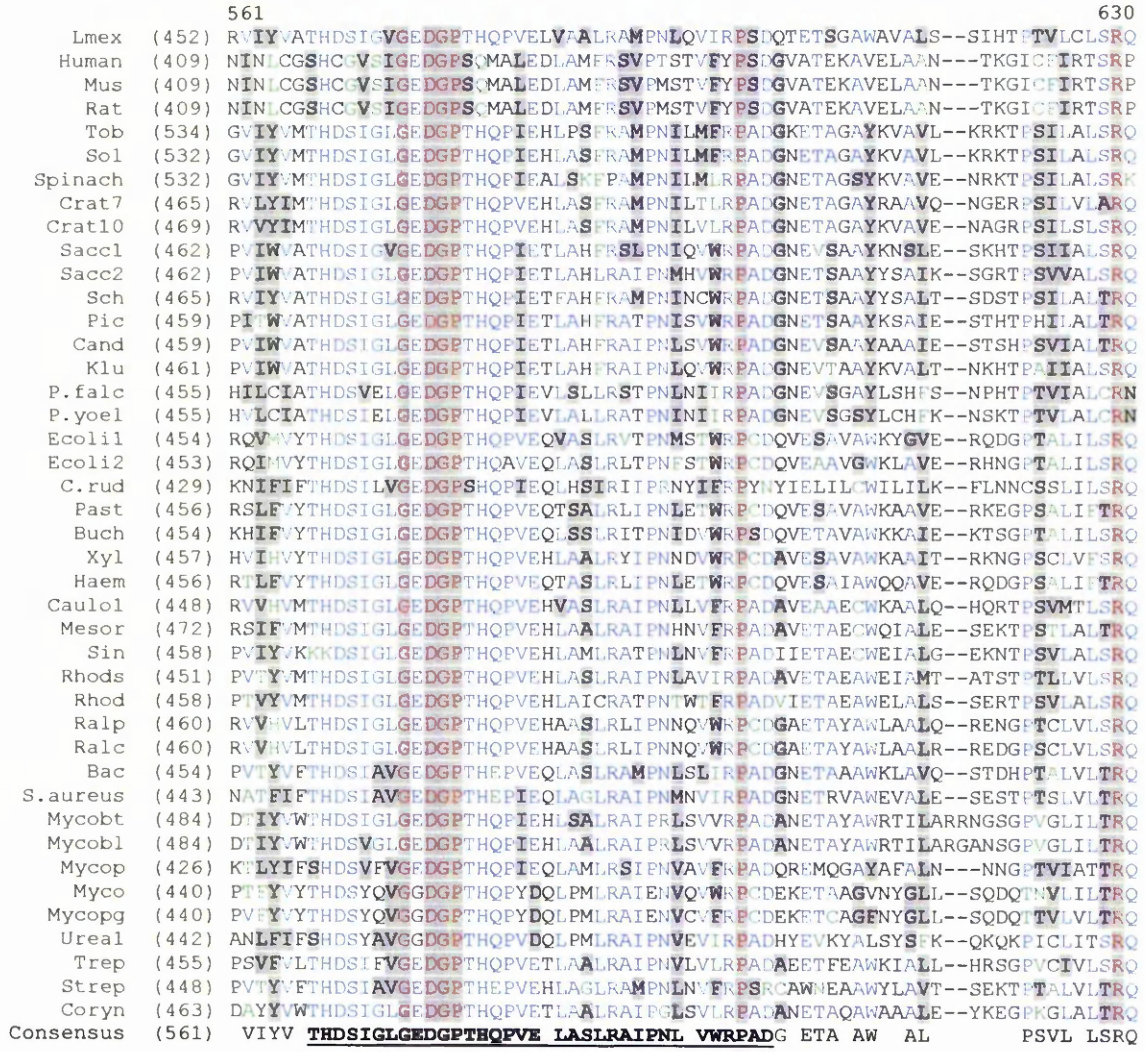


Figure 4.15 cont.: Alignment of transketolase amino acid sequences

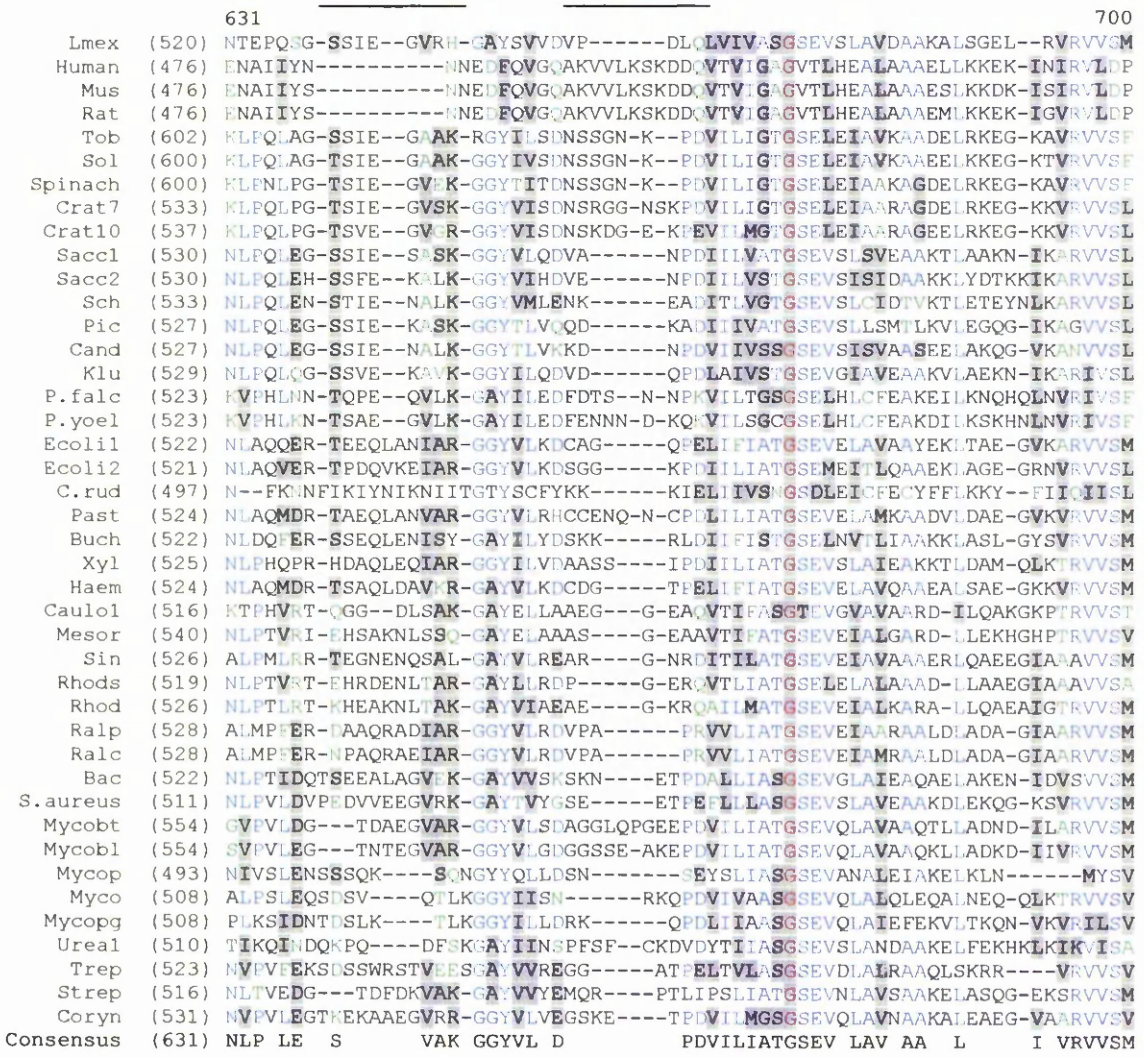


Figure 4.15 cont.: Alignment of transketolase amino acid sequences

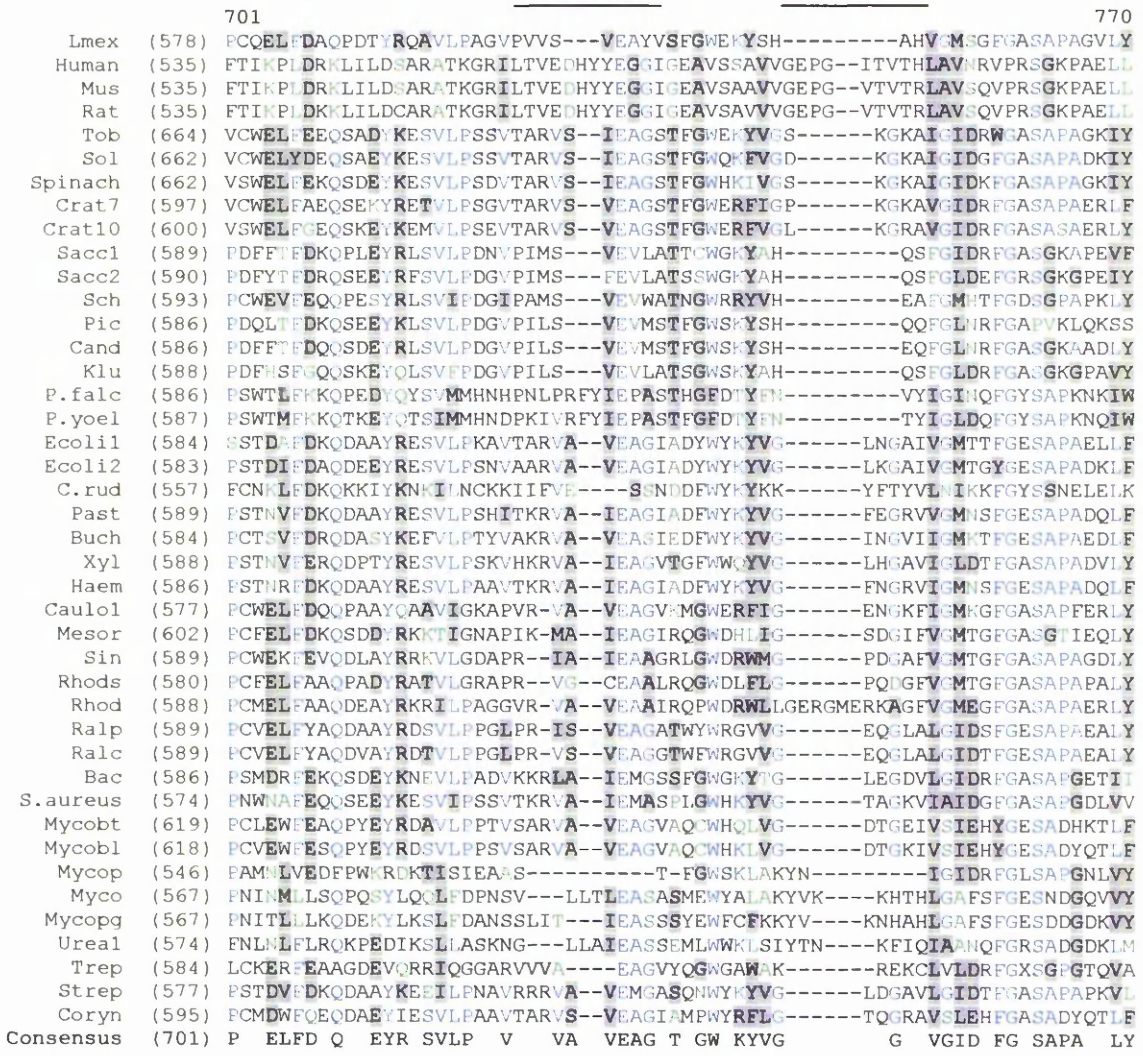


Figure 4.15 cont.: Alignment of transketolase amino acid sequences

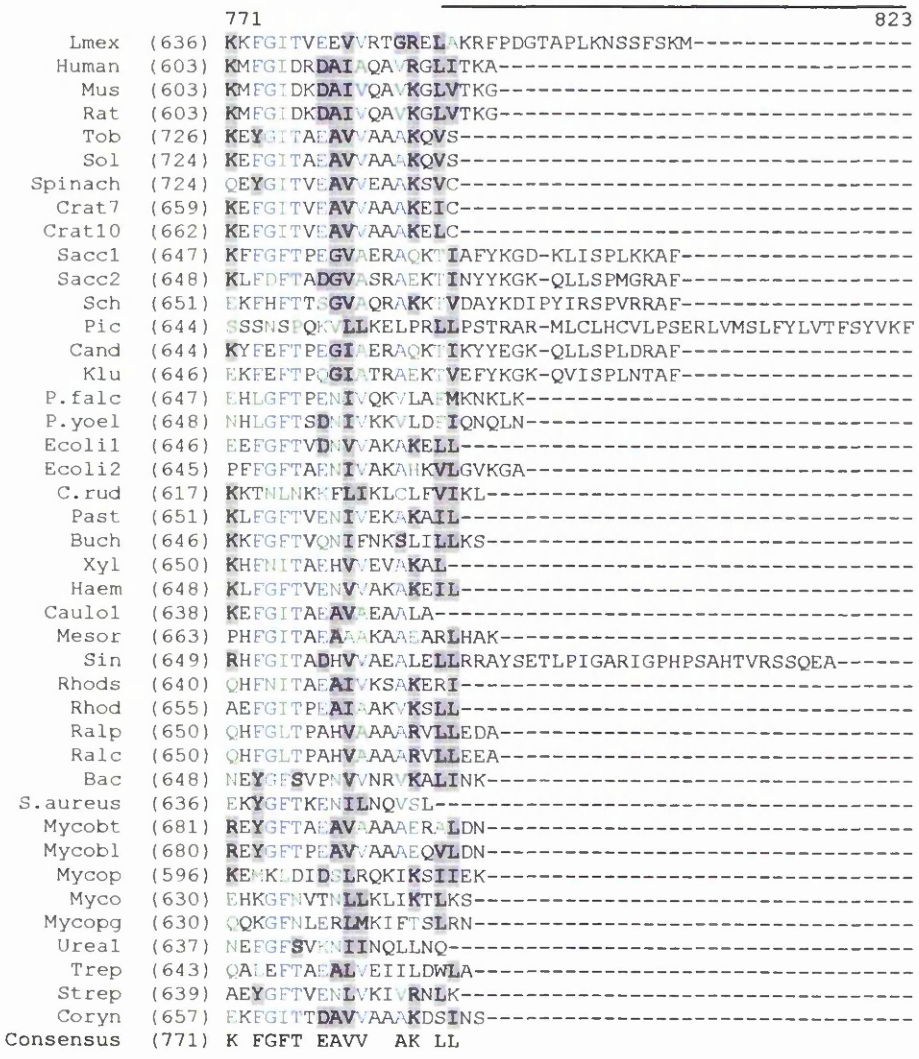


Figure 4.15 cont.: Alignment of transketolase amino acid sequences

4.8. Invariant residues of transketolase

The amino acid sequence for *L. mexicana* TKT (Lmex) was compared to *Saccharomyces cerevisiae* transketolase 1 (Sacc1) using the Clustal W alignment algorithm (Figure 4.16). Nikkola (Nikkola *et al.*, 1994) identified 41 invariant residues in transketolase, all of which are present in the *L. mexicana* TKT protein sequence and are shown in Figure 4.14 as underlined residues in the consensus sequence. Only 11 out of the 41 invariant residues were completely conserved in the alignment in Figure 4.15. 13 invariant amino acids were identified from this alignment that were not mentioned by Nikkola (Nikkola *et al.*, 1994), and are highlighted in Figure 4.16 with an asterisk below the residue.

The transketolase sequence from yeast and *Leishmania* have 50.6 % identity and 65.5 % similarity. 14 invariant residues have been investigated for function by site-directed mutagenesis in yeast, all of which are highlighted in Figure 4.16.

The 3-dimensional crystal structure of *S. cerevisiae* transketolase has been determined to 2.5 Å resolution (Lindqvist *et al.*, 1992) and subsequently refined to 2.0 Å resolution (Nikkola *et al.*, 1994). The crystal structure has provided information regarding the interactions of the enzyme subunits with the cofactors TPP and $MgCl_2$, and the substrates. Several interactions between amino acid residues on the surface of the transketolase protein and the cofactors were recognised. The holoenzyme is a dimer and the active site is created between the two-monomer subunits with the cofactor TPP located inside the cleft. Following the crystallisation of transketolase in yeast, various studies have been done involving the site directed mutagenesis of invariant amino acid residues, providing an insight into their function (Lindvist *et al.*, 1992). The residues discussed are highlighted in Figure 4.16.

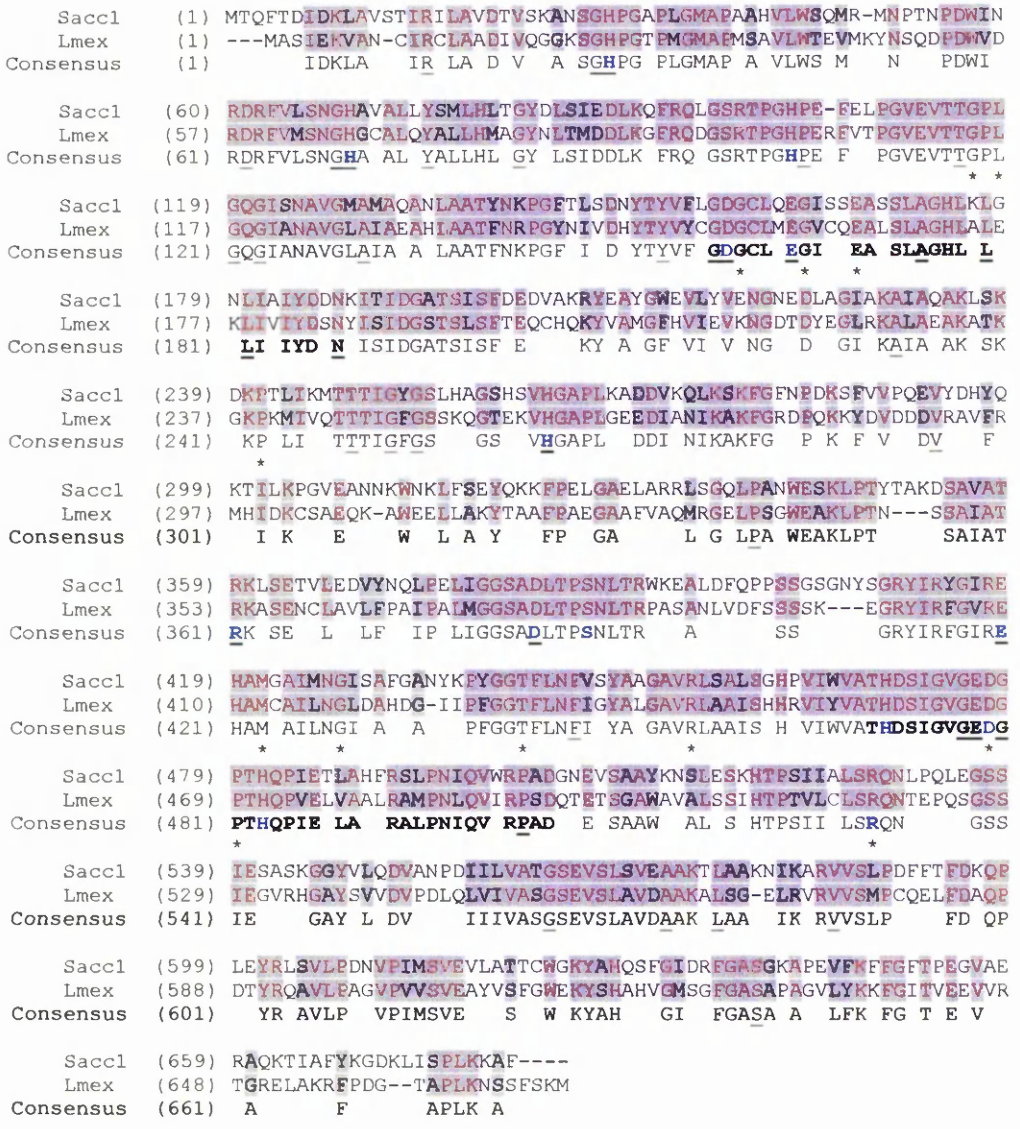


Figure 4.16: Alignment of the *L. mexicana* transketolase protein and the *S. cerevisiae* transketolase 1 protein. The numbers in brackets following the shortened name correspond to the amino acid number of the alignment. The amino acid residues that are identical in the alignment are highlighted in red type with a grey background. The amino acids that are similar in physiochemical properties are in black type with grey background. The letters in the consensus sequence that are invariant amino acid residues (as identified by Nikkola *et al.*, 1994) are underlined. The TPP binding motif and the TKT conserved motif are highlighted in bold text in the consensus sequence. The conserved residues that were identified from the alignment in Figure 4.15 are highlighted with an asterisk* below the residue. The amino acid discussed within the text regarding site-directed mutagenesis studies are highlighted in blue bold text.

4.8.1. Analysis of residues involved in TPP binding

The invariant residue Glu^{418} side-chain forms a hydrogen bond to the N-1' nitrogen atom of the pyrimidine ring of TPP (Lindqvist *et al.*, 1992). To investigate the role of this interaction with the cofactor, Glu^{418} was replaced with glutamine and alanine by site-directed mutagenesis (Wikner *et al.*, 1994). The hydrogen bond was shown not to be required for cofactor binding, but essential for catalytic activity. It was suggested that this hydrogen bond was involved in a cofactor-assisted deprotonation of the C-2 carbanion, the first step in thiamine catalysis. Two other conserved amino acid residues, Glu^{162} and Asp^{382}, present in the TPP binding site, were subject to site-directed mutagenesis studies (Meshalkina *et al.*, 1997). Glu^{162} is present in a cluster of glutamic acid residues present in the dimer close to the cofactor. Glutamine and alanine mutants of Glu^{162} displayed increased lag phases thus suggesting that Glu^{162} may be important for the dimerisation of transketolase. Asp^{382} is thought to compensate for the positive charge of the thiazolium ring of TPP, as it is the only conserved polar amino acid near it (Lindqvist *et al.*, 1992). The mutants D382A and D382N exhibit impaired TPP binding and decreased catalytic activity. This suggests that Asp^{382} is important in cofactor binding that may affect catalysis. Arg^{359}, Arg^{528} and His^{469} are thought to interact with the phosphate groups of the TPP molecule at the entrance to the cleft of the active site (Nilsson *et al.*, 1997). Site-directed mutagenesis of these residues resulted in increased K_m values for the acceptor substrate (50 times greater) and donor substrates (5-10 times greater), suggesting that these residues are important in substrate binding (Nilsson *et al.*, 1997). All of the residues investigated are present in the *L. mexicana* TKT protein sequence suggesting that this protein may be similar to the yeast enzyme in binding the cofactor.

4.8.2. Analysis of residues involved in substrate interactions

The histidine residues located at positions 30, 69, 103, 263 and 481 of the *S. cerevisiae* transketolase amino acid sequence is a cluster of conserved histidines, present at the active site where the C2 carbon of the thiazolium ring of TPP resides (Nikkola *et al.*, 1994). The invariant active-site residues His^{103} and His^{69} are thought to be involved in the binding of donor substrates, as they are within hydrogen bonding distance of the hydroxyl group on the donor substrate (Nikkola *et al.*, 1994). This proposal of function was confirmed when His^{103} was replaced with alanine, asparagine and phenylalanine using site-directed mutagenesis (Wikner *et al.*, 1995) and His^{69} was replaced with alanine (Wikner *et al.*, 1997). The mutants exhibited reduced catalytic activity and a 14 to 60-fold increase in K_m for the donor substrate xylulose 5-phosphate. The His^{103} mutants also had an increased K_m for the cofactor TPP. The mutants had a K_m of the acceptor substrate ribose 5-phosphate comparable to that of the wild-type enzyme. These findings implicate His^{103} and His^{69} in donor substrate binding and help to orientate the donor and the cofactor TPP. The mutation of His^{263} resulted in a transketolase with reduced catalytic activity and is involved in the deprotonation of the C3-hydroxyl group of the donor substrate during cleavage (Wikner *et al.*, 1997; Fiedler *et al.*, 2001). The crystallisation of *S. cerevisiae* transketolase-1 protein associated with TPP, Ca^{2+} and the acceptor substrate erythrose 4-phosphate, demonstrated that the substrate binds deeply into a cleft between the individual subunits of the enzyme (Nilsson *et al.*, 1997). The residues His^{30} and His^{263} form hydrogen bonds with the aldehyde oxygen (Nilsson *et al.*, 1997) and site directed mutagenesis of these residues results in decreased affinities for the substrates (Nilsson *et al.*, 1997; Wikner *et al.*, 1997). Asp^{477} is thought to form a hydrogen bond with the C2-hydroxyl group of the donor substrate and is replacement of this substrate results in impaired catalytic activity

(Nilsson *et al.*, 1997; Nilsson *et al.*, 1998). All of these residues that are important for substrate binding and catalysis are present in the *L. mexicana* TKT protein sequence. This analysis suggests that the mechanism by which the substrates interact with transketolase may be similar in both the *L. mexicana* and yeast transketolase.

4.9. Oligomeric structure of the *L. mexicana* transketolase protein

The *L. mexicana* transketolase protein has been submitted for crystallisation trials to Dr.Ylva Lindqvist at the Karolinska Institute, Sweden. The protein structure has been resolved to 3 Å and molecular replacement studies are currently being undertaken.

4.10. Discussion

The specific activity of the *L. mexicana* native transketolase proteins was 0.058 U/mg of total protein. The measurement of transketolase activity in crude lysates of the parasite has a high background of NADH oxidation occurring due to the number of other proteins present in the sample. This was taken into account when determining the activity of the protein. The measured specific activity of the *L. mexicana* transketolase is similar to the previously measured values from *L. mexicana* and *T. brucei* (Martin *et al.*, 1976; Cronin *et al.*, 1989) (see Section 4.2).

The recombinant *L. mexicana* TKT has a specific activity of 1.65 U/mg. The over-expression of *L. mexicana* recombinant soluble TKT protein was achieved only when the temperature of induction was 15 °C. Increasing the yield of the exogenous protein to 5-8 % of the total soluble lysate was achieved with overnight incubation of the bacterial culture in the presence of 0.1 mM IPTG. This transketolase was not over-expressed when induced at 37 °C and was insoluble when expressed at 25 °C. This may be due to the intrinsic folding mechanisms of the protein and the possible sensitivity of the expression system. At 25 °C the protein may have been aggregating due to the high rate of over expression of the exogenous protein and therefore remaining insoluble. The specific activity of the *L. mexicana* transketolase is low in comparison to other recombinant transketolases such as the *E. coli* and the human transketolase (see Table 4.5). This could be due to the expression system used, with the soluble protein perhaps not folding correctly prior to purification, producing a partially inactive preparation.

L. mexicana TKT does not require pre-incubation or dialysis with the cofactors TPP and $MgCl_2$ for optimal activity subsequent to purification using a nickel-affinity column. This may suggest that TPP was acquired from the *E. coli* subsequent to the

translation of transketolase and remains tightly bound to the protein during the purification procedure. Pre-incubation of transketolase with 100 mM TPP inhibits the transketolase reaction. TPP is known to sit in the active site of the protein, and is responsible for binding the donor substrate of the reaction (Lindqvist *et al.*, 1992). Therefore, a high concentration of TPP may cause a blockage of the active site and decrease the activity of the reaction. The *L. mexicana* TKT protein had an optimal pH of 7.5 in Tris buffer and could be stored for several weeks at –80 °C without glycerol. The K_m and the V_{max} of R5P for the *L. mexicana* TKT is 2.75 mM and 1.7 μmol/min/mg protein, respectively. The K_m values are similar to transketolases from other organisms (see Table 4.5).

X5P was successfully made in-house using a biotransformation method utilising the *E. coli* TKA. This method was extremely cost-effective in comparison to purchasing commercial X5P, however the method was time-consuming and the purification was problematic. The preparation did contain DHAP, a product made during the biotransformation, however was reduced enzymatically before using the X5P to assay transketolase.

The *L. mexicana* TKT protein primary sequence has been analysed using various methods. The primary sequence was aligned with other known transketolase protein sequences and examined. *L. mexicana* TKT has been submitted for crystallisation trials. Two new transketolase protein sequences from *Plasmodium falciparum* and *Plasmodium yoelli* were identified from the *Plasmodium* genome-sequencing database (PlasmoDB). The alignment of 42 full-length transketolase amino acid sequences indicates that the TPP binding motif and TKT motif are highly conserved regions within transketolases. The extended N-terminal domains present in certain plant transketolases are likely to contain chloroplast targeting signals. The direct comparison of *S.*

cerevisiae transketolase-1 and *L. mexicana* TKT showed that the sequences are highly conserved, with 50.6 % identity and 65.5 % similarity. All the invariant residues previously discussed in the literature are present in the *L. mexicana* amino acid sequence. Various invariant amino acid residues of the transketolase dimer have been recognised as important in cofactor and substrate interactions through the crystallisation studies and site-directed mutagenesis of transketolase from *S. cerevisiae*. The interactions of the residue side-chains in the active site cleft and close to the TPP and substrate binding regions are often essential for complete catalytic activity and the accurate affinity for binding. These studies reflect the complexity of the transketolase reaction, with the cofactors playing an essential role in the orientation of substrate binding.

Chapter V

Localisation and phylogenetic analysis of a transketolase protein in *Leishmania mexicana*

5.1. Introduction

Transketolase has been shown to be present in most organisms including mammals, plants, yeasts, protozoa and bacteria. In most organisms transketolase is present in the cytosol however, this protein has been localised to the chloroplast in some plant species and the endoplasmic reticulum in humans. The discovery of the peroxisome-targeting sequence (PTS) at the C-terminal of the *L. mexicana* transketolase protein suggested that this enzyme was perhaps glycosomal or displayed a bi-compartmental distribution. To address this, subcellular localisation and immunoflourescence experiments were performed and these are presented in this chapter. Different life-cycle stages of *L. mexicana* were examined for the presence of transketolase to determine if transketolase is present in the human host stage of the disease.

Genome database searching has provided several putative transketolase protein sequences from various organisms. Fragments of protein sequence from trypanosomatids with similarity to transketolase were aligned with the *L. mexicana* transketolase in order to determine if they also had a peroxisome-targeting sequence. Phylogenetic analysis of several transketolase sequences will give insight into the possible ancestry of the *L. mexicana*. A particular family of transketolases, the 1-deoxy-D-xylulose 5-phosphate synthases, are also subjected to phylogenetic analysis in this chapter and compared with other transketolase sequences. Xylulose 5-phosphate

synthase is an enzyme that is involved in exchanging a two-carbon unit from a ketose to an aldose and therefore is strictly a transketolase.

5.2. Expression of transketolase in different life-cycle stages of *L. mexicana*

Polyclonal antibody to the purified recombinant *L. mexicana* TKT protein was raised in a rabbit. Three bleeds were obtained, and dot-blots were performed for each bleed to determine the specificity of the antibodies. The third and final bleed contained antibodies highly specific for the transketolase protein in parasite crude lysates, and was used in all the experiments in this chapter. The antibody was used at a dilution of 1/10,000. To determine if transketolase is present in both of the life-cycle stages of *L. mexicana*, a western blot was performed with protein lysates from the promastigotes, amastigotes from a lesion from a mouse infected with *L. mexicana* and axenic amastigotes. Figure 5.1a shows that transketolase is expressed at approximately the same level in promastigotes and axenic amastigotes, however in this experiment it was not detected in amastigotes from lesions. An identical Western blot was probed with antibodies to an α-mercaptopyruvate sulfurtransferase protein (40 kDa) from *L. major* obtained from Roderick Williams, University of Glasgow. This protein is expressed to the same level in both life-cycle stages (Roderick Williams, personnel communication). The amastigotes obtained from lesions in mice did not produce a signal with transketolase or the control sulfurtransferase protein in this experiment. This was seen on multiple occasions and is perhaps due to a higher quantity of proteinase enzymes in the amastigotes stage of the parasite that degrades proteins in the crude lysates. However, proteinase inhibitors were used during the lysis of the crude extracts and no degradative bands are seen on the blot, which may be expected if proteolysis occurred. This result suggests that the transketolase protein is expressed both in the promastigote

stage and the axenic amastigotes stage of the parasite at the same level, however it cannot be shown in this experiment to be present in the amastigotes from lesions.

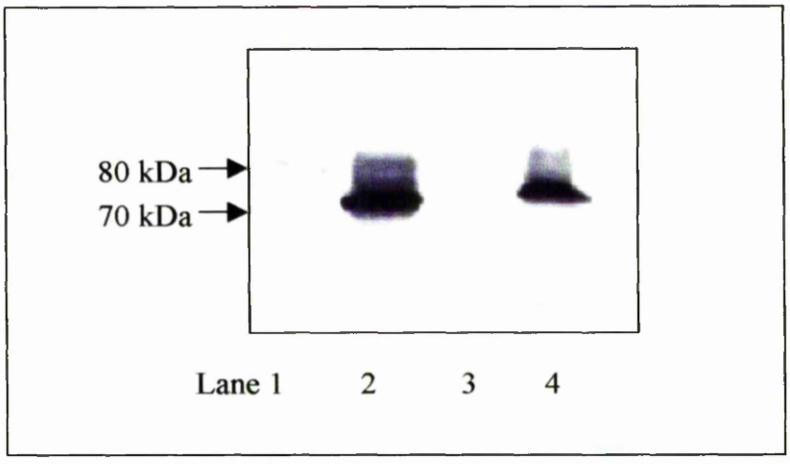


Figure 5.1a: Western blot showing transketolase expression in different life cycle stages of *L. mexicana*. Lanes 2, 3 and 4 contain 50 μg of protein. Promastigote, amastigotes and axenic amastigotes protein is present in these lanes respectively. The faint band in lane one corresponds to an 80 kDa marker protein and 70 kDa is indicated below this.

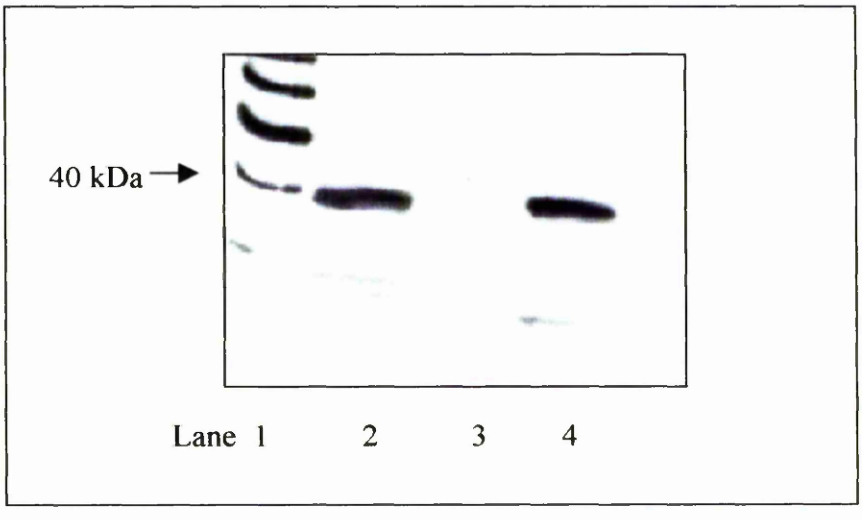


Figure 5.1b: Western blot showing expression of sulphur transferase in *L. mexicana*. The lanes in this figure correspond to those in Figure 5.1.a. Lane 1 shows various bands corresponding to the sizes indicated to the left of the blot.

5.3. Subcellular localisation of transketolase in *L. mexicana*

Transketolases are mainly present in the cytosol of most organisms, however there are examples of these proteins in organelles. Various plant transketolases contain chloroplast targeting sequences and have been localised to the chloroplast (Teige *et al.*, 1998). Mammalian cells are thought to have two sets of pentose phosphate pathway enzymes, one is cytosolic and the other is located in the endoplasmic reticulum (Bublitz and Steavenson, 1988). The pentose phosphate pathway enzyme 6PGL has recently been cloned and sequenced from *T. brucei* and was shown to contain the C-terminal PTS, AKF (Duffieux *et al.*, 2000). Subcellular localisation experiments showed that 10 % of 6PGL activity was associated with the glycosomal fractions. The G6PDH in *T. brucei* was also partially localised to the glycosome, however was shown not to contain a PTS (Duffieux *et al.*, 2000; Heise and Opperdoes, 1999). This data suggests that the pentose phosphate pathway may function in the glycosomes and the cytosol in *T. brucei* (Barrett, 1997). The *L. mexicana* transketolase protein sequence has a C-terminal PTS suggesting it may be glycosomal or may have a bi-compartmental distribution.

To determine in which compartment of *L. mexicana* promastigotes transketolase resides, subcellular fractionation experiments were performed in collaboration with Dante Maugeri and Prof. Juan Jose Cazzulo in Buenos Aires.

5.3.1. Enzyme analysis

The fractionated cells were assayed for enzyme activity immediately after the subcellular fractionation. The enzyme activities were recorded for each fraction in Table 5.1a and are expressed as percentages of the total homogenate activity in Table 5.1b. Hexokinase and pyruvate kinase are marker enzymes for the glycosomes and the cytosol, respectively. Hexokinase was previously shown to be mainly glycosomal in *L.*

mexicana, with 12 % of the activity being present in the soluble fraction (Coombs *et al.*, 1982). Pyruvate kinase has been shown to be mainly cytosolic in *L. mexicana*, however 1-3 % of activity was recovered from all of the other fractions (Coombs *et al.*, 1982).

79 % of transketolase activity is present in the soluble fraction and 8.8 % in the granular fraction (Table 5.1a and 5.1b). The granular fraction exhibits 60 % of the total hexokinase activity and 2.9 % of the pyruvate kinase activity, indicating that the granular fraction mainly contains the glycosomal fraction with a small cytosolic contamination (Table 5.1a and 5.1b).

Enzyme measured	**Marker**	**Total homogenate**	**Soluble Fraction**	**Granular fraction**	**Pellet**	**Microsomal**	**Recovery (%)**
		Units/total					
Hexokinase	Glycosome	10.42	1.14	6.2	0.582	0.580	78
Pyruvate Kinase	Cytosol	11.26	9.54	0.329	0.133	0.0911	89.6
Transketolase	N.A.	2.76	2.2	0.244	0	0	96

Table 5.1a: The activity of enzymes present in the subcellular compartments of *L. mexicana*. The fractions were obtained through subcellular fractionation using centrifugation. Activity is recorded in the total number of units (μg/min/ml) obtained from each fraction. The recovery represents the sum of the enzyme activity present in the individual fractions over the total homogenate and is represented by % recovery of activity. A breakdown of the recoveries in each individual fraction is shown in Figure 5.1b. N.A. means not applicable.

Enzyme	**Marker**	**Soluble**	**Granular**	**Pellet**	**Microsomal**
	Recovery (%)				
Hexokinase	Glycosome	10.9	60	5.5	5.5
Pyruvate Kinase	Cytosol	85	2.9	1.1	0.8
Transketolase	N.A.	79	8.8	0	0

Table 5.1b: The recovery of enzyme activity in individual subcellular fractions. The recovery is shown as a percentage (calculated as in Table 5.1a). N.A. stands for not applicable.

The granular fraction was further fractionated by isopycnic centrifugation and the individual fractions were assayed for enzyme activity. Transketolase displays two distinctive peaks of activity (Figure 5.2). The majority of the transketolase activity corresponds to the cytosolic marker that mainly resides at the top of the gradient as indicated by the pyruvate kinase activity. A clear second peak of transketolase activity corresponds to the hexokinase activity peak in the sucrose gradient at a density of 1.241 g/cm^3. However the glycosomal fractions are contaminated with the cytosolic enzyme pyruvate kinase, and therefore it cannot be concluded with certainty that the transketolase peak is glycosomal.

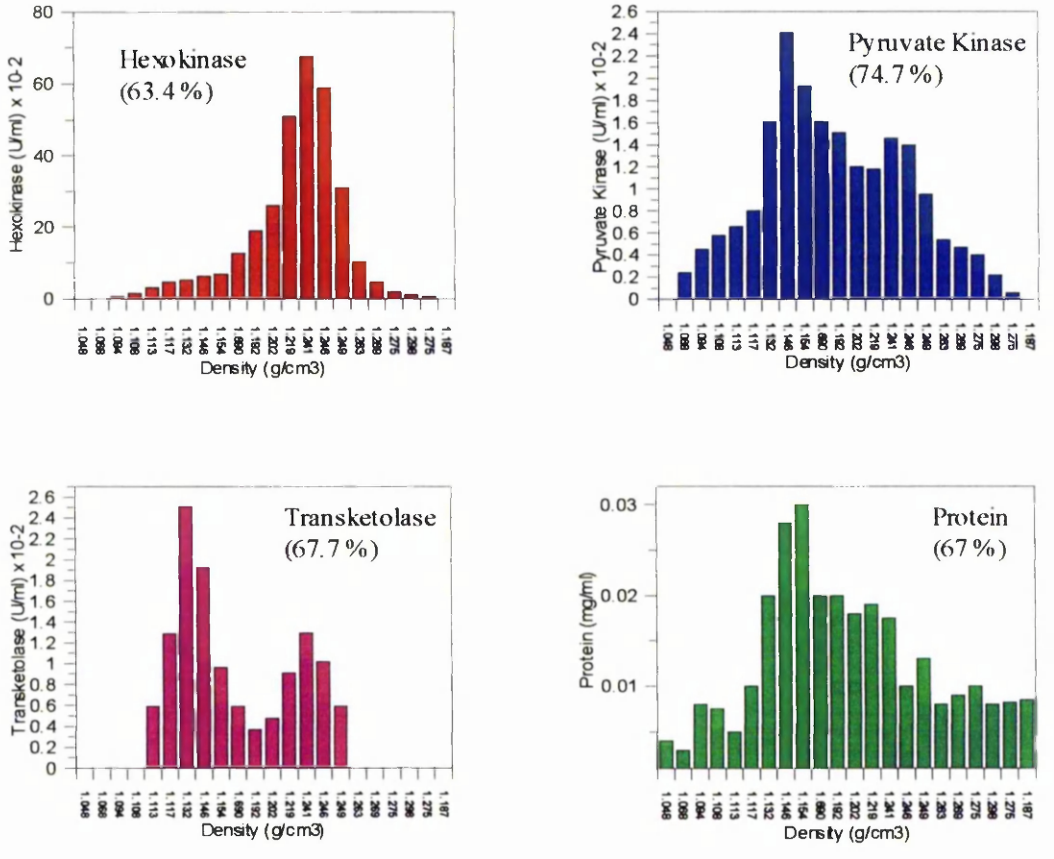


Figure 5.2: The density distribution of *L. mexicana* subcellular marker enzymes, transketolase and protein content. The x-axis shows the activity (U/ml) of the individual enzymes and the concentration (mg/ml) of protein. The sucrose gradient fractions are represented in density of sucrose (g/cm^3). The recoveries of the enzymes and protein from the sucrose gradient are given in brackets. The recovery represents the sum of the activities in the sucrose gradient fractions over the activity found in the granular fraction applied to the sucrose gradient.

5.3.2. Western blot analysis

A Western blot of the subcellular fractions with antibodies to transketolase from *L. mexicana* shows that transketolase is present in all fractions of the subcellular fractionation (Figure 5.3). The Western blot shows equal volumes of loading and not an equal volume of protein, therefore the fractions cannot be compared for levels in different compartments. The individual fractions of the sucrose gradient were Western blotted with transketolase antibodies. This shows a larger amount of transketolase protein being present in Lanes 8 and 9 (Figure 5.3b) corresponding to the hexokinase activity peak in Figure 5.2. A possible explanation for this discrepancy between the Western blot and the transketolase enzyme activity is that a proportion of the transketolase enzyme may have been inactive in the preparation. In a repeat of this experiment, the enzyme activity was greater in the peak corresponding to the hexokinase activity peak. Antibodies to a purified native hexokinase in *T. brucei* were obtained from Paul Michels, Brussels, however these did not cross-react with *L. mexicana* lysates. These results suggest that transketolase is mainly present in the cytosol of *L. mexicana* promastigotes. This suggests that although the protein sequence has the PTS, the promastigote parasites only transport at most small quantities of the enzyme into to the glycosome in the insect stage of the life-cycle. It cannot be ruled out that in other life-cycle stages of this parasite, a glycosomal localisation does occur.

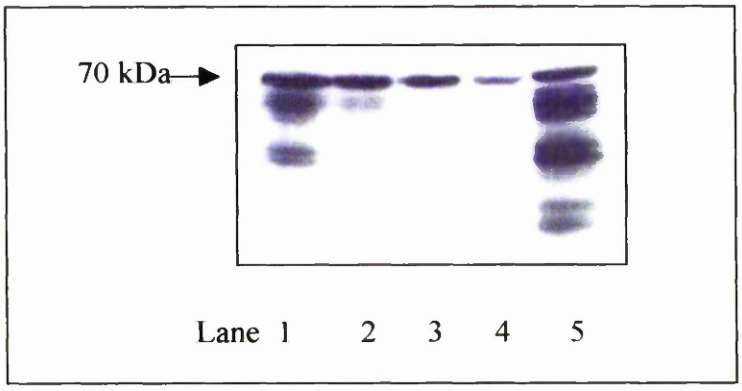


Figure 5.3a: Western blot of the subcellular fractionation of *L. mexicana* promastigotes reacting with transketolase antibodies. The fractions are the total homogenate (Lane 1), the pellet (Lane 2), the cytosol (lane 3), the granules (Lane 4) and the microsome (Lane 5).

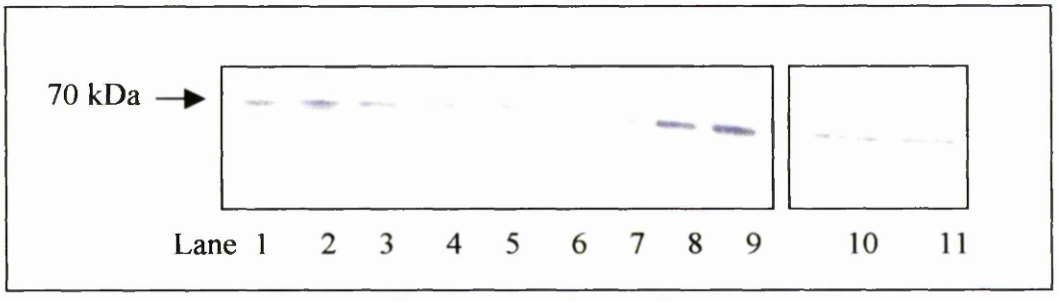


Figure 5.3b: Western blot of the individual fractions of the sucrose gradient reacting with transketolase antibodies. Lanes 1 to 11 correspond to an increasing density of sucrose from 1.113 – 1.248 g/cm^3.

5.4. Immunoflourescence of transketolase in *L. mexicana* and *T. cruzi*

Indirect immunoflourescence of transketolase antibodies in *L. mexicana* promastigotes shows a uniform distribution of transketolase over the cell (Figure 5.4). This also indicates that the transketolase protein occurs mainly in the cytosol of the parasite. A dilution of 1/5000 of the primary transketolase antibody was required to obtain a fluorescent image. Antibodies to hexokinase from *T. brucei* from Paul Michels, Brussels, did not cross-react with the *L. mexicana* promastigotes (at a dilution of 1/200), and therefore this glycosomal marker could not be used as a control in this experiment. *T. cruzi* epimastigotes (insect stage) were used also during this experiment and were incubated with transketolase antibodies (using a 1/1000 dilution). The hexokinase antibodies from *T. brucei* did cross-react with *T. cruzi* at a dilution of 1/1000 and therefore they were used as a control. This was determined by Western blotting of fractionated *T. cruzi* with the hexokinase antibody. This cross-reactivity suggests that the *T. brucei* hexokinase protein is more similar to the *T. cruzi* hexokinase protein than the *L. mexicana*. When incubated with hexokinase antibodies *T. cruzi* shows a punctate fluorescence consistent with the protein being localised to the glycosomes (Figure 5.5). Transketolase antibodies show a uniform distribution in the epimastigote form of *T. cruzi*, similar to the transketolase in *L. mexicana*, suggesting that this protein is mainly cytosolic in this organism too. Control experiments did not show any detectable background immunoflourescence under the flourescence microscope used.

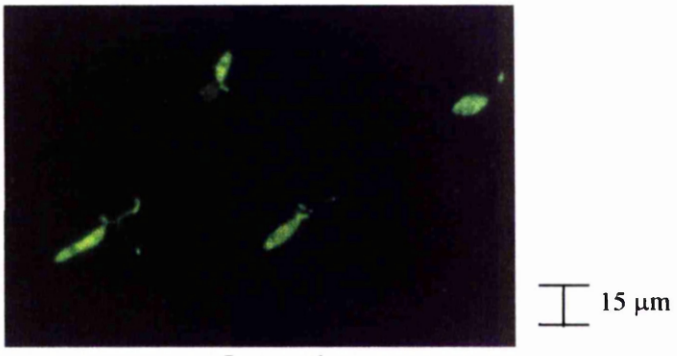


Image 1

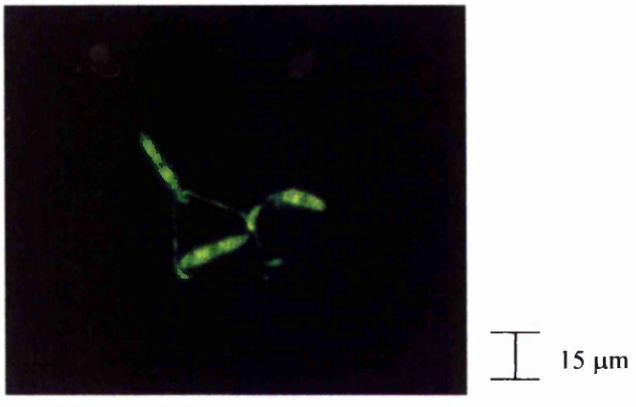


Image 2

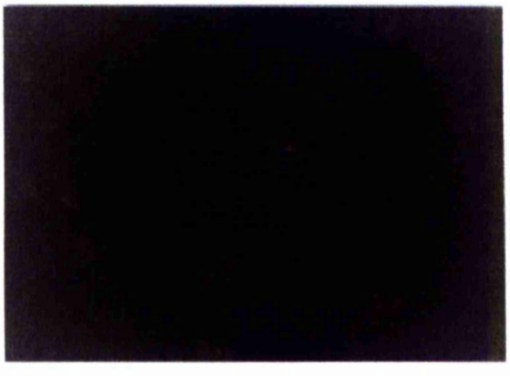
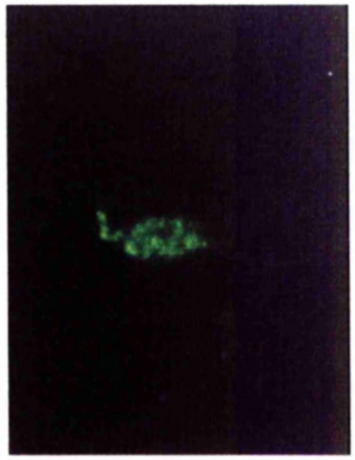


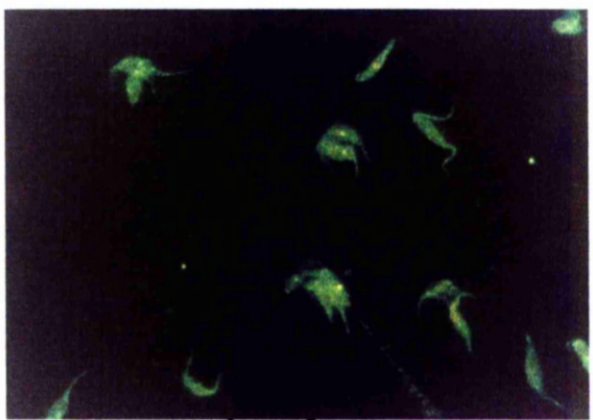
Image 3

Figure 5.4: Indirect immunoflourescence of transketolase antibodies in *L. mexicana* promastigotes. Image 1 and 2 are both from the same slide and show the parasites after incubation with transketolase antibodies. Image 3 shows a negative control in which the primary antibody was excluded from the method. All images have an exposure time of 15 seconds.



10μm

Image 1



20 μm

Image 2

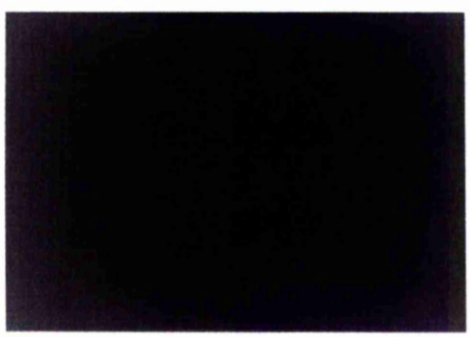


Image 3

Figure 5.5: Indirect immunoflourescence of transketolase and hexokinase antibodies in *T. cruzi* epimastigotes. Image 1 shows *T. cruzi* incubated with hexokinase antibodies, Image 2 shows *T. cruzi* after incubation with transketolase antibodies. Image 3 shows a negative control in which the primary antibody was excluded from the method. All images have an exposure time of 15 seconds.

5.5. Identification of the peroxisome-targeting signal present on the *L. mexicana* transketolase protein

The *L. mexicana* transketolase protein contains a potential PTS at the C-terminus with the residues SKM. This targeting signal has been shown to be present on the glyceraldehyde-3-phosphate dehydrogenase (G3PDH) protein localised to the glycosome in *T. brucei* (Hannaert *et al.*, 1992). To determine if the transketolase amino acid sequences in other trypansomatids also contain a peroxisome targeting signal, the parasite genome databases were searched and the sequences were aligned with the *L. mexicana* TKT protein sequence. The *L. mexicana* TKT protein sequence was aligned using the Clustal W algorithm with partial transketolase sequences from *Leishmania major*, *Trypanosoma brucei* and *Trypanosoma cruzi* (Figure 5.5). The *L. major* 3' transketolase sequence was obtained by sequencing of cosmid *L. major* DNA obtained from Dr. Al Ivens (Sanger Centre) (details of primers and conditions used for PCR in Section 3.3.3). *T. brucei* and *T. cruzi* DNA sequence with similarity to transketolase was obtained by searching the kinetoplastid database (EMBL) for homology to the *L. mexicana* TKT protein. The PTS is present in the putative *L. major* transketolase 3' sequence, with identical residues to the *L. mexicana* TKT protein. The *T. brucei* transketolase 3' gene fragment also contains a PTS with the residues SHL. This PTS is also present on the *L. mexicana* glucosephosphate isomerase (PGI) (Nyame *et al.*, 1994). The sequencing information from *T. cruzi* does not cover the 3' region of the gene. This analysis of putative transketolases reveals that *L. mexicana*, *L. major* and *T. brucei* all contain a PTS at the C-terminal of the protein.

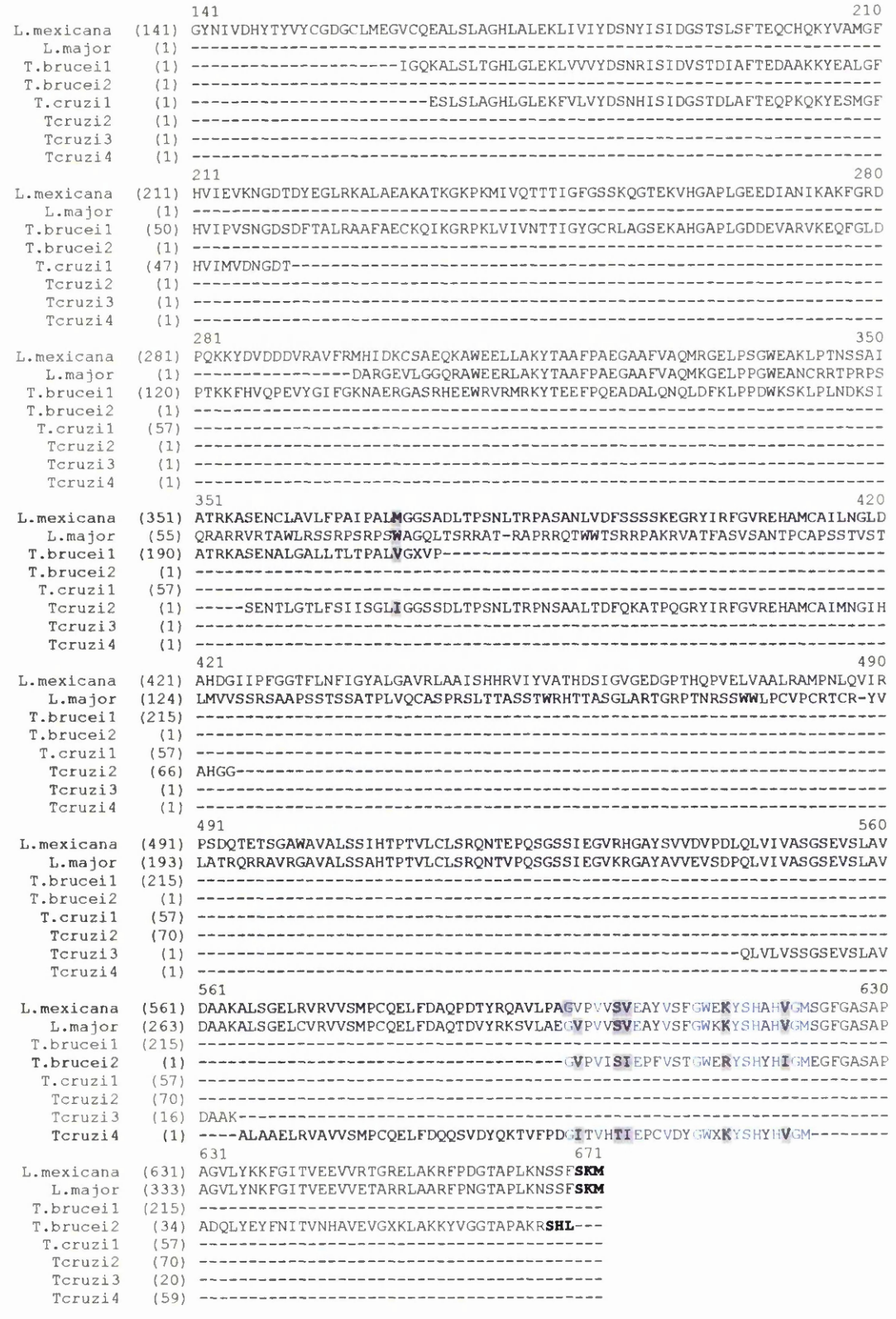


Figure 5.6: Alignment of the L. mexicana transketolase protein sequence with putative partial trypanosomatid transketolase sequences. Legend over the page.

Figure 5.6. Legend. The similarity searching was completed in December 2001. The *L. mexicana* transketolase sequence starts at amino acid 141 as there are no fragments from other Trypanosomatids that have homology to and align before this residue. A block of similar residues is indicated by black text with a grey background, and conserved residues are indicated by blue text. The numbers in brackets following the name of the sequence represents the amino acid position of the individual sequences. The *T. cruzi* sequences referred to as T. cruzi 1 and 2 are from one EST (EMBL accession number AI668722) and T. cruzi 3 and 4 are also from one EST (EMBL accession number AW330226). The *T. brucei* contig referred to as T. brucei 1 was obtained by aligning two sequences from sheared genomic DNA clones (EMBL accession numbers AL453110 and AQ948814). The T. brucei sequence T. brucei 2 was obtained from sheared genomic DNA clone (EMBL accession number AQ941039). The DNA sequences were translated using Vector NTI and aligned using Align X (Clustal W algorithm). The C-terminal PTS residues are highlighted in bold text.

5.6. Phylogenetic analysis of transketolase amino acid sequences

Evidence is emerging that glycosomes and other microbodies could be the remnants of early endosymbiotic events (Latruffe and Vamecq, 2000). Having shown that the transketolase from *L. mexicana* has a peroxisome targeting sequence and may in part be localised to the glycosome, it was decided to conduct a detailed phylogenetic analysis on this protein.

To determine the protein sequence similarity and clustering relationship of the *L. mexicana* TKT with other transketolase protein sequences, a phylogenetic tree was created. The sequences used in this study are all detailed in Table 4.6 (Chapter IV). One additional sequence that was added to this phylogenetic analysis was a cyanobacterial transketolase that was obtained from the *Synechocystis* species database, CyanoBase. This was obtained by searching the database for sequence similarity to the *L. mexicana* transketolase, as it was not obtained through a BLAST search of public databases.

The neighbour-joining method (Saitou and Nei, 1987) was used to produce a cladogram. The transketolase protein sequences cluster according to their individual phyla (Figure 5.7). The phylogenetic analysis suggests that the *L. mexicana* and *Plasmodium* species transketolase protein sequences are more closely related to the plant, cyanobacteria and yeast transketolase sequences compared to the transketolase sequences from other organisms. A direct comparison of various transketolases with the *L. mexicana* transketolase using Clustal W shows a 50.6% identity with yeast and 48.3% identity with *Synechocyctis* (cyanobacteria) protein sequences (Table 5.2). The gram-negative bacterial transketolases branch together with further clustering into the α-, β- and γ-subdivisions. The mammalian protein sequences cluster with various bacterial transketolase sequences suggesting that the sequences of the former two

groups and *Leishmania* are divergent. The branch length indicates the distance between protein sequences.

A second phylogenetic cladogram was made using the transketolase protein sequences from plants, yeasts, protozoans and cyanobacteria (Figure 5.8). The tree shows bootstrap values, indicating the confidence values of the tree from 1000 random samples. This tree suggests that the *L. mexicana* transketolase clusters specifically with the plants and cyanobacteria. Several other proteins from the pentose phosphate pathway and glycolysis have been shown to branch closely with their respective plant and cyanobacteria homologues (Hannaert *et al.*, unpublished). *Nicotiana tabacum* (Tob), *Solanum tuberosum* (Sol) and *Spinacia oleracea* (Spinach) cluster within the plant sequences. This is not surprising, as the transketolase proteins from these species of plants have been localised to the chloroplast. *Craterostigma plantigineum* transketolase 7 and 10 are cytosolic proteins. The *L. mexicana* transketolase does not cluster specifically to either the chloroplast or cytosolic protein in plants. The phylogenetic tree suggests that the *Plasmodium* transketolase sequences are more closely related to the yeast sequences.

Organism	% Identity
Spinacia oleracea	42.3
Craterostigma plantagineum 7	46.8
Craterostigma plantagineum 10	45.3
Nicotiana tabacum	42.6
Solanum tuberosum	44.1
Plasmodium falciparum	40.8
Plasmodium yoelli	41.0
Schizosaccharomyces pombe	46.4
Pichia stipitis	46.1
Saccharomyces cerevisiae 1	50.6
Saccharomyces cerevisiae 2	45.9
Kluyveromyces lactis	47.9
Candida albicans	47.7
Synechocystis	48.3

Table 5.2. Percentage identity of transketolase from other organisms in comparison to the *L. mexicana* transketolase. The Clustal W algorithm was used to determine the identities between the sequences.

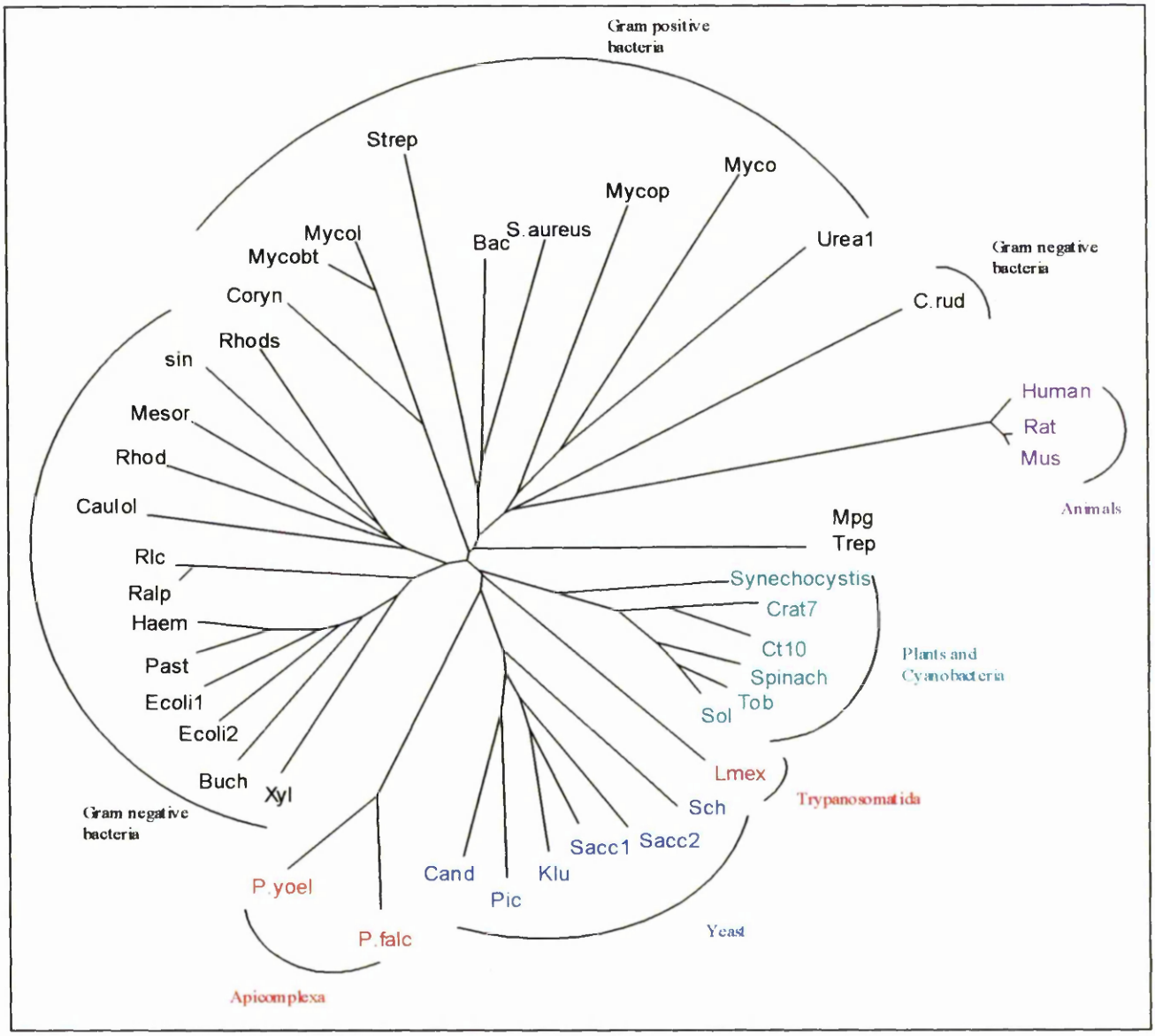


Figure 5.7: A phylogenetic cladogram representing the protein sequence comparison and the clustering relationship of various transketolases. The abbreviated names of the organisms are fully explained in Table 4.6. The *Synechocystis* species (strain PCC6803) transketolase protein was obtained from CyanoBase (accession number sll1070). The plant transketolase sequences are represented in green text, the yeast in blue, the protozoa in red and the mammalian sequences in purple. Clustal X (1.5b) was used to create this tree with a neighbour joining method, and was viewed using Tree View (Win32) version 1.6.5. The bar in the left bottom corner of this figure represents the observed pairwise distances calculated between sequences.

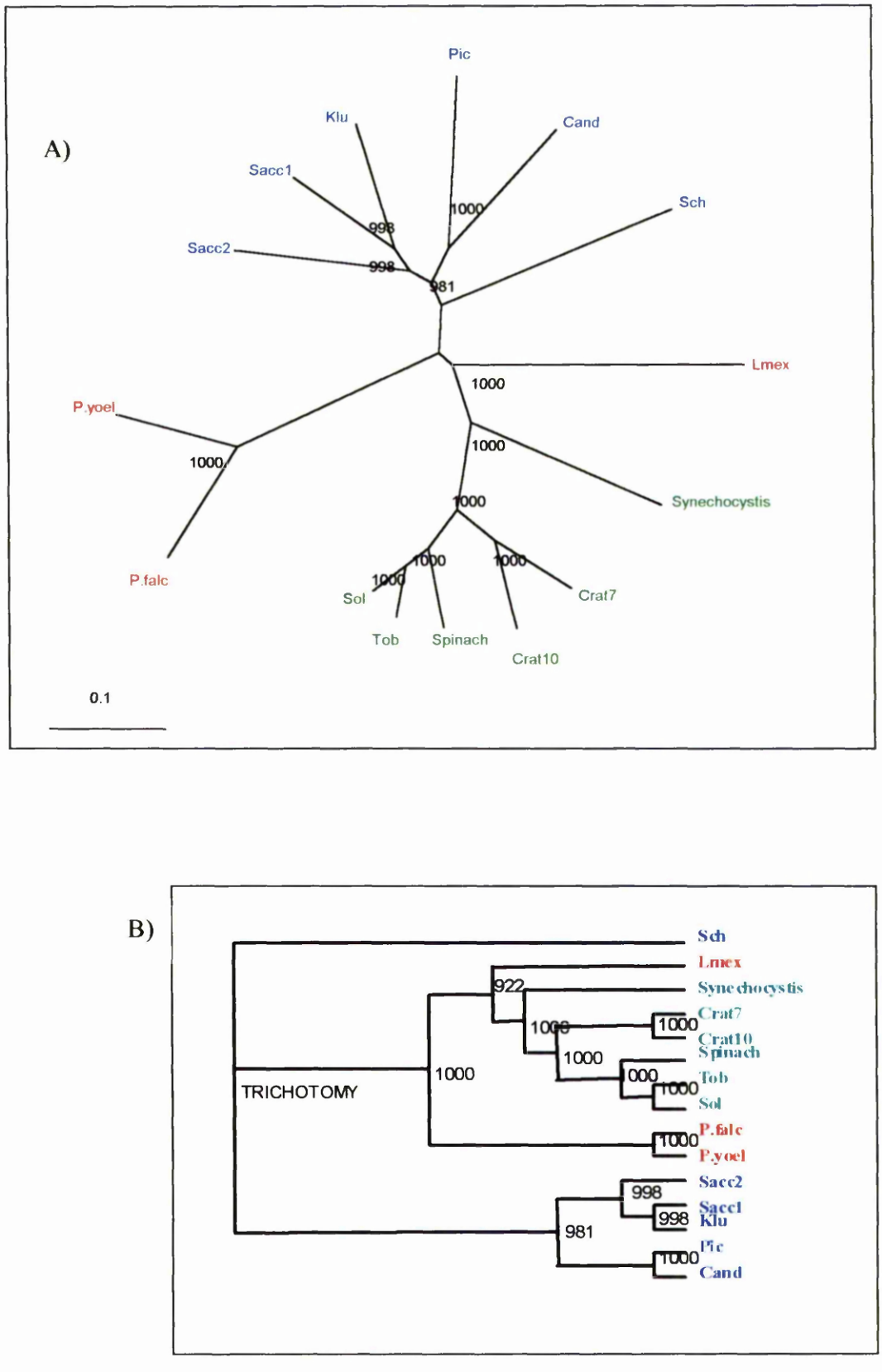


Figure 5.8: Phylogenetic tree for the plant, yeast, protozoa and cyanobacteria transketolase protein sequences. Details of abbreviations and colour of text in Figure 5.7. A) represents a bootstrapped cladogram, B) represents a rectangular cladogram.

5.7. 1-deoxy-D-xylulose 5-phosphate synthase, a novel family of transketolases

1-deoxy-D-xylulose 5-phosphate synthase (DXPS) represents a novel family of highly conserved transketolases. Isoprenoids can be additionally synthesised by a novel mevalonate-independent pathway in plants, protozoa and bacteria (Lois *et al.*, 1998; Lange *et al.*, 1998; Jomaa *et al.*, 1999). The first step in this pathway is the synthesis of 1-deoxy-D-xylulose 5-phosphate (DOXP) from pyruvate and glyceraldehyde 3-phosphate. Recently several DXPS protein sequences from various organisms have become available from the cloning of the genes and sequence similarity searching of genome sequencing databases. To assess the relatedness of DXPS proteins to transketolase more generally, a phylogenetic tree was created using a neighbour-joining method (Figure 5.9). The DXPS sequences used to create this phylogenetic tree are listed in Table 5.3 (transketolase sequences use the abbreviations from Table 4.6). A formaldehyde transketolase, dihydroxyacetone synthase (DHAS), was isolated from the peroxisome of the yeast *Pichia angusta* (Janowicz *et al.*, 1985) and was also placed in the alignment. DHAS utilises formaldehyde as an acceptor substrate producing glyceraldehyde 3-phosphate and glycerone. A DXPS protein or a DHAS protein in *Leishmania* species has not been identified and could not be found in the *Leishmania major* sequencing database (December 2001).

The phylogenetic tree (Figure 5.9) indicates that the DXPS protein sequences are distinct from the transketolase protein sequences as they form two independent clusters clearly separated. Lange (Lange *et al.*, 1998) also noted this novel class of transketolases that branch independently to the transketolase proteins when the sequence was analysed. This phylogenetic tree suggests the DHAS protein of *Pichia angusta* is more related to transketolases than to DXPS proteins. The overall reaction of

the formaldehyde transketolase is more similar to the transketolase reaction in comparison to the DXPS reaction, possibly explaining why the sequences have a more recent common ancestry.

Phyla	Name of organism and protein	Shortened name	Length of sequence (AA)	Genbank accession number
Plant	Peppermint (*Menthal x piperita*), DXPS	Mint DXPS	724	AF019383
Yeast	*Pichia angusta*, DHAS	Pic DHAS	710	P06834
Protozoa	*Plasmodium falciparum*, DXPS	Plasm DXPS	1206	AF111814
Bacteria	*Escherichia coli*, DXPS	Ecoli DXPS	620	AF035440
	Pseudomonas aeruginosa, DXPS	Pseudo DXPS	627	AF282878
	Rhodobacter capsulatus, DXPS-A	Rhodo DXPS	641	Z11165
	Neisseria meningitidid serogroup A, DXPS	Neis DXPS	637	Q9JW13 (P)
	Mycobacterium tuberculosis, DXPS	Mycob DXPS	638	O07184 (P)
	Bacillus subtilis, DXPS	Bac DXPS	633	Z99112
	Synechococcus leopoliensis, DXPS	Syncoc DXPS	636	Q9R6S7
	Synechocystis sp. PCC 6803, DXPS	Syncys DXPS	640	P73067 (P)
	Deinococcus radiodurans, DXPS	Dein DXPS	629	Q9RUB5 (P)
	Aquifex aelicus, DXPS	Aqui DXPS	628	O67036 (P)
	Thermotoga martima, DXPS	Thermo DXPS	608	Q9X291 (P)

Table 5.3: A list of the DXPS and DHAS sequences used in this study to create the phylogenetic tree represented in Figure 5.8. The sequences are split into phyla, with the full name of the organism followed by the protein 1-deoxy-D-xylulose 5-phosphate synthase (DHXS) or dihydroxy-acetone synthase (DHAS). The shortened name is used in Figure 5.9. The length of the protein sequences are represented in amino acids (AA) and the genbank accession numbers represent putative protein sequences when followed by (P).

Plastid derived isoprenoids are thought to originate from this non-mevalonate pathway. The peppermint (*Menthal x piperita*) DXPS has an N-terminal plastid targeting signal suggesting a microbody localisation (Lange *et al.*, 1998). The *Plasmodium falciparum* DXPS contains an extended N-terminal sequence and is thought to be localised to the to apicoplast (Jomaa *et al.*, 1999). The *P. falciparum* transketolase sequence was obtained from the *Plasmodium* genome sequencing project (PlasmoDB, accession number PF00456). Analysis of the *P. falciparum* transketolase protein sequence for targeting signals (PSORT) suggested that it contains no apicoplast-

targeting signal and will therefore be cytosolic. No DXPS has been identified in *Leishmania* species and the sequence is currently not present in the *L. major* genome sequencing database.

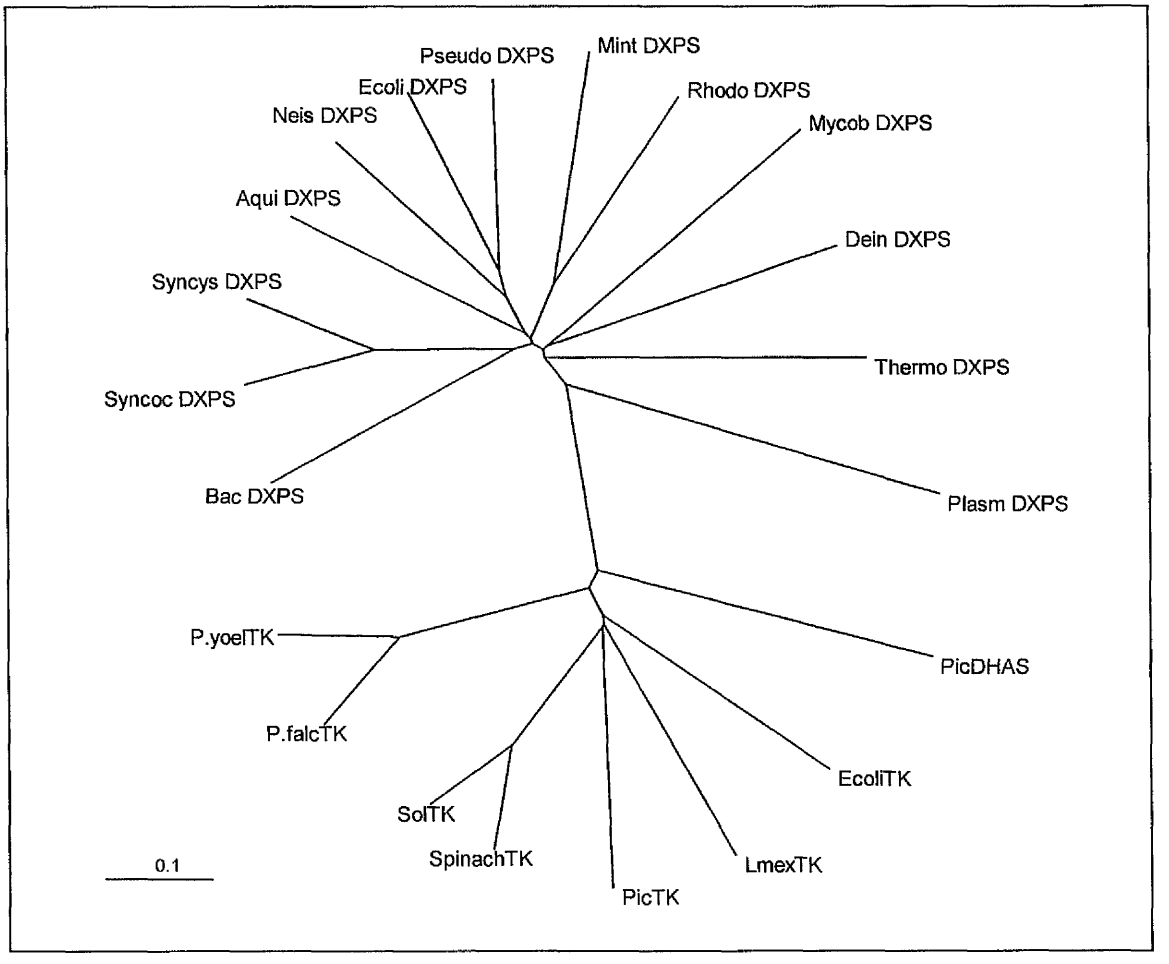


Figure 5.9: A phylogenetic cladogram representing a protein sequence comparison and the clustering relationship of transketolase with DXPS and a DHAS protein sequence. Abbreviations are listed in Table 5.2 for DXPS and DHAS sequences and Table 4.6 for transketolase sequences. Clustal X (1.5b) was used to create this tree with a neighbour joining method, and was viewed using Tree View (Win32) version 1.6.5. The bar in the left bottom corner of this figure represents the observed pairwise distances calculated between sequences.

5.8. Discussion

Transketolase was detected in both promastigote and axenic amastigote crude-lysates of *L. mexicana*. The transketolase protein was not present in the amastigotes from lesions, and therefore differences must exist between the axenic amastigotes and amastigotes from lesions. This is expected, as the axenic amastigotes are grown in medium whereas the amastigotes from lesions are intracellular and inhabit a very different environment. However, in this experiment, the control protein did not show any cross-reaction with this preparation of amastigotes either, so the presence of transketolase in amastigotes from lesions cannot be ruled out. The preparation was stored at –70 °C for a number of months, which may have had an effect on the preparation.

The transketolase protein was shown to be mainly cytosolic when fractionated with 8 % being localised to the granular fraction. Approximately 3 % of the total activity of transketolase in *L. mexicana* promastigotes was localised to granular material with similar characteristics of the glycosome in the sucrose gradient. However these glycosomal fractions were contaminated with cytosolic material, most likely present on the surface of the glycosomes. Therefore, none of the transketolase activity seen can be strictly related to the glycosome, and it can only be concluded that no transketolase is present in the glycosome in this experiment. Preliminary digitonin subcellular fractionation experiments of all the enzymes in the pentose phosphate pathway in *L. mexicana* indicate that all are present mainly in the cytosol, but all have a minor particulate compartmentation, including transketolase (Cazzulo, poster, Crete, May 2000).

A C-terminal PTS with the residues SKM is present in both the *L. mexicana* transketolase and a putative *L. major* transketolase. A putative *T. brucei* transketolase fragment identified by homology searching a genome-sequencing database shows a C-terminal PTS with the residues SHL. The putative *T. cruzi* transketolase sequence obtained did not have homology to the C-terminal of the protein, and therefore no PTS could be determined.

Phlyogenetic analysis of transketolase protein sequences suggested that the *L. mexicana* transketolase protein sequence clusters closely with plant and cyanobacteria transketolase protein sequences. Several genes encoding proteins in the Calvin cycle in the chloroplast in plants, including transketolase, are thought to be of cyanobacterial origin (Martin and Schnarrenberger, 1997). Plant-like proteins have been noted to be present in trypanosomatids, for example sedoheptulose-1, 7-bisphosphate and fructose-1, 6-bisphosphate, normally present only together in the chloroplast of plants. Hannaert *et al.*, unpublished, have suggested that the Kinetoplastida acquired an algal endosymbiont that was subsequently lost, transferring some of the algal genes to the host's genome. Trypanosomatids are from the order Kinetoplastida, and share the phylum Euglenozoa with members of the order Euglenida with the distinction being that Euglenids possess chloroplasts, whereas the Kinetoplastids do not. It is thought that these two groups of organisms split in lineage, with only the Euglenids possessing a plastid, however, phylogenetic analysis suggests a common ancestry of Kinetoplastids, plants and cyanobacteria. This may explain why several of the enzymes involved in carbohydrate metabolism in trypanosomatids cluster with plant and cyanobacteria homologues when subjected to phylogenetic analysis (Hannaert *et al.*, unpublished; Krepinsky *et al.*, 2001), including the *L. mexicana* transketolase protein. In trypanosomatids, the glycolytic pathway occurs mainly in the glycosome, with the

pentose phosphate pathway possibly also being partially localised to this specialised organelle. A recent influx of information regarding the ancestry of the enzymes involved in these pathways suggests a plant or cyanobacterial origin. It may be possible that the Kinetoplastida lineage did acquire an algal endosymbiont that was subsequently lost after a transfer of genes occurred into the nuclear genome, and that these proteins were localised to another organelle, now known as the glycosome. The PPP may be partially glycosomal, with a role in both the cytosol and the glycosome. So far, only the enzymes from the oxidative pathway have been localised to the glycosome, suggesting that this part of the pathway may be providing NADPH to this organelle.

Phylogenetic analysis of the DXPS protein sequences suggests they are a novel family of transketolases with distinctive branching in a phylogenetic tree. A DXPS protein has not been located in trypanosomatids as yet. The DHAS sequence of the yeast *Pichia angusta* is more related to the transketolase protein sequences than to the DXPS sequences.

Chapter VI

General Discussion

The main aim of this study was to clone and characterise a transketolase from the parasite *Leishmania mexicana*. This study is part of an ongoing collaborative effort between a number of laboratories to characterise the pentose phosphate pathway in *Leishmania* species. The transketolase gene was successfully cloned and sequenced during the course of this study. A recombinant transketolase was expressed, purified and characterised in order to assess the characteristics of the protein. Subcellular localisation studies were performed to determine the location of the transketolase within the *L. mexicana* cell. Experiments were initiated to knock-out the transketolase in *L. mexicana* promastigotes to assess the importance of transketolase in the parasite.

A transketolase gene was cloned from *L. mexicana* using a variety of molecular methodologies. The gene was amplified using several PCR-based methods and an ORF was obtained for the transketolase gene. The translated protein contained features specifically characteristic of transketolase proteins. A TPP-box, known to be conserved in all TPP-dependent enzymes (Hawkins *et al.*, 1989), and crucial for the binding of TPP to transketolase and activation of the enzyme (Wang, J.J.L. *et al.*, 1997; Meshalkina *et al.*, 1997) is present in the sequence between amino acid positions 144 to 183. A TKT-box, specific to transketolase and known to form part of the active site, is present on the *L. mexicana* TKT sequence between amino acid residues 465 to 493. The conserved residues that form the TKT-box are directly involved in the co-factor and donor (aldehyde) substrate binding (Nikkola *et al.*, 1994; Nilsson *et al.*, 1997). The *L. mexicana* transketolase also contains a peroxisome targeting signal (PTS) on the

carboxy-terminal of the protein. This suggests that the protein may be targeted to the glycosome, a unique organelle present in all trypanosomatids that is primarily involved in glycolysis.

The *L. mexicana* TKT protein was aligned with transketolase proteins from a variety of organisms. In addition to published sequences, several transketolases were identified from other organisms by searching DNA and protein sequences, which are available in genome sequencing databases, for homology to transketolase. Transketolase genes were identified from *Plasmodium falciparum* and *Plasmodium yoelli* that have not been previously reported. A cyanobacterial transketolase was obtained from the fully sequenced genome of *Synechocystis*. The *L. mexicana* TKT is most similar to the *S. cerevisiae* (yeast) transketolase-1 and the *Synechocystis* (cyanobacteria) transketolase, with 50.6 % and 48.3 % identities, respectively. All of the conserved residues from the yeast transketolase that have been previously shown to be crucial for substrate and cofactor binding, and overall enzyme function, are present in the *L. mexicana* TKT protein sequence.

A Southern blot of *L. mexicana* genomic DNA probed with transketolase labelled with ^{32}P indicates that there is only one isoform of transketolase present in the *Leishmania* genome. Only one transketolase has been found in the *L. major* database and is present on chromosome 24. Transketolase has been found to be present in certain genomes in more than one copy, with *S. cerevisiae* and *E. coli* both having two copies and *Craterostigma plantagineum* having three copies. However, in several organisms including humans, spinach, tobacco and *Plasmodium* only one copy of the gene has been found.

During this study, the flanking regions upstream and downstream of the transketolase gene were amplified. In total, 700 bp of 5' flank and 900 bp of 3' flank

were amplified from genomic DNA. No other gene sequence was detected in the flanking regions, suggesting that the transketolase coding region is surrounded by approximately 1kb of intergenic region on either flank. This density of protein-coding gene sequences is constant throughout the *L. major* sequenced genome to date with one gene to every 3.7 kb (Myler *et al.*, 2001). The 5'-flanking region contains the spliced leader acceptor site, AG, 145 bp upstream to the transketolase start codon. This was shown to be the site where the oligonucleotide to the 39-nucleotide spliced leader binds to the cDNA of transketolase to cleave the mRNA transcript. Polypyrimidine tracts (>90 % T/C rich) occur upstream of the AG spliced acceptor site. These tracts have been implicated in the accurate polyadenylation of adjacent upstream mRNA transcripts and spliced site recognition (LeBowitz *et al.*, 1993; Schürch *et al.*, 1994; Hug *et al.*, 1994). The flanking regions were used to make a construct to replace the transketolase gene *in vivo* with an antibiotic resistance marker gene by homologous recombination. Two constructs were made, one with a hygromycin B phosphotransferase gene and the other with a streptothricin acetyltransferase gene, encoding resistance to the antibiotics Hygromycin B and Nourseothricin, respectively. Transfection of the constructs into the *L. mexicana* promastgotes did yield resistant clones of *Leishmania*, however the constructs did not recombine at the transketolase locus according to PCR. This may have been due to the short 5' flanking region (120 bp) used to make the constructs. This investigation is ongoing within the laboratory.

The transketolase gene from *L. mexicana* was cloned into a pET-16b vector and expressed in *E. coli*. The recombinant transketolase was abundantly expressed using this system, with up to 15 mg being obtained from one litre of *E. coli* culture. Expression of the protein was induced with IPTG, and only at an induction temperature of 15 °C was the protein partially soluble. At 25 °C, all of the expressed protein was in

the insoluble phase of the cell lysate. Under standard conditions, expression of a recombinant protein using this system usually occurs at 37 °C (for example, the recombinant *E. coli* transketolase expressed in this study). The unusual expression of the recombinant *L. mexicana* could be due to aggregation of the protein as it is being abundantly overexpressed. A reduction of temperature from 25 °C to 15 °C caused a shift in the protein expression from insoluble to soluble protein. It is possible that at 15 °C the rate of protein expression is reduced, producing more favourable conditions for correct folding of the protein in the soluble phase.

The recombinant *L. mexicana* transketolase had a highest measured specific activity of 1.65 U/mg of protein. This value is low in comparison to other recombinant transketolases from other organisms (for example, *E. coli* has a specific activity of 50 U/mg). This could be due to the expression system, with some aggregation in the final preparation leading to protein that is not correctly folded and therefore a partially inactive preparation. However, the biochemical characteristics are similar to other transketolases. The *L. mexicana* TKT K_m for R5P is 2.75 mM, which is similar to the *E. coli* K_m for R5P of 1.4 mM. Substrate inhibition did occur with R5P concentrations exceeding 50 mM when the donor substrate X5P was used at a constant saturating concentration. Transketolase is thought to have one active site where the donor and acceptor substrates bind sequentially. The donor substrate initially binds to the cofactor TPP and then releases the first product of the reaction, allowing the acceptor substrate to enter the active site cleft and bind to the 2-carbon glycoaldehyde unit. The final product is then released. Competition for this site occurs if there is excess substrate, leading to enzyme inhibition (Solov'eva *et al.*, 2000). This proposal is consistent with the Bi-Bi Ping-Pong mechanism of action, with the initial substrate being released prior to the

binding of the second substrate (Cavalier and Sable, 1972; Gyamerah and Willetts, 1997; Fiedler *et al.*, 2001).

A further biochemical characterisation of the *L. mexicana* TKT was hindered by the lack of a commercial supply of X5P. X5P was made in the laboratory using a biotransformation that utilised the recombinant *E. coli* transketolase. The biotransformation was difficult to repeat, and time consuming, however enough X5P was made in one preparation to do a limited kinetic study on the *L. mexicana* transketolase.

Currently, the *L. mexicana* recombinant TKT is undergoing crystallisation trials by Prof. Ylva Lindqvist in Sweden and has been crystallised to a resolution of 3 Å. This study may lead to the identification of differences between the *L. mexicana* and mammalian transketolase that could be exploited as chemotherapeutic targets assuming the enzyme is essential.

The activity of the native transketolase from promastigote *L. mexicana* was determined and was shown to be similar to that previously recorded in *Leishmania* species and *Trypanosoma brucei*. Anti-transketolase antibody reacted with proteins from *L. mexicana* promastigotes and *L. mexicana* axenic amastigotes, however it did not recognise protein from amastigotes from a lesion from BALB/C mice. This may reflect biochemical differences between axenic amastigotes and amastigotes from lesions. If for example the antibody failed to recognise the transketolase in amastigotes, as a different form may be present. Multiple forms of transketolase have been shown to exist. Yeast has two genes that code for transketolase, however three forms of transketolase have been purified from this organism (Kuimov, 1990). The third form of yeast transketolase is most likely a post-translationally modified protein, and a transketolase-RNA has been shown to exist in yeast (Tikhomirova *et al.*, 1990).

Another possibility is that transketolase is not present in the amastigotes form of *Leishmania*. Transketolase could not be found in the bloodstream form of *T. brucei* and a novel mechanism of 2-carbon transfer was suggested to exist (Cronin *et al.*, 1989). This may also be the case in the amastigote form of *Leishmania*. Developmental regulation of genes is known to occur in trypanosomatids. The glycolytic enzymes phosphoglyceratekinase-C and fructose bisphosphate aldolase are both upregulated in the bloodstream form *T. brucei* (Clayton, 1985; Graham, 1995). Perhaps the main role of the PPP in *L. mexicana* amastigotes is the production of NADPH and maintenance of reduced trypanothione to protect against reactive oxygen species. The oxidative pathway may be upregulated in this case, with transketolase and the non-oxidative pathway enzymes having a lesser role. However, as the control protein did not cross-react with the amastigotes proteins, no conclusion can be drawn as to the presence of transketolase in amastigotes from lesions.

The *L. mexicana* transketolase contained a C-terminal PTS with the residues SKM. This is suggestive of a glycosomal localisation of the protein, as this signal allows the protein to be imported into the glycosome (Sommer *et al.*, 1992). This specific PTS is also present on the glyceraldehyde-3-phosphate dehydrogenase protein from *T. brucei* that has been localised to the glycosome (Hannaert *et al.*, 1992). Searching the trypanosomatid databases for nucleotide homology to the *Leishmania* transketolase yielded DNA fragments with similarity in *L. major*, *T. brucei* and *T. cruzi*. The DNA sequence was translated into protein sequence and aligned with the *L. mexicana* transketolase, indicating that *L. major* and *T. brucei* also contain a PTS with the residues SKM and SHL, respectively. The *T. cruzi* sequence was not complete at the 3'-region and therefore the presence of a PTS could not be determined. The PPP enzyme 6-phosphogluconolactonase (6PGL) of *T. brucei* does contain the PTS C-

terminal tripeptide AKF and has been shown to have a dual localisation, with 10 % of the activity associated with the glycosome (Duffieux *et al.*, 2000). The *T. brucei* Glucose-6-phosphate dehydrogenase (G6PDH), the first enzyme in the PPP, is also partially present in the glycocome (~ 45 %), however does not contain a PTS-1 or a PTS-2 (Heise and Opperdoes, 1999; Duffieux *et al.*, 2000). It may be possible that the G6PDH enters the glycosome as a complex with 6PGL, or it may have an alternative PTS within the sequence. Subcellular localisation studies indicated that transketolase in promastigote *L. mexicana* was predominantly cytosolic. However, the localisation of transketolase in other life cycle stages such as metacyclic or amastigote *Leishmania* cannot be ruled out. Studies to determine the localsation of all of the PPP enzymes from *L. mexicana* are currently ongoing (Prof. Cazzulo, Buenos Aires). The PPP could potentially have a role in the glycosome, as NADPH is required for enzymatic reactions of the ether-lipid and sterol biosynthesis pathways that are partially present in the glycosome (Michels *et al.*, 2000). R5P may also play a role in the glycosome, as this phosphorylated sugar is incorporated into nucleotides via phosphoribosyl pyrophosphate, and enzymes of the purine salvage and pyrimidine biosynthesis pathways have been localised to the glycosome (Michels *et al.*, 2000). Transketolase has been found in other microbodies within other organisms such as the endoplasmic reticulum (ER) in human cells and the chloroplast in plants (Bublitz and Steavenson, 1988; Teige *et al.*, 1998). It has also been suggested that in *T. cruzi*, transketolase may be present in the ER (Prof. Cazzulo, Buenos Aires, personnel communication).

Sequence comparison and phylogenetic analysis of the *L. mexicana* transketolase revealed that it is closely related to cyanobacterial, plant and yeast transketolases. The mammalian and bacterial transketolase were phylogenetically distinct to the other transketolase, clustering independently in a phylogenetic tree. The

Plasmodium transketolases also clustered with the *Leishmania* and the plant sequences. *Plasmodium* species do contain an apicoplast, an organelle thought to have been acquired by endosymbiosis and be reminiscent of the chloroplast (Andersson and Roger, 2002). Several enzymes in carbohydrate metabolism in trypanosomatids cluster with plant and cyanobacteria homologues when subjected to phylogenetic analysis (Hannaert *et al.*, unpublished; Krepinsky *et al.*, 2001). *T. brucei* 6PGDH and G6PDH both branch with plants under phylogenetic analysis, with 6PGD specifically clustering with cyanobacteria (Krepinsky *et al.*, 2001). It has been suggested that there is a common ancestry of the genes in Kinetoplastids and higher plants with a cyanobacterial origin. It has been postulated that an organism ancestral to Euglenida and Kinetoplastida acquired a photosynthetic algal endosymbiont that was subsequently lost in the Kinetoplastid lineage, after a transfer of genes occurred into the nuclear genome (Hannaert *et al.*, unpublished). Cyanobacteria are thought to be closely related to the chloroplasts in plants and algae (Martin and Schnarrenberger, 1997), and therefore this theory of a loss of a photosynthetic endosymbiont explains why some of the enzymes of carbohydrate metabolism of trypanosomatids are similar to those in cyanobacteria and plants. The presence of the PTS on several enzymes of carbohydrate metabolism must have arisen subsequent to the acquisition of the glycosome. The glycosome is a specialised peroxisome and may have occurred with the entry of an endosymbiont and subsequent gene transfer to the nucleus, or a specialisation of a pre-existing compartment (Borst, 1989). Glycosomes are not present in Euglenids, suggesting that these organelles must have been acquired in an early ancestral Kinetoplastid (Michels *et al.*, 2000) or may have been lost from the Euglenida lineage. An evolutionary advantage must have existed in order for a set of cytosolic enzymes to become targeted to one microbody. It is likely that the compartmentation of glycolysis occurred as it

increased the efficiency of the breakdown of glucose and therefore an increase in energy availability.

The phylogenetic analysis of the formaldehyde transketolase dihydroxyacetone synthase (DHAS) and the novel transketolase family, 1-deoxy-D-xylulose 5-phosphate synthase (DXPS), show that they are evolutionarily distinct. Two groups clearly form between DXPS and transketolases, with DHAS being more closely related to the transketolases. DXPS catalyses the production of 1-deoxy-D-xylulose 5-phosphate (DOXP) from pyruvate and glyceraldehyde 3-phosphate, that represents the first step in the non-mevalonate pathway. The mevalonate-independent pathway may represent a good drug-target, as it is not present in mammalian cells. DOXP reductoisomerase is the second enzyme in this pathway, the final product being isopentenyl diphosphate, an isoprenoid precursor The recombinant *P. falciparum* DOXP reductoisomerase was inhibited by fosmidomycin, a compound that causes antimalarial activity in multi-drug resistant strains of *Plasmodium falciparum*. To date, no DXPS or DOXP reductoisomerase has been identified in trypanosomatids.

Transketolase may be interesting as a chemotherapeutic target as it is a key enzyme in the PPP, specifically in the production of R5P and E4P, two important molecules for nucleotide synthesis and aromatic amino acid biosynthesis. Transketolase inhibitors are being investigated as anti-cancer agents, as it is thought that transketolase produces R5P via the PPP specifically for the proliferation of the cancer cells. Treatment of adenocarcinoma cells with the transketolase inhibitor oxythiamine resulted in inhibition of cellular proliferation (Boros *et al.*, 1997). *E. coli* mutants lacking both of the genes encoding for transketolase are auxotrophic for aromatic amino acids, pyridoxine (vitamin B_6) and vitamins (Zhao and Winkler, 1994). Similar findings were reported for a yeast strain lacking one transketolase gene, with the organisms requiring a

supplement of aromatic amino acids (Sundström *et al.*, 1993). Using an RNA interference (RNAi) approach, it was shown that a decrease in tobacco transketolase can lead to decreased levels of aromatic amino acids, in turn inhibiting phenylpropanoid metabolism and also effects photosynthesis by causing a loss of chlorophyll (Henkes, *et al.*, 2001). These observations suggest that transketolasė is important for a variety of biochemical pathways in various organisms and may therefore be a potential drug target.

The pentose phosphate pathway as a whole may represent a potential chemotherapeutic target as it supplies important substrates such as NADPH to the cell. It has been shown to be important for the protection of yeast against oxidative stress, as mutants that lacked genes encoding for certain proteins in the PPP were more sensitive to hydrogen peroxide (Juhnke *et al.*, 1996). Protection against oxidative stress is crucial for the successful colonisation of the *Leishmania* parasite within the mammalian host. The PPP may play a crucial role in supplying NADPH to the cell, and therefore may be a potential drug target.

To determine if the *L. mexicana* transketolase is a potential drug target, the recombinant protein should be fully characterised, with inhibition studies using various transketolase inhibitors previously described, such as oxythiamine (Boros *et al.*, 1997). The gene replacement studies initiated during this project are to be continued, and will play an important role in determining if transketolase is essential to the parasite, and what biochemical pathways are effected by the gene deletion. Transketolase may play a role in *Leishmania* in providing R5P to the cell, for the incorporation into nucleic acids and therefore be necessary for cell proliferation. Aromatic amino acids are scavenged from the environment and therefore E4P may not play such a crucial role in *Leishmania* as it does in *E. coli* and yeast.

In conclusion, I have isolated and cloned a transketolase from the parasite *L. mexicana*. The production of the recombinant transketolase and the characterisation of this enzyme showed that it has similar biochemical characteristics to transketolases from other organisms. The identification of the PTS on the *L. mexicana* TKT protein led to the subcellular localisation of the enzyme that showed a cytosolic location for this enzyme. Transketolase may be a potential chemotherapeutic target for parasites, and experiments assessing the importance of this enzyme have been initiated.

Bibliography

Agabian, N. (1990). Trans Splicing of Nuclear Pre-mRNAs. Cell *61*, 1157-1160.

Agami, R. and Shapira, M. (1992). Nucleotide Sequence of the Spliced Leader RNA Gene from *Leishmania mexicana amazonensis*. Nucleic Acids Research *20*, 1804.

Altincicek, B., Hintz, M., Sanderbrand, S., Wiesner, J., Beck, E. and Jomaa, H. (2000). Tools for Discovery of Inhibitors of the 1-Deoxy-D-Xylulose 5-Phosphate (DXP) Synthase and DXP Reductoisomerase: An Approach with Enzymes from the Pathogenic Bacterium *Pseudomonas aeruginosa*. FEMS Microbiology Letters *190*, 329-333.

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. Nucleic Acids Research *25*, 3389-3402.

Andersson, J.O. and Roger, A.J. (2002). A Cyanobacterial Gene in Nonphotosynthetic Protists- An Early Chloroplast Acquisition in Eukaryotes? Current Biology *12*, 115-119.

Anon. Report of a WHO Expert Committee. 701. (1984). Geneva, WHO. WHO Technical Report series.

Antoine, J., Prina, E., Lang, T. and Courret, N. (1998). The Biogenesis and Properties of the Parasitophorous Vacuoles that Harbour *Leishmania* in Murine Macrophages. Trends in Microbiology *7*, 392-401.

Antonenkov, V.D. (1989). Dehydrogenases of the Pentose Phosphate Pathway in Rat Liver Peroxisomes. European Journal of Biochemistry *183*, 75-82.

Arana, B., Rizzo, N. and Diaz, A. (2001). Chemotherapy of Cutaneous Leishmaniasis: A Review. Medical Microbiology and Immunology *190*, 93-95.

Bakker, B.M., Westerhoff, H.V., Opperdoes, F.R. and Michels, P.A.M. (2000). Metabolic Control Analysis of Glycolysis in Trypanosomes as an Approach to Improve Selectivity and Effectiveness of Drugs. Molecular and Biochemical Parasitology *106*, 1-10.

Barrett, M.P. and Le Page, W.F. (1993). A 6-Phosphogluconate Dehydrogenase Gene from *Trypanosoma brucei*. Molecular and Biochemical Parasitology *57*, 89-100.

Barrett, M.P., Phillips, C., Adams, M.J. and Le Page, W.F. (1994). Overexpression in *Escherichia coli* and Purification of the 6-Phosphogluconate Dehydrogenase of *Trypanosoma brucei*. Protein Expression and Purification *5*, 44-49.

Barrett, M.P. (1997). The Pentose Phosphate Pathway and Parasitic Protozoa. Parasitology Today *13*, 11-16.

Barrett, M.P., Mottram, J.C. and Coombs, G.H. (1999). Recent advances in Identifying and Validating Drug Targets in Trypanosomes and Leishmaniasis. Trends in Microbiology *7*, 82-88.

Bart, G., Frame, M.J., Carter, R., Coombs, G.H. and Mottram, J.C. (1997). Cathepsin B-like Cysteine Proteinase-Deficient Mutants of *Leishmania mexicana*. Molecular and Biochemical Parasitology *88*, 53-61.

Basu, T.K. and Dickerson, J.W.T. (1976). The Thiamin Status of Early Cancer Patients with Particular Reference to Those with Breast and Bronchial carcinomas. Oncology *33*, 250-252.

Berens, R.L., Brun, R. and Krassner, S.M. (1976). A Simple Monophasic Medium for Axenic Culture of Hemoflagellates. The Journal of Parasitology *62*, 360-365.

Berens, R.L., Deutsch-King, L.C. and Marr, J.J. (1980). *Leishmania donovani* and *Leishmania braziliensis*: Hexokinase, Glucose 6-Phosphate dehydrogenase and Pentose Phosphate Shunt Activity. Experimental Parasitology *49*, 1-8.

Bergmeyer, H.U. (1974). Pyruvate Kinase. Methods of Enzymatic Analysis *1*, 509-510.

Berman, J.D. (1988). Chemotherapy for Leishmaniasis: Biochemical Mechanisms, Clinical Efficacy, and Future Strategies. Reviews of Infectious Diseases *10*, 560-586.

Bernacchia, G., Schwall, G., Lottspeich, F., Salamini, F. and Bartels, D. (1995). The Transketolase Gene Family of the Resurrection Plant *Craterostigma plantagineum:* Differential Expression During the Rehydration Phase. The EMBO Journal *14*, 610-618.

Blass, J.P. and Gibson, G.E. (1977). Abnormality of a Thiamine-Requiring Enzyme in Patients with Wernicke-Korsakoff Syndrome. The New England Journal of Medicine *297*, 1367-1370.

Blattner, J., Swinkels, B., Dorsam, H., Prospero, T. and Subramani, S. (1992). Glycosome Assembly in Trypanosomes: Variations in the Acceptable Degeneracy of a COOH-terminal Microbody Targeting Signal. The Journal of Cell Biology *119*, 1129-1136.

Blattner, J., Dörsam, H. and Clayton, C.E. (1995). Function of N-terminal Import Signals in Trypansome Microbodies. FEBS Letters *360*, 310-314.

Blum, J.J. (1993). Intermediary Metabolism of *Leishmania*. Parasitology Today *9*, 118-122.

Booth, C.K. and Nixon, P.F. (1993). Reconstitution of Holotransketolase is by a Thiamin-diphosphate-Magnesium Complex. European Journal of Biochemistry *218*, 261-265.

Boros, L.G., Puigjaner, J., Cascante, M., Paul Lee, W., Brandes, J.L., Bassilian, S., Yusuf, F.I., Williams, R.D., Muscarella, P., Melvin, W.S. and Schirmer, W.J. (1997). Oxythiamine and Dehydroepiandrosterone Inhibit the Nonoxidative Synthesis of Ribose and Tumour Cell Proliferation. Cancer Research *57*, 4242-4248.

Boros, L.G., Lee, P.W.N., Brandes, J.L., Casacante, M., Muscarella, P., Schirmer, W.J., Melvin, W.S. and Ellison, E.C. (1998). Nonoxidative Pentose Phosphate Pathways and their Direct Role in Ribose Synthesis in Tumours: Is Cancer a Disease of Cellular Glucose Metabolism? Medical Hypothesis *50*, 55-59.

Boros, L.G. (2000). Population Thiamine Status and Varying Cancer Rates between Western, Asian and African Countries. Anticancer Research *20*, 2245-2248.

Borst, P. (1989). Peroxisome Biogenesis Revisited. Biochimica et Biophysica Acta *1008*, 1-13.

Bradford, M.M. (1976). A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilising the Principle of Protein-Dye Binding. Analytical Biochemistry 72, 248-254.

Brittingham, A. and Mosser, D.M. (1996). Exploitation of the Complement System by *Leishmania* Promastigotes. Parasitology Today *12*, 444-447.

Bublitz, C. and Steavenson, S. (1988). The Pentose Phosphate Pathway in the Endoplasmic Reticulum. The Journal of Biological Chemistry *263*, 12849-12853.

Burchmore, R.J. and Landfear, S.M. (1998). Differential Regulation of Multiple Glucose Transporter Genes in *Leishmania mexicana*. The Journal of Biological Chemistry *273*, 29118-29126.

Bykova, I.A., Solovjeva, O.N., Meshalikina, L., Kovina, M.V. and Kochetov, G. (2001). One-Substrate Transketolase-Catalysed Reaction. Biochemical and Biophysical Research Communications *280*, 845-847.

Calvier, R.L. and Sable, H.Z. (1972). Kinetic Study of Transketolase from Baker's yeast. Federation Procceedings *31*, 846.

Cazzulo, J.J. (1992). Aerobic Fermentation of Glucose by Trypanosomatids. The FASEB Journal *6*, 3153-3161.

Clarke, J.L., Scopes, D.A., Sodeinde, O. and Mason, P.J. (2001). Glucose-6-Phosphate Dehydrogenase-6-Phosphogluconolactonase. A Novel Bifunctional Enzyme in Malaria Parasites. European Journal of Biochemistry *268*, 2013-2019.

Clarke, T.B. and Wallace, F.G. (1960). A Comparative Study of Kinetoplast Ultrastructure in the Trypanosomatidae. Journal of Protozoology *7*, 115-124.

Clayton, C.E. (1985). Structure and Regulated Expression of Genes Encoding Fructose Bisphosphate Aldolase in *Trypanosoma brucei*. The EMBO Journal *4*, 2997-3003.

Coburn, C.M., Otteman, K.M., McNeely, T., Turco, S.J. and Beverley, S.M. (1991). Stable DNA Transfection of a Wide Range of Trypanosomatids. Molecular and Biochemical Parasitology *46*, 169-180.

Collard, F., Collet, J.F., Gerin, I., Veiga-da-Chunha, M. and Van Schaftingen, E. (1999). Identification of the cDNA Encoding Human 6-Phosphogluconolactonase, the Enzyme Catalysing the Second Step of the Pentose Phosphate Pathway. FEBS Letters *459*, 223-226.

Coombs, G.H., Craft, J.A. and Hart, D.T. (1982). A Comparative Study of *Leishmania mexicana* Amastigotes and Promastigotes. Enzyme Activities and Subcellular Locations. Molecular and Biochemical Parasitology *5*, 199-211.

Cronin, C.N., Nolan, D.P. and Voorheis, H.P. (1989). The Enzymes of the Classical Pentose Phosphate Pathway Display Differential Activities in Procyclic and Bloodstream Forms of *Trypanosoma brucei*. FEBS Letters *244*, 26-30.

Cruz, A., Coburn, C.M. and Beverley, S.M. (1991). Double Targeted Gene Replacement for Creating Null Mutants. Proceedings of the National Academy of Sciences of the United States of America *88*, 7170-7174.

Cunningham, M.L., Titus, R.G., Turco, S.J. and Beverley, S.M. (2001). Regulation of Differentiation to the Infective Stage of the Protozoan by Tetrahydobiopterin. Science *292*, 285-287.

Datta, A.G. and Racker, E. (1961). Mechanisms of Action of Transketolase. The Journal of Biological Chemistry *236*, 617-623.

De Walque, S., Opperdoes, F.R. and Michels, P.A.M. (1999). Cloning and Characterisation of *Leishmania mexicana* Fructose-1, 6-bisphosphate Aldolase. Molecular and Biochemical Parasitology *103*, 279-283.

Draths, K.M., Pompliano, D.L., Conley, D.L., Frost, J.W., Berry, A., Disbrow, G.L., Staversky, R.L. and Lievense, J.C. (1992). Biocatalytic Synthesis of Aromatics from D-Glucose: The Role of Transketolase. J. Am. Chem. Soc. *114*, 3956-3962.

Duffieux, F., Roy, J.V., Michels, P.A.M. and Opperdoes, F.R. (2000). Molecular Characterisation of the First Two Enzymes of the Pentose-Phosphate Pathway of *Trypanosoma brucei*. The Journal of Biological Chemistry *275*, 27559-27565.

Egan, R.M. and Sable, H.Z. (1981). Transketolase Kinetics. The Journal of Biological Chemistry *256*, 4877-4883.

Eisenreich, W., Sagner, S., Zenk, M. and Bacher, A. (1997). Monoterpenoid Essential Oils are Not of Mevalonoid Origin. Tetrahedron Letters *38*, 3889-3892.

Fairlamb, A.H. and Cerami, A. (1992). Metabolism and Functions of Trypanothione in the Kinetoplastida. Annual Review of Microbiology *46*, 695-729.

Fenech, F.F. (1997). Leishmaniasis in Malta and the Mediterranean Basin. Annals of Tropical Medicine and Parasitology *91*, 747-753.

Fiedler, E., Golbik, R., Schneider, G., Tittmann, K., Neef, H., König, S. and Hübner, G. (2001). Examination of Donor Substrate Conversion in Yeast Transketolase. The Journal of Biological Chemistry *276*, 16051-16058.

Fischer, C., Voss, A. and Engel, J. (2001). Development Status of Miltefosine as First Oral Drug in Visceral and Cutaneous Leishmaniasis. Medical Microbiology and Immunology *190*, 85-87.

Follstad, B.D. and Stephanopoulos, G. (1998). Effect of Reversible Reactions on Isotope Label Redistribution Analysis of the Pentose Phosphate Pathway. European Journal of Biochemistry *252*, 360-371.

Fraser, C.M. *et al.* (1998). Complete Genome Sequence of *Treponema pallidum*, the Syphilis Spirochete. Science *281*, 375-388.

Galvao-Quintao, L., Alfier, S.C., Ryter, A. and Rabinovitch, M. (1990). Intracellular Differentiation of *Leishmania amazonensis* Promastigotes to Amastigotes: Presence of Megasomes, Cysteine Proteinase Activity and Susceptibility to Leucine-Methyl Ester. Parasitology *101*, 7-13.

Ganem, B. (1978). From Glucose to Aromatics: Recent Developments in Natural Products of the Shikimic Acid Pathway. Tetrahedron *34*, 3353-3383.

Ghosh, D.K. and Datta, A.G. (1971). *Leishmania donovani*: Assay for a Functional Pentose Phosphate Pathway. Experimental Parasitology *29*, 103-109.

Gould, S.J., Keller, G.A., Hosken, N., Wilkinson, J. and Subramani, S. (1989). A Conserved Tripeptide Sorts Proteins to Peroxisomes. The Journal of Cell Biology *108*, 1657-1664.

Graham, S.V. (1995). Mechanisms of Stage-Regulated Gene Expression in Kinetoplastida. Parasitology Today *11*, 217-223.

Gutteridge, W.E. (1997). Designer Drugs: Pipe-Dreams or Realities? Parasitology *114*, S145-S151.

Gyamerah, M. and Willetts, A.J. (1997). Kinetics of Overexpressed Transketolase from *Escherichia coli* JM109/pQR 700. Enzyme and Microbial Technology *20*, 127-134.

Hahn, F.M., Eubanks, L.M., Testa, C.A., Blagg, B.S.J., Baker, J.A. and Poulter, C.D. (2001). 1-Deoxy-D-Xylulose 5-Phosphate Synthase, the Gene Product of Open Reading Frame (ORF) 2816 and ORF 2895 in *Rhodobacter capsulatus*. Journal of Bacteriology *183*, 1-11.

Hammond, D.J., Gutteridge, W.E. and Opperdoes, F.R. (1981). A Novel Location for Two Enzymes of *de novo* Pyrimidine Biosynthesis in Trypanosomes and *Leishmania*. FEBS Letters *128*, 27-29.

Hanau, S., Rippa, M., Bertelli, M., Dallocchio, F. and Barrett, M.P. (1996). 6-Phosphogluconate Dehydrogenase from *Trypanosoma brucei*. European Journal of Biochemistry *240*, 592-599.

Hannaert, V., Blaauw, M., Kohl, L., Allert, S., Opperdoes, F.R. and Michels, P.A.M. (1992). Molecular Analysis of the Cytosolic and Glycosomal Glyceraldehyde 3-Phosphate Dehydrogenase in *Leishmania mexicana*. Molecular and Biochemical Parasitology *55*, 115-126.

Hannaert, V., Saavedra, E., Duffieux, F., Szikora, J.P., Rigden, D.J., Michels, P.A.M. and Opperdoes, F.R. The Trypanosomatidae have Lost a Secondarily Aquired Chloroplast. Unpublished.

Hansen, H., Didion, T., Thiemann, A., Veenhuis, M. and Roggenkamp, R. (1992). Targeting Sequences of the Two Major Peroxisomal Proteins in the Methylotrophic Yeast *Hansenula polymorpha*. Mol Gen Genet *235*, 269-278.

Hart, D.T. and Coombs, G.H. (1981). The Effects of Carbon Dioxide and Oxygen Upon the Growth and *In Vitro* Transformation of *Leishmania mexicana mexicana*. Molecular and Biochemical Parasitology *4*, 117-127.

Hart, D.T. and Coombs, G. (1982). *Leishmania mexicana*: Energy Metabolism of Amastigotes and Promastigotes. Experimental Parasitology *54*, 397-409.

Hart, D.T. and Opperdoes, F.R. (1984). The Occurrence of Glycosomes (Microbodies) in the Promastigote Stage of Four Major *Leishmania* Species. Molecular and Biochemical Parasitology *13*, 159-172.

Hassan, H.F., Mottram, J.C. and Coombs, G.H. (1985). Subcellular Localisation of Purine-Metabolising Enzymes in *Leishmania mexicana mexicana*. Comparative Biochemistry and Physiology *81B*, 1037-1040.

Hawkins, C.F., Borges, A. and Perham, R.N. (1989). A Common Structural Motif in Thiamin Pyrophosphate-Binding Enzymes. FEBS Letters *255*, 77-82.

Hecquet, L., Bolte, J. and Demuynck, C. (1996). Enzymatic Synthesis of "Natural-labelled" 6-deoxy-L-sorbose Precursor of an Important Food Flavour. Tetrahedron *52*, 8223-8232.

Heinrich, P.C., Steffen, H., Janser, P. and Wiss, O. (1972). Studies on the Reconstitution of Apotransketolase with Thiamine Pyrophosphate and Analogs of the Coenzyme. European Journal of Biochemistry *30*, 533-541.

Heise, N. and Opperdoes, F.R. (1997). The Dihydroxyacetonephosphate Pathway for Biosynthesis of Ether Lipids in *Leishmania mexicana* promastigotes. Molecular and Biochemical Parasitology *89*, 61-72.

Heise, N. and Opperdoes, F.R. (1999). Purification, Localisation and Characterisation of Glucose-6-Phosphate Dehydrogenase of *Trypanosoma brucei*. Molecular and Biochemical Parasitology *99*, 21-32.

Henkes, S., Sonnewald, U., Badur, R., Flachmann, R. and Stitt, M. (2001). A small Decrease of Plastid Transketolase Activity in Antisense Tobacco Transformants Has Dramatic Effects on Photosynthesis and Phenylpropanoid Metabolism. The Plant Cell *13*, 535-551.

Henze, K., Badr, A., Wettern, M., Cerff, R. and Martin, W. (1995). A Nuclear Gene of Eubacterial Origin in *Euglena gracilis* reflects Cryptic Endosymbioses During Protist Evolution. Proc. Natl. Acad. Sci. USA *92*, 9122-9126.

Huang, J. and Van der Ploeg, L.H.T. (1991). Requirement of a Polypyrimidine Tract for *Trans*-Splicing in Trypanosomes: Discriminating the PARP Promoter from the Immediately Adjacent 3' Splice Acceptor Site. The EMBO Journal *10*, 3877-3885.

Hug, M., Hotz, H., Hartmann, C. and Clayton, C. (1994). Hierarchies of RNA-Processing Signals in a Trypanosome Surface Antigen mRNA Precursor. Molecular and Cellular Biology *14*, 7428-7435.

Iida, A., Teshiba, S. and Mizobuchi, K. (1993). Identification and Characterisation of the *tktB* Gene Encoding a Second Transketolase in *Escherichia coli* K-12. Journal of Bacteriology *175*, 5375-5383.

Ilg, T. (2000). Lipophosphoglycan is Not Required for Infection of Macrophages or Mice By *Leishmania mexicana*. The EMBO Journal *19*, 1953-1962.

Ivens, A.C., Lewis, S.M., Bagherzadeh, A., Zhang, L., Chan, H.M. and Smith, D. (1998). A Physical Map of the *Leishmania major* Friedlin Genome. Genome Research *8*, 135-145.

Janovy, Jr. J. (1972). Temperature and Metabolism in *Leishmania* III. Some Dehydrogenases of *L. donovani*, *L. mexicana* and *L. tarentolae*. Experimental Parasitology *32*, 196-205.

Janowicz, Z.A., Eckart, M.R., Drewke, C., Roggenk, R.O., Hollenberg, C.P., Maat, J., Ledeboer, A.M., Visser, C. and Verrips, C.T. (1985). Cloning and Characterisation of the DAS Gene Encoding the Major Methanol Assimilatory Enzyme from the Methylotrophic Yeast *Hansenula-Polymorpha*. Nucleic Acids Research *13*, 3043-3062.

Jester, J.V., Moller-Pedersen, T., Huang, J., Sax, C.M., Kays, W.T., Cavangh, H.D., Petroll, W.M. and Piatigorsky, J. (1999). The Cellular Basis of Corneal Transparency: Evidence for 'Corneal Crystallins'. Journal of Cell Science *112*, 613-622.

Jomaa, H., Wiesner, J., Sanderbrand, S., Altincicek, B., Weidemeyer, C., Hintz, M., Türbachova, I., Eberl, M., Zeidler, J., Lichtenthaler, H.K., Soldati, D. and Beck, E. (1999). Inhibitors of the Nonmevalonate Pathway of Isoprenoid Biosynthesis as Antimalarial Drugs. Science *285*, 1573-1576.

Josephson, B.L. and Fraenkel, D.G. (1969). Transketolase Mutants of *Escherichia coli*. Journal of Bacteriology *100*, 1289-1295.

Josephson, B.L. and Fraenkel, D.G. (1974). Sugar Metabolism in Transketolase Mutants of *Escherichia coli*. Journal of Bacteriology *118*, 1082-1089.

Joshi, P.B., Webb, J.R., Davies, J.E. and McMaster, W.R. (1995). The Gene Encoding Streptothricin Acetyltransferase (*sat*) as a Selectable Marker for *Leishmania* Expression Vectors. Gene *156*, 145-149.

Juhnke, H., Krems, B., Kotter, P. and Entian, K.E. (1996). Mutants That Show Increased Sensitivity to Hydrogen Peroxide Reveal an Important Role for the Pentose Phosphate Pathway in Protection of Yeast Against Oxidative Stress. Molecular General Genetetics *252*, 456-464.

Keegan, F.P., Sansone, L. and Blum, J.J. (1987). Oxidation of Glucose, Ribose, Alanine, and Glutamine by *Leishmania braziliensis panamensis*. Journal of Protozoology *34*, 174-179.

Kelly, J.M., Ward, H.M., Miles, M.A. and Kendall, G. (1992). A Shuttle Vector Which Facilitates the Expression of Transfected Genes in *Trypanosoma cruzi* and *Leishmania*. Nucleic Acid Research *20*, 3963-3969.

Kendall, G., Wilderspin, A.F., Ashall, F., Miles, M.A. and Kelly, J.M. (1990). *Trypanosoma cruzi* Glycosomal Glyceraldehyde-3-Phosphate Dehydrogenase does Not Conform to the 'Hotspot' Topogenic Signal Model. The EMBO Journal *9*, 2751-2758.

Kochetov, G.A. (1982). Transketolase from Yeast, Rat Liver, and Pig Liver. Methods in Enzymology *90*, 209-223.

Kochetov, G.A. (2001). Functional Flexibility of the Transketolase Molecule. Biochemistry (Moscow) *66*, 1077-1085.

Kohl, L., Callens, M., Wierenga, R.K., Opperdoes, F.R. and Michels, P.A.M. (1994). Triose-Phosphate Isomerase of *Leishmania mexicana mexicana*. European Journal of Biochemistry *220*, 331-338.

Kovina, M.V., Tikhonova, O.V., Solov'eva, O.N., Bykova, I.A., Ivanov, A.S. and Kochetov, G.A. (2000) Influence of Transketolase Substrates on Its Conformation. Biochemical and Biophysical Communications *275*, 968-972.

Krauth-Siegel, R.L. and Coombs, G.H. (1999). Enzymes of Parasite Thiol Metabolism as Drug Targets. Parasitology Today *15*, 404-409.

Krepinsky, K., Plaumann, M., Martin, W. and Schnarrenberger, C. (2001). Purification and Cloning of Chloroplast 6-Phosphogluconate Dehydrogenase from Spinach. European Journal of Biochemistry *268*, 2678-2686.

Kuimov, A., Filippov, M. and Kochetov, G. (1990). Multiple Forms of Transketolase. Biochemistry International *21*, 1081-1087.

Laban, A., Tobin, J.F., Curotto de Lafaille, M.A. and Wirth, D.F. (1990). Stable Expression of the Bacterial neo^r Gene in *Leishmania enriettii*. Nature *343*, 572-574.

Laemmli, U.K. (1970). Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. Nature *227*, 680-685.

Lainson, R. (1983). The American Leishmaniases: Some Observations On their Ecology and Epidemiology. Transactions of the Royal Society of Tropical Medicine and Hygiene *77*, 569-566.

Lander, E.S. *et.al.* (2001). Initial Sequencing and Analysis of the Human Genome. Nature *409*, 860-921.

Lange, B.M., Wildung, M.R., McCaskill, D. and Croteau, R. (1998). A Family of Transketolases that Directs Isoprenoid Biosynthesis Via a Mevalonate-Independent Pathway. Proc. Natl. Acad. Sci. USA *95*, 2100-2104.

Latruffe, N. and Vamecq, J. (2000). Evolutionary Aspects of Peroxisomes as Cell Organelles, and of Gene Encoding Peroxisomal Proteins. Biology of the Cell *92*, 389-395.

LeBowitz, J.H., Coburn, C.M., McMahon-Pratt, D. and Beverley, S.M. (1990). Development of a Stable *Leishmania* Expression Vector and Application to the Study of Parasite Antigen Genes. Proc. Natl. Acad, Sci. USA *87*, 9736-9740.

LeBowitz, J.H., Smith, H.Q., Rusche, L. and Beverley, S.M. (1993). Coupling of Poly(A) Site Selection and *Trans*-Splicing in *Leishmania*. Genes and Developement *7*, 996-1007.

LeBowitz, J.H. (1994). Transfection Experiments with *Leishmania*. Methods in Cell Biology *45*, 65-78.

Lichtenthaler, H.K., Schwender, J., Disch, A. and Rohmer, M. (1997). Biosynthesis of Isoprenoids in Higher Plant Chloroplasts Proceeds Via a Mevalonate-Independent Pathway. FEBS Letters *400*, 271-274.

Lilly, M.D. *et al.* (1996). Carbon-Carbon Bond Synthesis: The Impact of rDNA Technology on the Production and use of *Escherichia coli* Transketolase. Annals of the New York Academy of Sciences *782*, 513-525.

Lindqvist, Y., Schneider, G., Ermler, U. and Sundström, M. (1992). Three-Dimensional Structure of Transketolase, a Thiamine Diphosphate Dependent Enzyme. The EMBO Journal *11*, 2373-2379.

Littlechild, J. and Turner, N. (1995). Crystallisation and Preliminary X-ray Crystallographic Data with *Escherichia coli* Transketolase. Acta Crystallographica *D51*, 1074-1076.

Lois, L.M., Campos, N., Putra, S.R., Danielsen, K., Rohmer, M. and Boronat, A. (1998). Cloning and Characterisation of a Gene from *Escherichia coli* Encoding a Transketolase-like Enzyme that Catalyses the Synthesis of D-1-Deoxyxylulose 5-Phosphate, a Common Precursor for Isoprenoid, Thiamin, and Pyridoxol Biosynthesis. Proceedings of the National Academy of Sciences of the United States of America *95*, 2105-2110.

Mancilla, R. and Naquira, C. (1964). Comparative Metabolism of C^{14} Glucose in Two Strains of *Trypanosoma cruzi*. Journal of Protozoology *11*, 509-513.

Marchand, M., Kooystra, U., Wierenga, R.K., Lambeir, A.M., Beemen, J.V., Opperdoes, F.R. and Michels, P.A.M. (1989). Glucosephosphate Isomerase from *Trypanosoma brucei*. European Journal of Biochemistry *184*, 455-464.

Martin, E., Simon, M.W., Schaefer, F.W. and Mukkada, A.J. (1976). Enzymes of Carbohydrate Metabolism in Four Human Species of *Leishmania*: A Comparative Study. Journal of Protozoology *23*, 600-607.

Martin, W. and Schnarrenberger, C. (1997). The Evolution of the Calvin Cycle from Prokaryotic to Eukaryotic Chromosomes: A Case Study of Functional Redundancy in Ancient Pathways through Endosymbiosis. Current Genetics *32*, 1-18.

Matlashewski, G. (2001). *Leishmania* Infection and Virulence. Medical Microbiology and Immunology *190*, 37-42.

Matthews, K.R., Tschudi, C. and Ullu, E. (1994). A Common Pryrimidine-rich Motif Governs Trans-splicing and Polyadenylation of Tubulin Polycistronic Pre-mRNA in Trypanosomes. Genes and Development *8*, 491-501.

McCool, B.A., Plonk, S.G., Martin, P.R. and Singleton, C.K. (1993). Cloning of Human Transketolase cDNAs and Comparison of the Nucleotide Sequence of the Coding Region, in Wernicke-Korsakoff and Non-Wernicke-Korsakoff Individuals. The Journal of Biological Chemistry *268*, 1397-1404.

Medina-Acosta, E. and Cross, G.A.M. (1993). Rapid Isolation of DNA from Trypanosomatid Protozoa using a Simple 'Mini-Prep' Procedure. Molecular and Biochemical Parasitology *59*, 327-330.

Meshalkina, L., Nilsson, U., Wikner, C., Kostikowa, T., and Schneider, G. (1997). Examination of the Thiamin Binding Site in yeast Transketolase by Site-Directed Mutagenesis. European Journal of Biochemistry *244*, 646-652.

Michels, P.A.M., Hannaert, V. and Bingaud, F. (2000). Metabolic Aspects of Glycosomes in Trypanosomatidae-New Data and Views. Parasitology Today *16*, 482-489.

Mishra, M., Biswas, U.K., Jha, D.N. and Khan, A.B. (1992). Amphotericin Versus Pentamidine in Antimony-Unresponsive Kala-Azar. The Lancet *340*, 1256-1257.

Misset, O., Van Beeumen, J., Lamber, A., Van der Meer, R. and Opperdoes, F.R. (1987). Glyceraldehyde-phosphate dehydrogenase from *Trypanosoma brucei*. Comparison of the Glycosomal and Cytosolic Isoenzymes. European Journal of Biochemistry *162*, 501-507.

Misset, O. and Opperdoes, F.R. (1987). The Phosphoglycerate Kinases from *Trypanosoma brucei*. A Comparison of the Glycosomal and the Cytosolic Isoenzymes and their Sensitivity Towards Suramin. European Journal of Biochemistry *162*, 493-500.

Mottram, J.C. and Coombs, G.H. (1985). *Leishmania mexicana*: Subcellular Distribution of Enzymes in Amastigotes and Promastigotes. Experimental Parasitology *59*, 265-274.

Mottram, J.C., Souza, A.E., Hutchison, J.E., Carter, R., Frame, M.J. and Coombs, G.H. (1996). Evidence from Disruption of the *lmcpb* Gene Array of *Leishmania mexicana* that Cysteine Proteinases are Virulence Factors. Proceedings of the National Academy of Sciences of the United States of America *93*, 6008-6013.

Myler, P.J. *et al.* (1999). *Leishmania major* Friedlin Chromosome 1 has an Unusual Distribution of Protein-Coding Genes. Proceedings of the National Academy of Sciences of the United States of America *96*, 2902-2906.

Myler, P.J. *et al.* (2001). The *Leishmania* Genome Project: New Insights into Gene Organisation and Function. Medical Microbiology and Immunology *190*, 9-12.

Nikkola, M., Lindqvist, Y. and Schneider, G. (1994). Refined Structure of Transketolase from *Saccharomyces cerivisiae* at 2.0 Å Resolution. Journal of Molecular Biology *238*, 387-404.

Nilsson, U., Meshalikina, L., Lindqvist,Y. and Schneider, G. (1997). Examination of Substrate Binding in Thiamin Diphosphate-dependent Transketolase by Protein Crystallography and Site-directed Mutagenesis. The Journal of Biological Chemistry *272*, 1864-1869.

Nilsson, U., Hecquet, L., Gefflaut, T., Guerard, C. and Schneider, G. (1998). Asp^{477} is a Determinant of the Enantioselectivity in Yeast Transketolase. FEBS Letters *424*, 49-52.

Ninfali, P., Malatesta, M., Biagiotti, E., Aluigi, G. and Gazzanelli, G. (2001). Glucose 6-Phosphate Dehydrogenase in Small Intestine of Rabbit: Biochemical Properties and Subcellular Localisation. Acta Histochemica *103*, 287-303.

Nyame, K., Do-Thi, C.D., Opperdoes, F.R. and Michels, P.A.M. (1994). Subcellular Distribution and Characterisation of Glucosephosphate Isomerase in *Leishmania mexicana mexicana*. Molecular and Biochemical Parasitology *1994*, 269-279.

Olliaro, P.L. and Bryceson, A.D.M. (1993). Practical Progress and New Drugs for Changing Patterns of Leishmaniasis. Parasitology Today *9*, 323-328.

Opperdoes, F.R. and Borst, P. (1977). Location of Nine Glycolytic Enzymes in a Microbody-Like Organelle in *Trypanosoma brucei*: The Glycosome. FEBS Letters *80*, 360-364.

Opperdoes, F.R. (1984). Localisation of the Initial Steps of Alkoxyphospholipid Biosynthesis in Glycosomes (Microbodies) of *Trypanosoma brucei*. FEBS Letters *169*, 35-39.

Opperdoes, F.R. (1987). Compartmentation of Carbohydrate Metabolism in Trypanosomes. Annual Review of Microbiology *41*, 127-51.

Opperdoes, F.R. and Michels, P.A.M. (2001). Enzymes of Carbohydrate Metabolism as Potential Drug Targets. International Journal for Parasitology *31*, 482-490.

Page, R.D.M. (1996). TREEVIEW: An Application to Display Phylogenetic Trees on Personel Computers. Computer Applications in the Biosciences *12*, 357-358

Paolettii, F. (1983) Purification and Properties of Transketolase from Fresh Rat Liver. Archives of Biochemistry and Biophysics *222*, 489-496.

Peterson, G.C., Sommer, J.M., Klosterman, S., Wang, Ç.C. and Parsons, M. (1997). *Trypanosoma brucei*: Identification of an Internal Region of Phosphoglycerate Kinase Required for Targeting to Glycosomal Microbodies. Experimental Parasitology *85*, 16-23.

Phillips, C., Barrett, M.P., Gover, S., Le Page, W.F. and Adams, M.J. (1993). Preliminary Crystallographic Study of 6-Phosphogluconate Dehydrogenase from *Trypanosoma brucei*. Journal of Molecular Biology *233*, 317-321.

Phillips, C., Dohnalek, J., Gover, S., Barrett, M.P. and Adams, M.J. (1998). A 2.8 Å Resolution Structure of 6-Phosphogluconate Dehydrogenase from the Protozoan Parasite *Trypanosoma brucei*: Comparison with the Sheep Enzyme Accounts for Differences in Activity with Coenzyme and Substrate Analogues. Journal of Molecular Biology *282*, 667-681.

Puentes, S.M., Da Silva, R.P., Sacks, D.L., Hammer, C.H. and Joiner, K.A. (1990). Serum Resistance of Metacyclic Stage *Leishmania major* Promastigotes is Due to Release of C5b-9. The Journal of Immunology *145*, 4311-4316.

Querol, J., Rodriguez-Concepcion, M., Boronat, A. and Imperial, S. (2001). Essential Role of Residue H49 for Activity of *Escherichia coli* 1-Deoxy-D-xylulose 5-Phosphate Synthase, the Enzyme Catalysing the First Step of the 2-C-Methyl-D-erythritol 4-Phosphate Pathway for Isoprenoid Synthesis. Biochemical and Biophysical Research Communications *289*, 155-160.

Racker, E. (1961). Transketolase. In The Enzymes, V, P.D.Boyer, H.Lardy, and E.Myrback, eds. (New York: Academic Press), pp. 397-406.

Robinsons, B.H. and Chun, K. (1993). The Relationships Between Transketolase, Yeast Pyruvate Decarboxylase and Pyruvate Dehydrogenase of the Pyruvate Dehydrogenase Complex. FEBS Letters *328*, 99-102.

Rohmer, M., Knani, M., Simonin, P., Sutter, B. and Sahm, H. (1993). Isoprenoid Biosynthesis in Bacteria: A Novel Pathway for the Early Steps Leading to Isopentenyl diphosphate. Journal of Biochemistry *295*, 517-524.

Ryan, K.A., Dasgupta, S. and Beverley, S.M. (1993). Shuttle Cosmid Vectors for the Trypanosomatid Parasite *Leishmania*. Gene *131*, 145-150.

Saitou, N. and Nei, M. (1987). The Neighbour-joining Method: A New Method for Reconstructing Pylogenetic Trees. Molecular Biology and Evolution *4*, 406-425.

Salamon, C., Chervenak, M., Piatigorsky, J. and Sax, C.M. (1998). The Mouse Transketolase (TKT) Gene: Cloning, Characterisation and Functional Promoter Analysis. Genomics *48*, 209-220.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular Cloning, A Laboratory Manual. Cold Spring Harbour Laboratory Press.

Sax, C.M., Salamon, C., Kays, W.T., Guo, J., Yu, F.X., Cuthbertson, R.A. and Piatigorsky, J. (1996). Transketolase is a Major Protein in the Mouse Cornea. The Journal of Biological Chemistry *271*, 33568-33574.

Schaff-Gerstenschläger, I. and Zimmermann, F.K. (1993). Pentose-Phosphate Pathway in *Saccharomyces cerevisiae*: Analysis of Deletion Mutants for Transketolase, Transaldolase, and Glucose 6-Phosphate Dehydrogenase. Current Genetics *24*, 373-376.

Schenk, G., Layfield, R., Candy, J.M., Duggleby, R.G. and Nixon, P.F. (1997). Molecular Evolutionary Analysis of the Thiamine-Diphosphate-Dependent Enzyme Transketolase. Journal of Molecular Evolution *44*, 552-572.

Schenk, G., Duggleby, R.G. and Nixon, P.F. (1998(a)). Properties and Functions of the Thiamin Diphosphate Dependent Enzyme Transketolase. The International Journal of Biochemistry and Cell Biology *30*, 1297-1318.

Schenk, G., Duggleby, R.G. and Nixon, P.F. (1998(b)). Heterologous Expression of Human Transketolase. The International Journal of Biochemistry and Cell Biology 30, 369-378.

Schlein, Y., Jacobson, R.L. and Messer, G. (1992). *Leishmania* Infections Damage the Feeding Mechanism of the Sandfly Vector and Implement Parasite Transmission by Bite. Proceedings of the National Academy of Sciences of the United States of America *89*, 9944-9948.

Schlein, Y. (1993). *Leishmania* and Sandflies: Interactions in the Life Cycle and Transmission. Parasitology *9*, 255-257.

Schürch, N., Hehl, A., Vassella, E., Braun, R. and Roditi, I. (1994). Accurate Polyadenylation of Procyclin mRNAs in *Trypanosoma brucei* is Determined by Pyrimidine-Rich Elements in the Intergenic Regions. Molecular and Cellular Biology *14*, 3668-3675.

Smeets, E.H.J., Muller, H. and De Wael, J. (1971). An NADH-Dependent Transketolase Assay in Erythrocyte Hemolysates. Clinica Chimica Acta *33*, 379-386.

Solov'eva, O.N., Bykova, I.A., Meshalikina, L., Kovina, M.V. and Kochetov, G. (2001). Cleaving of Ketosubstrates by Transketolase and the Nature of the Products Formed. Biochemistry (Moscow) *66*, 932-936.

Solov'eva, O.N., Meshalkina, L.E., Kovina, M.V., Selivanov, V.A., Bykova, I.A. and Kochetov, G.A. (2000). Acceptor Substrates Inhibits Transketolase Competitively with Respect to Donor Substrate. Biochemistry (Moscow) *65*, 1202-1205.

Sommer, J.M., Cheng, Q., Keiler, G. and Wang, C.C. (1992). *In Vivo* Import of Firefly Luciferase into the Glycosomes of *Trypanosoma brucei* and Mutational Analysis of the C-Terminal Targeting Signal. Molecular Biology of the Cell *3*, 749-759.

Souza, A.E., Bates, P.A., Coombs, G.H. and Mottram, J.C. (1994). Null Mutants for the *lmcpa* Cysteine Proteinase Gene in *Leishmania mexicana*. Molecular and Biochemical Parasitology *63*, 213-220.

Sprenger, G.A. (1993). Nucleotide Sequence of the *Escherichia coli* K-12 Transketolase (tkt) Gene. Biochimica et Biophysica Acta *1216*, 307-310.

Sprenger, G.A., Schorken, U., Sprenger, G. and Sahm, H. (1995). Transketolase A of *Escherichia coli* K12. Purification and Properties of the Enzyme from Recombinant Strains. European Journal of Biochemistry *230*, 525-532.

Stiles, J.K., Hicock, P.I., Shah, P.H. and Meade, J.C. (1999). Genomic Organisation, Transcription, Splicing and Gene Regulation in *Leishmania*. Annals of Tropical Medicine and Parasitology *93*, 781-807.

Stryer, L. (1988). Pentose Phosphate Pathway and Glucogenesis. In Biochemistry, (New York: W.H. Freeman and Company), pp. 427-444.

Studier, F.W. and Moffat, B.A. (1986). Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Genes. Journal of Molecular Biology *189*, 113-130.

Sundström, M., Lindqvist, Y., Schneider, G., Hellman, U. and Ronne, H. (1993). Yeast TKL1 Gene Encodes a Transketolase that is Required for Efficient Glycolysis and Biosynthesis of Aromatic Amino Acids. Journal of Biological Chemistry *268*, 24346-24352.

Teige, M., Melzer, M. and Süss, K.H. (1998). Purification, Properties and *in situ* Localisation of the Amphibolic Enzymes D-Ribulose 5-phosphate 3-epimerase and Transketolase from Spinach Chloroplasts. European Journal of Biochemistry *252*, 237-244.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994). CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignments through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. Nucleic Acids Research *22*, 4673-4680.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. and Higgins, D.G. (1997). The CLUSTAL_X windows Interface: Flexible Strategies for Multiple Sequence Alignment Aided by Quality Analysis Tools. Nucleic Acids Research *25*, 4876-4882.

Tielens, A.G.M. and Van Hellemond, J.J. (1998). Differences in Energy Metabolism Between Trypanosomatidae. Parasitology Today *14*, 265-271.

Tikhomirova, N.K., Merchan, A.Y. and Kochetov, G.A. (1990). A New Form of Baker's Yeast Transketolase. An Enzyme-RNA complex. FEBS Letters *274*, 27-29.

Tobin, J.F., Laban, A. and Wirth, D.F. (1991). Homologous Recombination in *Leishmania enriettii*. Proceedings of the National Academy of Sciences of the United States of America *88*, 864-868.

Tobin, J.F. and Wirth, D.F. (1992). A Sequence Insertion Targeting Vector for *Leishmania enriettii*. Journal of Biological Chemistry *267*, 4752-4758.

Tovar, J., Wilkinson, S., Mottram, J.C. and Fairlamb, A.H. (1998). Evidence that Trypanothione Reductase is an Essential Enzyme in *Leishmania* by Targeted Replacement of the *tryA* Gene Locus. Molecular Microbiology *29*, 653-660.

Van Hellemond, J.J. and Tielens, A.G.M. (1997). Inhibition of the Respiratory Chain Results in a Reversible Metabolic Arrest in *Leishmania* Promastigotes. Molecular and Biochemical Parasitology *85*, 135-138.

Van Hellemond, J.J., Van der Meer, P. and Tielens, A.G.M. (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. and Tielens, A.G.M. (1998). Trypanosomatidae Produce Acetate Via a Mitochondrial Acetate:Succinate CoA Transferase. Proceedings of the National Academy of Sciences of the United States of America *95*, 3036-3041.

Vanhamme, L. and Pays, E. (1995). Control of Gene Expression in Trypanosomes. Microbiological Reviews *59*, 223-240.

Villafranca, J.J. and Axelrod, B. (1971). Heptulose Synthesis from Nonphosphorylated Aldoses and Ketoses by Spinach Transketolase. Journal of Biological Chemistry *246*, 3126-3131.

Visser, N., Opperdoes, R. and Borst, P. (1981). Subcelluar Compartmentation of Glycolytic Intermediates in *Trypanosoma brucei*. European Journal of Biochemistry *118*, 521-526.

Walsh, J.F., Molyneux, D.H. and Birley, M.H. (1993). Deforestation: Effects on Vector-borne Disease. Parasitology *106*, 555-575.

Wang, J.J.L., Martin, P.R. and Singleton, C.K. (1997). Aspartate 155 of Human Transketolase is Essential for Thiamine Diphosphate-Magnesium Binding, and Cofactor Binding is Required for Dimer formation. Biochimica et Biophysica Acta *1341*, 165-172.

WHO (2001/2002) web site: http//www.who.int/

Wiemer, E.A.C., Hannaert, V., Van den Ijssel, P.R.L.A., Van Roy, J., Opperdoes, F.R. and Michels, P.A.M. (1995). Molecular Analysis of Glyceraldehyde 3-Phosphate Dehydrogenase in *Trypanosoma borelli*: An Evolutionary Scenario of Subcellular Compartmentation in Kinetoplastida. Journal of Molecular Evolution *40*, 443-454.

Wierenga, R.K. *et al.* (1987). Common Elements on the Surface of Glycolytic Enzymes from T*rypanosoma brucei* May Serve as Topogenic Signals for Import into Glycosomes. The EMBO Journal *6*, 215-221.

Wikner, C., Meshalikina, L., Nilsson, U., Nikkola, M., Lindqvist, Y., Sundström, M. and Schneider, G. (1994). Analysis of an Invariant Cofactor-Protein Interaction in Thiamin Diphosphate-Depndent Enzymes by Site-Directed Mutagenesis. The Journal of Biological Chemistry *269*, 32144-32150.

Wikner, C., Meshalkina, L., Nilsson, U., Backstrom, S., Lindvuist, Y. and Schneider, G. (1995). His 103 in Yeast Transketolase is Required for Substrate Recognition and Catalysis. European Journal of Biochemistry *233*, 750-755.

Wikner, C., Milsson, U., Meshalikina, L., Udekwu, C., Lindquist, Y. and Schneider, G. (1997). Identification of Catalytically Important Residues in Yeast Transketolase. Biochemistry *36*, 15643-15649.

Wilson, M.E., Hardin, K.K. and Donelson, J.E. (1989). Expression of the Major Surface Glycoprotein of *Leishmania donovani chagasi* in Virulent and Attenuated Promastigotes. The Journal of Immunology *143*, 678-684.

Wood, T. (1986(a)). Physiological Functions of the Pentose Phosphate Pathway. Cell Biochemistry and Function *4*, 241-247.

Wood, T. (1986(b)). Distribution of The Pentose Phosphate Pathway in Living Organisms. Cell Biochemistry and Function *4*, 235-240.

Zaheer Baquer, N. and McLean, P. (1972). Evidence for the Existence and Functional Activity of the Pentose Phosphate Pathway in the Large Particle Fraction Isolated from Rat Tissues. Biochemical and Biophysical Research Communications *46*, 167-174.

Zhao, G. and Winkler, M.E. (1994). An *Escherichia coli* K-12 *tktA tktB* Mutant Deficient in Transketolase Activity Requires Pyroxidine (Vitamin B_6) as Well as the Aromatic Amino Acids and Vitamins for Growth. Journal of Bacteriology *176*, 6134-6138.

Zimmermann, F.T., Schneider, A., Schorken, U., Sprenger, G.A. and Fessner, W. (1999). Efficient Multi-Enzymatic Synthesis of D-Xylulose 5-Phosphate. Tetrahedron: Asymmetry *10*, 1643-1646.

Zomer, A.W.M., Michels, P.A.M. and Opperdoes, F.R. (1999). Molecular Characterisation of *Trypanosoma brucei* Alkyl Dihydroxyacetone-Phosphate Synthase. Molecular and Biochemical Parasitology *104*, 55-66.