DEVELOPMENTALLY REGULATED GENES
IN *TRYPANOSOMA CRUZI*

Thesis submitted to the University of London
for the degree of PhD

EVA GLUENZ

London School of Hygiene
& Tropical Medicine

2005

Numerous Originals in Colour

ABSTRACT

Trypanosoma cruzi, the causative agent of Chagas disease, is transmitted by blood feeding triatomine bugs. As *T. cruzi* cycles between the insect vector and vertebrate host, it goes through several distinct developmental stages. The molecular mechanism that regulates these differentiation steps is poorly understood. We have initiated a systematic dissection of metacyclogenesis, the differentiation step from replicating non-infective epimastigotes to non-replicating infective metacyclic trypomastigotes. Our aim is to functionally characterise novel stage-regulated genes, using transfection-based approaches. We demonstrate that one of these genes, *MET3*, encodes a protein that localises to the nucleus and associates with nucleolar antigens. During metacyclogenesis the structure of the nucleus is modified: heterochromatin spreads as the nucleus changes shape from round to elongated, and nucleolar antigens are dispersed in the metacyclic nucleus. Our data show that while dispersal of nucleolar antigens occurs in all cells of a stationary phase culture, MET3 protein is expressed only in the subset of cells that have differentiated to metacyclics or to intermediate forms. By expressing MET3-GFP fusion proteins in the parasite, we identified two sequence elements that can independently direct localisation. To address function, we generated *MET3* null mutants. These knockout cells produced significantly fewer fully developed metacyclics than wild-type controls. However, this phenotype could not be complemented with an ectopic copy of *MET3*. The reduced rate of metacyclogenesis may therefore be unrelated to the loss of *MET3*. The knockout cells are able to complete the entire life-cycle *in vitro*. This demonstrates that *MET3* is not essential for development of the infective stage. RNA transcripts of the second gene, (*MET2*) are induced during metacyclogenesis. Recombinant MET2 protein associates with the kinetoplast when expressed in epimastigotes. We analysed in detail its genomic locus near a putative centromere on chromosome 3, and generated deletion mutants for further investigating its role in differentiation.

2

CONTENTS

3

4

5

6

LIST OF TABLES AND FIGURES

8

ABBREVIATIONS

bp	base pairs
cAMP	5'-3'-cyclic adenosine monophosphate
cDNA	copy DNA
CHEFE	clamped homogenous electric fields electrophoresis
DIC	differential interference contrast
DEAE	diethylaminethanol
EDTA	ethylene diamine tetraacetic acid
GFP	green fluorescent protein
GAPDH	glyceraldehyde-3 phosphate dehydrogenase
GPI	glycosylphosphatidylinositol
kb	kilobase pairs
kDa	kilo Daltons
kDNA	kinetoplast DNA
Mb	megabases
NLS	nuclear localisation signal
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PFA	paraformaldehyde
RNAi	RNA interference
rDNA	ribosomal DNA
SDS	sodium dodecyl sulfate
SL	spliced leader sequence
TAU	triatomine artificial urine
TBS	tris buffered saline
Tet	tetracycline
UTR	untranslated region

ACKNOWLEDGEMENTS

1. INTRODUCTION

Trypanosoma cruzi, the causative agent of Chagas disease, is a protozoan flagellate of the order Kinetoplastida, subfamily Trypanosomatidae. Among the trypanosomatids, there are a number of important parasites that cause disease in humans and their domestic animals. Different species of *Leishmania*, which are transmitted by phlebotomine sandflies cause cutaneous and visceral Leishmaniasis. *Trypanosoma brucei*, causes sleeping sickness and the cattle disease Nagana. The last 30 years have seen a re-emergence of sleeping sickness in Sub Saharan Africa, where tsetse flies, which transmit the parasite, are endemic. The World Health Organisation estimates that together, these three parasites cause the deaths of over 120,000 people every year, and disable millions, mostly in the poorest parts of the world (WHO, World Health Report 2002).

Trypanosomes are parasites of huge medical importance. They are also biologically fascinating. Kinetoplastids diverged early from the main eukaryotic lineage and their biochemistry and genetics differs in many ways from that of higher eukaryotes. The study of trypanosomes has led to the discovery of several interesting and novel biological phenomena, including unusual mechanisms of gene expression. The relative ease with which they can be cultured and their genetic tractability have made them the focus of research in many laboratories.

1.1 *TRYPANOSOMA CRUZI* AND CHAGAS DISEASE

Chagas disease is endemic in large parts of Latin America. The disease was first described in 1909 by the Brazilian doctor Carlos Chagas (Chagas, 1909). Remarkably, Chagas also discovered both the causative agent, named *Trypanosoma cruzi* in honour of his mentor Oswaldo Cruz, and the insect vector responsible for transmission of the parasite, and was the first to describe the parasite's life-cycle. Today, Chagas disease is considered the most significant parasitic disease in Latin America, posing a direct risk to 25% of the population. It is estimated that 11-18 million people are infected, with 3-5 million symptomatic cases. The WHO estimates that

every year 13,000 people die of the disease, and a further 200,000 people are newly infected (Barrett et al., 2003; WHO, 2003).

Chagas disease has three phases in humans (Bogitsh and Cheng, 1990). The first, acute phase of the disease is characterised by high parasitemia and can last for several weeks. Although it develops in just a few percent of patients, it can be fatal, especially in small children. The disease can then go unnoticed for many years, during which patients remain asymptomatic. Approximately 10-30% of patients proceed to the chronic form of the disease which ultimately leads to death. Classical symptoms of the chronic phase include premature and progressive heart disease and swelling of the oesophagus and colon (mega syndromes). The factors that trigger development of the chronic phase of the disease remain controversial. An important disputed question is whether the pathogenesis of the chronic disease requires the continued presence of the parasite, or whether the condition is an autoimmune response that occurs even in the absence of parasite (Tarleton, 2003). A recent study demonstrates integration of *T. cruzi* DNA into the host genome, and thus suggests a potential mechanism for pathology in the absence of persistent infection. As this appears to occur at random, it could provide an explanation for the differential activation of the chronic disease (Nitz et al., 2004). Other evidence strongly suggests that the persistence of the parasite is necessary and sufficient for chronic disease pathology (Tarleton et al., 1997). It will be important to resolve this issue, as it has direct implications for determining research priorities.

Trypanosoma cruzi, the causative agent of Chagas disease is transmitted by blood-sucking triatomine bugs (order Hemiptera, family Reduviidae, subfamily Triatominae). *T. cruzi* has a digenetic life-cycle (section 1.3), cycling between the insect vector and a wide range of mammalian host species, which include humans, domestic and wild mammals. Amphibians and birds are refractory to persistent *T. cruzi* infections (Brener, 1973). Over 100 species of triatomine are important in the sylvatic cycle of the disease but few infest houses. Important vectors of Chagas disease in the domestic cycle include *Rhodnius prolixus* (Figure 1.1) and *Triatoma infestans* (Kollien and Schaub, 2000). *T. cruzi* can also be transmitted congenitally

from infected mother to child, and via transfusions with infected blood (WHO, 2003).

Chagas disease is very much associated with poverty. Grass thatched huts with cracked stone or mud walls are often highly infested with triatomine bugs, which will transmit the parasites when they feed on the inhabitants. Public health measures to control the vector, improved housing conditions and the introduction of a more effective health system play decisive roles in stopping transmission to humans. This is illustrated by the success of the 'Southern Cone Initiative'. This initiative, launched in 1991 by seven South American countries (Argentina, Bolivia, Brazil, Chile, Paraguay, Peru and Uruguay), aimed to eliminate the main vector by large scale spraying of houses with insecticides, and to interrupt transfusional transmission by screening blood donors (Moncayo, 1999). As a result, Uruguay and Chile, central and southern Brazil, and parts of Argentina and Paraguay are now certified free of transmission (Miles et al., 2003). Yet, there are still 11-18 million people already infected with the parasite, who urgently require less toxic and more efficient drugs.

Figure 1.1: Adult Rhodnius prolixus taking a blood meal through human skin
(Image: WHO/TDR/Stammers)

There is no immediate prospect of a vaccine against *T. cruzi*, and only two drugs are available to treat the acute phase of Chagas disease, nifurtimox and benznidazole. These drugs have been used for the treatment of Chagas patients since the early 1970s. Both have limited efficacy and often cause severe side effects (Rodriques Coura and de Castro, 2002). Comparative biochemistry and genomics research has discovered a number of interesting differences between trypanosomatid and human cells. These represent potential targets for the rational development of new drugs (Barrett et al., 2003; Rodriques Coura and de Castro, 2002). Sadly, the pharmaceutical industry has little economic incentive to invest in research and development projects that focus on parasitic diseases that affect the poorest people. A joining of forces between the pharmaceutical industry, the WHO and non-governmental organisations, and the launch of the 'Drugs for Neglected Diseases Initiative' (www.dndi.org) offers hope that progress will be made in the development of new drugs on a not-for-profit basis (Barrett et al., 2003).

1.2 BIOLOGY OF *T. CRUZI*

1.2.1 Classification of *T. cruzi* strains

T. cruzi is a highly divergent species. Strains differ in their biological properties and genome organisation. A number of methods have been used to compare and classify different *T .cruzi* strains (Devera et al., 2003). These include isoenzyme patterns, restriction maps of minicircle DNA (Morel et al., 1980) and ribosomal RNA sequences. The current nomenclature divides the species into two principal groups: *T. cruzi* I (sylvatic strains) and *T. cruzi* II (peridomestic strains) (Anonymous, 1999). Recent immunological evidence suggests that chronic disease in human infections is caused predominantly by *T. cruzi* subgroup II (Di Noia et al., 2002).

The reference organism for the genome sequencing project is *T. cruzi* CL Brener (Zingales et al., 1997c). The CL Brener strain was originally cloned from the blood of mice infected with the CL strain. The rationale for choosing CL Brener as the genome reference strain was that it exhibits the following characteristics that make it a relevant model. (i) the strain was isolated from a strictly domiciliary vector (*Tria-*

toma infestans); (ii) it preferentially parasitizes heart and muscle cells; (iii) it shows a clear acute phase in human patients; (iv) it is highly susceptible to the drugs benznidazole and nifurtimox (Zingales et al., 1997a; Zingales et al., 1997b). Phylogenetic analysis later showed that CL Brener appears to be a hybrid of two distantly related lineages of *T. cruzi* II (Machado and Ayala, 2001). This has had major implications in terms of completing the genome project (see below).

<p align="center">1.2.2 The unusual cell biology of *T. cruzi*</p>

T. cruzi epimastigote cells are polar with a single flagellum at the anterior end and a single mitochondrion (Figure 1.2; de Souza, 1999). Like other eukaryotic cells they possess a membrane bounded nucleus, Golgi apparatus and endoplasmic reticulum (ER). In addition, *T. cruzi* shares with other trypanosomatids some unique features not found in higher eukaryotes (de Souza, 2002; de Souza, 2003). A brief overview is presented below.

Unique vesicular organelles. The most extensively studied vesicles of trypanosomatids are the glycosomes, which contain the first seven enzymes of the glycolytic pathway and other metabolic enzymes (Parsons, 2004). Acidocalcisomes are acidic vacuoles which contain a Ca^{2+} - H^+ translocating ATPase activity and a high Ca^{2+} concentration (Docampo et al., 1995; Vercesi et al., 1994). Their physiological role is not well understood. Reservosomes contain cysteine proteinase, and endocytosed macromolecules (de Souza, 2002). In *T. cruzi* they are developmentally regulated and have been implicated in the process of metacyclogenesis (section 1.4.2).
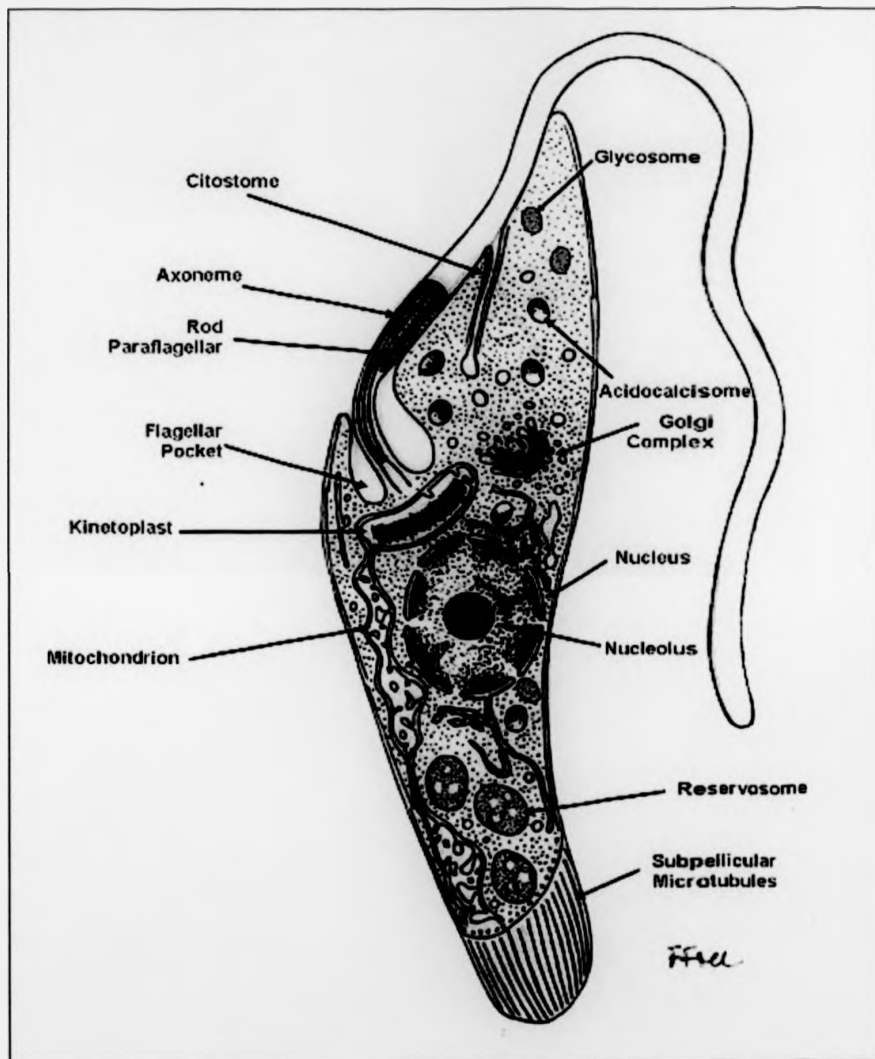
Figure 1.2: Trypanosoma cruzi *epimastigote*

Schematic drawing, based on information obtained with the transmission electron microscope, showing the various structures found in the epimastigote form of *T. cruzi* (from de Souza, 1999).

Flagellum. The single flagellum of trypanosomatids has the classical 9+2 pattern of axonemal microtubules, but also contains an unusual structure known as the para-flagellar rod. At the base of the flagellum, where it emerges from the cell, a special-ised structure is formed, called the flagellar pocket. This is the site where all vesicu-lar traffic to and from the cell surface occurs (Bastin et al., 2000b). The membrane of the flagellar pocket also contains receptor molecules, such as the transferrin receptor (Borst and Fairlamb, 1998). Components of signal transduction pathways, an adenylyl cyclase of *T. brucei* (Paindavoine et al., 1992) and a cAPM-dependent phosphodiesterase of *T. cruzi* (D'Angelo et al., 2004) were found to be concentrated in the flagellar membrane. Thus, the flagellum of trypanosomes has multiple func-tions (Bastin et al., 2000b). It confers motility, it mediates attachment to surfaces (Kleffmann et al., 1998), and these structures are thought to play important roles in environmental sensing.

Cytoskeleton. The cytoskeleton of trypanosomatids is characterised by a corset of subpellicular microtubules, localised below the plasma membrane, which confer rigidity and dictate cell shape (Kohl and Gull, 1998). Different life-cycle stages are defined by distinctive morphologies. The position of kinetoplast and flagellum rela-tive to the nucleus are easily recognised by microscopic observation and commonly used as markers. Two recent studies have provided new insights into the underlying mechanisms of morphogenesis and shown that cytoskeletal components are impor-tant determinants of trypanosomatid cell morphology. During cell division of *T. brucei*, a specialised structure at the tip of the flagellum, called the flagellar attachment zone, is connected to cytoskeletal components. Thus as the new flagellum grows, the old cell guides the morphology of the daughter cell (Moreira-Leite et al., 2001). Kinetoplast repositioning during differentiation of *T. brucei* bloodstream forms was demonstrated to be a microtubule-dependent process. Treatment of differ-entiating cells with the drug rhizoxin (an inhibitor of microtubule assembly) resulted in an almost complete inhibition of kinetoplast movement (Matthews et al., 1995).

Mitochondrion. The mitochondrion is active throughout the life-cycle of *T. cruzi* (Brener, 1973). In this respect *T. cruzi* differs from *T. brucei*, where the mitochon-drion is repressed in bloodstream forms, which obtain their energy from glycolysis

18

alone (Vickerman, 1985). The mitochondrion of *T. cruzi* cells has been shown to have a dynamic structure, however, whose complexity depends on the developmental stage and the cell cycle. The simplest conformation is found in trypomastigotes, and the most complex in replicating epimastigotes (Tyler and Engman, 2001).

The mitochondrial genome. The mitochondrial DNA (kDNA), which represents about 20% of total cellular DNA in *T. cruzi*, is contained in a specialised structure, the kinetoplast. The kDNA is organised in a network of intercatenated circular molecules. *T. cruzi* harbours about 30 maxicircles (23 kb), and 20-30,000 minicircles (1.4 kb) (de Souza, 2003). The maxicircles encode some of the genes for mitochondrial proteins. A peculiar feature of trypanosomes is that in order to generate functional mRNA from these genes, the cryptic transcripts need to undergo extensive modification (RNA editing), which involves posttranscriptional addition or deletion of uridine residues (Estevez and Simpson, 1999). This RNA editing is directed by small RNAs (guide RNAs) which are encoded on both the maxi- and minicircles and mediated by a multi-protein editing complex, the editosome (Stuart and Panigrahi, 2002).

The structure of the *T. cruzi* kinetoplast and packing of kDNA differs between replicating and non-replicating cells. In epimastigotes, the kinetoplast is disc shaped, with compactly organised kDNA, whereas in trypomastigote forms it appears spherical, and the kDNA is less densely packed (Brack, 1968; de Souza, 1999; de Souza, 2003).

Nucleus. The trypanosome nucleus is 2 μm in diameter with a single nucleolus. It is enclosed by a double membrane which is continuous with that of the ER, and contains nuclear pore complexes that are similar to those of other eukaryotes (Rout and Field, 2001). During mitosis the nuclear envelope stays intact (closed mitosis) (De Souza and Meyer, 1974; Ersfeld et al., 1999).

The architecture of the *T. cruzi* nucleus changes markedly during the course of the life-cycle (Belli, 2000; Elias et al., 2001b), a fact that was already noted in Chagas' original recordings (Chagas, 1909). An important determinant of nuclear architecture is chromatin structure (Arney and Fisher, 2004). In trypanosomes, the degree of

19

chromatin condensation and histone profiles, modifications, and abundance vary between different developmental stages (Alsford and Horn, 2004; Belli, 2000; Rout and Field, 2001). In *T. cruzi* epimastigotes and amastigotes, heterochromatin is found at the nuclear periphery. Through an as yet unknown mechanism it becomes dispersed in metacyclics and cell-derived trypomastigotes, where continuous and interconnected patches of heterochromatic material take up most of the nucleoplasmic space (Elias et al., 2001b). Analysis of *T. cruzi* histone H1, which in is the most divergent histone in trypanosomes, provided evidence of stage regulated modifications. A phosphorylated form of histone H1, identified by its distinct mobility pattern in triton-acid-urea polyacrylamide gels, was found to be specifically associated with non-replicating stages. (Marques Porto et al., 2002). Rearrangement of nuclear structure was correlated with decreased transcription rates of RNA polymerase I and II (Elias et al., 2001b). The significance of these changes is not yet fully understood, but it is expected that epigenetic mechanisms play a role in the control of trypanosome gene expression. Work on other organisms has clearly established that during most differentiation processes the nuclear architecture is dramatically altered. How exactly this affects gene expression is an area of intense investigation in many laboratories (Dundr and Misteli, 2001; Manuelidis, 1990).

Nucleolus. The trypanosome nucleolus is visible by fluorescence microscopy as a single 1 μm area of weak DNA staining in the interphase nucleus. In higher eukaryotes, nucleoli are arranged in three morphologically and functionally distinct regions: (i) fibrillar centres (FC), which contain the rDNA, (ii) dense fibrillar centres (DFC) where rRNA transcription occurs, and (iii) granular centres (GC) where ribonucleoprotein particles are assembled (Melese and Xue, 1995). Similarly, in EM images of *T. brucei*, GC and fibrillar centres FC can be distinguished, but DFC are not clearly distinct from FC (Ogbadoyi et al., 2000). During the mitotic cycle in *T. brucei*, the nucleolus elongates, but persists throughout division. In contrast, the nucleolus has been seen to disappear during division of *T. cruzi* amastigotes (De Souza and Meyer, 1974). Dispersal of nucleolar antigens was also found to occur in non-replicating trypomastigote stages of *T. cruzi* (Elias et al., 2001b).

Karyotype. Chromosomes of trypanosomes do not condense during mitosis, and pulse field gel electrophoresis techniques are commonly used to determine the karyotype. According to the latest estimate, the nuclear genome of *T. cruzi* CL Brener is organised in approximately 55 chromosomes which range in size from 0.4 to 4.5 Mb (Andersson, 2004). *T. cruzi* is generally considered to be diploid. The haploid DNA content of *T. cruzi* CL Brener is 40 Mb of DNA and it contains 10-14,000 genes (Andersson, 2004; Taylor and Kelly, 2002).

In nuclei of replicating *T. cruzi* cells, the position of the chromosomes was shown to change as a function of the cell cycle. During S-phase and G2, chromosomes were found to localise to the nuclear periphery. During G1 phase they became dispersed and apparently randomly distributed. A random distribution was also observed in non-replicating forms (Elias et al., 2002).

The mechanism of chromosome segregation is incompletely understood. Results from ultrastructural studies indicate that separation of the genetic material and movement to the daughter nuclei is mediated by nuclear microtubules, analogous to the process in higher eukaryotes. Ultrastructural studies on the *T. cruzi* amastigote nucleus have revealed the presence of intranuclear spindle microtubules. These were found to be associated with electron dense plaques, which could represent kineto-chore-like structures (de Souza, 2003; De Souza and Meyer, 1974). In *T. brucei*, whose nucleus contains three size classes of chromosomes (Ersfeld et al., 1999), the two larger classes were shown to be segregated via interaction with the mitotic spindle (Ogbadoyi et al., 2000). An intriguing problem is how trypanosomatids manage to faithfully segregate their chromosomes, given the observation that the number of kinetochores (visible by EM), is always much smaller than the total number of chromosomes (Ogbadoyi et al., 2000).

Reading the blueprint: the T. cruzi genome project. The *T. cruzi* genome project started in 1994, as part of the WHO-sponsored Parasite Genome Initiative (Zingales et al., 1997c). Since *T. cruzi* is a highly divergent species, one reference clone (CL Brener, section 1.2.1) was chosen for the project. The initial part of the genome project was dedicated to the construction of genomic BAC, YAC and cDNA libraries

and the generation of over 5000 ESTs from a normalised epimastigote cDNA library (Porcel et al., 2000). Large scale shotgun sequencing began in 2000. An annotated version of the genome is expected to be published in 2004, together with the genomes of *T. brucei* and *L. major*. The most immediate impact of this rich source of information on all three trypanosomatid parasites has been the ease with which new genes can be discovered *in silico*, at the click of a button. As in other organisms, a large proportion of open reading frames shows no similarity to previously character-ised genes. For example, the function of 62% of the predicted protein coding sequences on *T. brucei* chromosome I is currently unknown (Hall et al., 2003). The functional analysis of these novel genes will be a major challenge for future research. Interestingly, comparative genomic analysis revealed a high level of synteny (con-served gene order) between the three trypanosomatid species (Bringaud et al., 1998; Ghedin et al., 2004).

Genome. The genome of *T. cruzi* is highly repetitive (Aguero et al., 2000). This is due to non-coding repeat sequences, tandemly repeated housekeeping genes, and multi-copy gene families. There are, for example, more than 600 genes of the *trans*-sialidase super family (Cross and Takle, 1993) dispersed throughout the genome. These include the metacyclic specific *gp90* and *gp82*, and the trypomastigote specific *gp85* genes. The *mucin* gene family (Salazar et al., 1996) is represented by over 700 genes. The genes for the major cysteine proteinase cruzipain are arranged in large clusters of tandemly repeated genes and total nearly 100. The challenging task of assembling such a repetitive genome is made considerably more difficult by the fact that the genome of CL Brener is highly polymorphic. This is due to the hybrid origin of the strain (Machado and Ayala, 2001) which is a likely explanation for the un-usually high sequence divergence between allelic loci, and the considerable size difference (up to 2 Mb) between homologous chromosomes (Andersson, 2004).

22

1.3 THE LIFE-CYCLE OF *T. CRUZI*

1.3.1 Overview

T. cruzi cycles between the insect vector and mammalian host and alternates between replicating forms and non-replicating forms (which are adapted for transmission). Each developmental stage is characterised by a distinct morphology (de Souza, 1999; Elias et al., 2001b; Tyler and Engman, 2000), and differentiation is concomitant with changes in metabolism (Adroher et al., 1990; Urbina, 1994; Vickerman, 1985) and gene expression (Krieger et al., 1999). Molecular approaches are beginning to shed new light on development of the parasite and its remarkable ability to persist and proliferate under the potentially most hostile conditions. The following is a broad overview of the *T. cruzi* life-cycle (Figure 1.3). A detailed account of metacyclogenesis, which is the subject of this study, is presented in section 1.4.
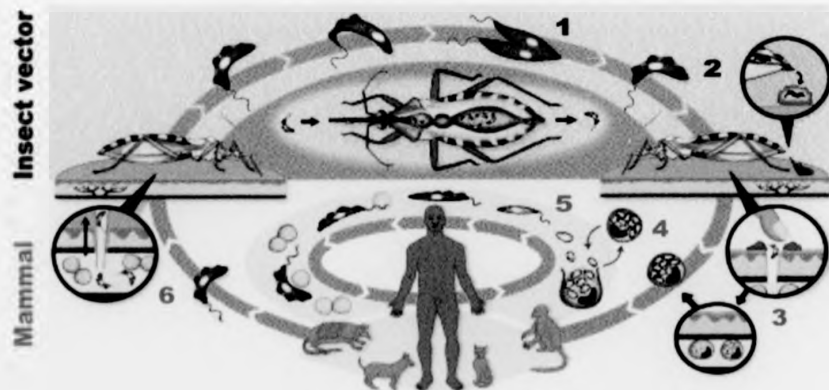
Figure 1.3: The life-cycle of T. cruzi
(1) Epimastigotes multiply in the midgut of the triatomine insect vector. (2) Metacyclics develop in the hindgut, and get excreted with the faeces. (3) Metacyclics that enter through the bite wound cause initial infection of the mammalian host. (4) Parasites multiply intracellularly as amastigotes. (5) Cell-derived trypomastigotes enter the peripheral bloodstream and infect other cells. (6) Infected blood is taken up by a feeding triatomine. Image adapted from WHO/TDR.

Development in the vector. When a triatomine bug feeds on an infected mammal, it takes up *T. cruzi* trypomastigotes with the bloodmeal. As the parasites travel via the foregut and stomach to the bug's midgut, the trypomastigotes transform into epimastigotes, which divide and proliferate by binary fission. Some epimastigotes reach the hindgut where they continue to proliferate, and after about 4 weeks, depending on the temperature, they reach about 10^6 cells per insect. Epimastigotes in the hindgut have elongated flagella and are found attached to the rectal cuticle lining via non-specific interaction between the hydrophobic cuticle wax and proteins on the parasite flagellum. Up to 50% of the attached epimastigotes differentiate into infective metacyclic trypomastigotes (referred to in this text as metacyclics), which are excreted with the faeces and urine when the bugs feeds on a mammal (Kollien and Schaub, 2000).

Development in the host. Through contamination of the bite wound or mucous membrane with the bug faeces, metacyclics enter the mammalian host, where they can invade a wide range of nucleated cells. The mechanisms and signal transduction events involved in host cell invasion have been characterised in considerable detail (Burleigh and Andrews, 1995; Tan and Andrews, 2002). The parasite enters the host cell through receptor-mediated attachment. Attachment causes parasite-induced Ca^{2+} signalling, resulting in recruitment of host cell lysosomes to the point of attachment (Burleigh and Andrews, 1995; Caler et al., 1998). The lysosome fuses with the plasma membrane, allowing the parasite entry into the resulting vacoular compartment. Inside the acidified parasitophorous vacuole, the trypomastigote secretes a porin-like molecule, TcTOX, and escapes into the host cell cytoplasm, where it differentiates to the amastigote form (Andrews et al., 1990). This differentiation takes about 3h. Following a lag phase of 35-48 h after differentiation, amastigotes start to divide (Elias et al., 2001b; Vickerman, 1985). Amastigotes proliferate in the cytoplasm by binary fission. At high densities, after 4 to 5 days, amastigotes differentiate to trypomastigotes. The infected cell lyses, releasing parasites into blood and lymph. These cell-derived trypomastigotes then perpetuate the cycle of infection in the host by invasion of new cells, preferentially those of the heart and smooth muscle of the digestive tract. The life-cycle is completed when some trypanosomes are taken up from the peripheral blood by a feeding triatomine (Tyler and Engman, 2001).

24

1.3.2 Dissection of the trypanosome life-cycle – a molecular approach

The basic aspects of the *T. cruzi* life-cycle have been known for nearly a century, but a number of unresolved questions still need to be answered. Surprisingly, it has remained controversial how many distinct life-cycle stages actually exist, and at which point one form becomes committed to a certain developmental pathway (Brack, 1968; Tyler and Engman, 2001).

One of the central questions regarding parasite differentiation is how the cell senses its environment, and how this information is transduced. The review of the literature on trypanosome signal transduction pathways, presented in section 1.5, shows that the mechanisms have remained almost completely obscure. How the elusive signals are processed to result in changes of gene expression is equally in the dark. It will therefore be of importance to functionally analyse stage-specific genes and dissect the mechanisms that regulate their expression. Given the many unusual features of trypanosome gene expression (section 1.6) it is likely that the study of developmental regulation will uncover more mechanisms that are distinct from those operating in other eukaryotes. It is hoped that a detailed understanding of these mechanisms will suggest ways to break the cycle, and block replication and/or transmission of the parasite.

1.3.3 A model case of trypanosome differentiation

At present, the best characterised model for trypanosome differentiation is the differentiation of *T. brucei* bloodstream to procyclic forms (Hendriks et al., 2000; Matthews, 1999; Matthews et al., 2004). Slender bloodstream forms proliferate in the blood of the mammalian host. During increasing parasitemia, the proportions of intermediate forms and non-dividing stumpy bloodstream forms increase. These stumpy forms are apparently pre-adapted for differentiation to the procyclic form (which occurs in the midgut of the tsetse fly).

Differentiation involves the repositioning of the kinetoplast from a posterior location to a position closer to the nucleus (Matthews et al., 1995), a complete change of the

surface glycoproteins (Roditi et al., 1989), re-activation of mitochondrial respiration and a switch from using glucose to using proline as the main energy source (Vickerman, 1985). Bloodstream forms carry on their surface a dense coat of variant surface glycoprotein (VSG), and only one VSG variant out of a large repertoire of VSG genes is expressed. A mechanism of antigenic variation allows some cells to escape the mammalian immune response by switching to the expression of a different VSG variant (Vanhamme and Pays, 1995). During differentiation to the procyclic form, the VSG coat is shed by an active process, involving a zinc metalloprotease and glycosylphosphatidylinositol (GPI) hydrolysis, (Gruszynski et al., 2003). The coat is replaced with two types of procyclin protein (Roditi et al., 1989), which are important for development in the tsetse fly (Ruepp et al., 1997).

The molecular and cell biology of this differentiation step has been dissected in considerable detail (Matthews et al., 2004), thanks to the development of a synchronous *in vitro* differentiation system, (Ziegelbauer et al., 1990) a catalogue of cellular and molecular markers (Hendriks et al., 2000; Matthews et al., 2004), and the availability of genetic manipulation techniques (Clayton, 1999). Analysis of the strict stage-specificity of VSG and procyclin expression has revealed important mechanisms by which trypanosomes control gene expression (Vanhamme and Pays, 1995; section 1.6). Importantly, this research is also beginning to provide molecular explanations for the observed changes in cellular morphology (Matthews et al., 1995). It has also identified molecules that are involved in life-cycle regulation such as protein kinases (Vassella et al., 2001) and novel RNA binding proteins (Hendriks et al., 2001).

The study of *T. cruzi* differentiation is less advanced. Robust *in vitro* differentiation assays have yet to be established, and very few specific molecular markers are known. A major advantage of studying *T. cruzi*, however, is that its life-cycle can be completely reproduced *in vitro* (Miles, 1993). Cell cultures of phagocytic or non-phagocytic cells can be used to grow *T. cruzi* amastigotes and bloodstream trypomastigotes. Shifting supernatants from infected cell cultures from 37°C to 27°C causes the amastigotes to differentiate to epimastigotes, which can be readily grown in axenic culture, using semi-defined or commercially available cell culture medium.

Transformation of epimastigotes to metacyclics occurs spontaneously in stationary phase epimastigote cultures of some strains of *T. cruzi*. While some culture media have been reported to enhance metacyclogenesis (section 1.4.2), the factors that induce this transition stage remain poorly understood, and no method has yet been devised that reliably induces metacyclogenesis in a wide variety of strains.

1.4 *T. CRUZI* METACYCLOGENESIS

The development of metacyclics is central to the *T. cruzi* life-cycle, as this form initiates the infection in the mammalian host. Metacyclics differ from epimastigotes in their morphology, their infectivity, resistance to mammalian serum, and their inability to replicate. The factors that trigger this differentiation, and the changes and mechanisms associated with metacyclogenesis at the molecular level have yet to be determined.

1.4.1 Metacyclogenesis in the insect vector

In the vector, the rate of metacyclogenesis depends on environmental factors such as nutritional state of the vector and temperature, as well as on the *T. cruzi* strain (Kollien and Schaub, 2000). Quantitative studies of metacyclogenesis in the vector (Schaub, 1989) have revealed that the proportion of metacyclics in the vector hindgut does not exceed 50%, and that metacyclics can develop in three different ways. They can develop through differentiation of a sphaeromastigote cell (flagellated spherical cells whose role in the life-cycle is incompletely understood), through differentiation of a slender epimastigote cell, or through unequal division during the growth phase. Attachment of the trypanosome to the insect appears only to be necessary for the differentiation from slender epimastigotes, but not for the others. A complete account of metacyclogenesis has to explain the role of these various morphologically distinct cells, as well as the persistence of non-dividing epimastigotes (Brack, 1968; Vickerman, 1985).

27

1.4.2 Factors that affect metacyclogenesis *in vitro*

Over the past three decades, attempts have been made to establish robust *in vitro* systems that cause epimastigotes to differentiate to metacyclics, so that a system in which factors that influence metacyclogenesis and the associated physiological changes in the parasite can be studied experimentally. There are three commonly used methods to obtain metacyclics: purification of spontaneously differentiated cells from old epimastigote cultures using a DEAE-column (Alvarenga and Brener, 1979), transfer of epimastigotes to Grace's Insect Medium (Sullivan, 1982), or transfer to a minimal medium supplemented with amino acids (Contreras et al., 1985b). Unfortunately, none of these methods is universally applicable, and the factors that control metacyclogenesis have remained elusive.

Grace's Insect Medium. A simple procedure for *in vitro* metacyclogenesis of *T. cruzi*, using a commercially available medium was developed by Sullivan (1982): Transfer of epimastigotes of the Brazilian strain from liver infusion tryptone (LIT) medium (pH 7.2) to Grace's Insect Tissue Culture Medium (pH 6.6; a defined medium) yielded an average of 72% metacyclics after 5 days. A variation in the level of transformation achieved by other strains or isolates of *T. cruzi* was noted. Ucros (1983) confirmed that Grace's medium supported metacyclogenesis in the Peru, Y and CL strains. However, contrary to the findings of Sullivan and others, maximal transformation was observed at pH 9.0.

Triatomine artificial urine (TAU). A chemically defined minimal medium capable of supporting *T. cruzi* metacyclogenesis was developed by Contreras et al. (1985b). TAU is a simple salt solution based on the composition of triatomine urine. When epimastigotes are transferred to TAU supplemented with L-proline (TAUP), the cells attach to the walls of the culture flask, and within 3 to 5 days transform to metacyclics, which are released into the medium. The best results were obtained with the Dm28c strain (87 ± 4% metacyclics within 3 days). Initial differentiation rates were considerably less with other strains, but in a second cycle of incubation in TAUP, following treatment of differentiated cells with fresh guinea pig serum (thereby

28

eliminating remaining epimastigotes), all parasite strains showed differentiation rates of 60 to 90%.

Detailed studies in other laboratories have confirmed that TAU supplemented with a carbon source can induce metacyclogenesis. However, three important differences to the above findings have been noted. First, attachment appeared not to be required for differentiation. Second, the yield of metacyclics is usually 30-50% rather than 80-90%. Third, TAU supplemented with glucose can also induce differentiation, albeit at a lower rate than L-proline (Homsy et al., 1989). Using TAU, the effects and mechanisms of action of a number of factors that potentially affect metacyclogenesis have been studied in great detail in a variety of *T. cruzi* strains. A summary of these findings is presented below.

Energy metabolism. The highest transformation rates are invariably obtained when TAU is supplemented with L-proline, but metacyclogenesis is also induced in the presence of D-glucose, D-fructose, L-glutamate, L-glutamine, L-asparagine, L-proline + sodium acetate, L-glutamine + L-hydroxyproline, D-fructose + L-hydroxy-proline. However, major differences between *T. cruzi* strains have been observed (Krassner et al., 1990). The tricarbolic acid cycle intermediates oxaloacetate, iso-citrate and α-ketoglutarate do not induce transformation on their own, and inhibit proline-induced transformation (Homsy et al., 1989). It was also noted that the amino acids capable of inducing metacyclogenesis can be metabolised by eukaryotic cells to pyrroline-5-carboxylate (P5C), the first product of proline catabolysm, whereas leucine and isoleucine, which both inhibit P5C dehydrogenase also inhibited proline and glutamate induced transformation. Results of further studies using inhibitors and intermediates of proline metabolism are consistent with a proline-P5C cycle induced cascade system for cell activation. However, ornithine and arginine, which are common precursors for proline did not stimulate transformation, and hexoses, which are not readily metabolised to P5C, induced metacyclogenesis suggesting that alternative explanations must be considered (Krassner et al., 1990).

The importance of energy metabolism in metacyclogenesis was recently underlined, when it was shown that reduced levels of glucose trigger flagellar elongation in

proliferating epimastigotes, and subsequent attachment to a hydrophobic surface (Tyler and Engman, 2000). The authors propose the following model for differentiation: attachment generates the signal to differentiate. Differentiation itself requires the remaining glucose to be above a threshold. If the glucose concentration is below the threshold, cells cannot differentiate and they die (Tyler and Engman, 2001). Further studies are required to validate this model, and it will be interesting to determine how attachment generates a signal to differentiate.

A potential role for the reservosome in energy metabolism and metacyclogenesis. The studies discussed above suggest that depletion of glucose combined with heightened tricarbolic acid cycle activity is concomitant with transformation (Krassner et al., 1990). The growth of epimastigotes in culture after depletion of carbohydrates is most likely supported by catabolic breakdown of amino acids (Urbina, 1994). A possible amino acid source are the proteins stored in reservosomes, (acidic [pH 6.0] compartments located at the posterior end of epimastigote cells, but absent in metacyclics) (de Souza et al., 2000). A role for reservosomes in metacyclogenesis has been proposed (Soares, 1999), and is supported by the following studies which used transfection-based approaches to study reservosome-associated proteins.

Reservosomes contain cruzipain, the major cysteine proteinase of *T. cruzi*. Cruzipain enzyme levels are developmentally regulated (Tomas and Kelly, 1996). Enzyme activity is highest in epimastigote cultures, and the activity increases as epimastigotes reach the stationary phase of growth (Tomas et al., 1997). Interestingly, over-expression of cruzipain leads to a 3-4 fold increase in the number of metacyclics produced by the Silvio X10/6 strain in stationary phase cultures (Tomas et al., 1997). Conversely, metacyclogenesis is blocked in the presence of cysteine-protease inhibitors (Bonaldo et al., 1991).

A second reservosome-associated protein whose involvement in metacyclogenesis has been demonstrated is Tc52. Tc52 plays a role in the regulation of the intracellular thiol-disulphide redox balance by reducing glutathione disulphide and it appears to be localised in the reservosome. Tc52 expression is developmentally regulated and the highest levels were observed in stationary phase epimastigote cultures (Ouaissi et

al., 1995). Monoallelic disruption of the *Tc52* gene significantly reduced development of metacyclics. Overexpression of Tc52 on the other hand did not enhance metacyclogenesis (Allaoui et al., 1999).

Calcium and other ions. In view of the role played by K^+ and Ca^{2+} ions in many signal transduction pathways, the action of inorganic ions was investigated in TAU. It was shown that calcium, potassium, phosphate and carbonate ions stimulate, while chloride ions inhibit metacyclogenesis in a concentration and strain dependent manner (Krassner et al., 1991). External Ca^{2+} ions were not required for transformation and, surprisingly, the calcium chelator EGTA in calcium free medium stimulated transformation. High external $[Ca^{2+}]$ inhibited transformation, but this inhibition was reversed when the divalent cation ionophore A23187 was added (Krassner et al., 1993). Induction of metacyclogenesis in the CL strain was not accompanied by changes in cytosolic $[Ca^{2+}]$ and thus no supporting evidence for participation of a protein kinase C (PKC) mediated phosphoinositide cascade in metacyclogenesis was found (Krassner et al., 1993). Thus, the importance of Ca^{2+} signalling in metacyclogenesis is unclear at present.

Cyclic AMP. An effect of cyclic AMP (5'-3'-cyclic adenosine monophosphate; cAMP) on metacyclogenesis in TAU was suggested by the frequently cited study of Gonzales-Perdomo et al. (1988). Under conditions which induce metacyclogenesis, addition of membrane permeable cAMP analogues increased the yield of metacyclics of the Dm28c strain in TAUP during the first two days, but did not affect the final yield after 4 days. Activators of mammalian adenylyl cyclases (cholera toxin, epinephrine) slightly stimulated the kinetics of metacyclogenesis compared to controls in TAUP (but not to the same extent as cAMP analogues). Compounds 48/80 and R24571, which both inhibit phosphodiesterase activation by calmodulin, suppressed metacyclogenesis in TAUP. When proline was replaced with glucose (TAUG), conditions which on their own did not support metacyclogenesis, differentiation was induced when cAMP analogues were added. Similarly, a non-hydrolysable GTP analogue, pNHppG, and adenosine supported metacyclogenesis even in the absence of proline. Based on these results, the existence of heterotrimeric G-proteins in-

31

volved in the activation of *T. cruzi* adenylyl cyclase and a function for cAMP in differentiation were proposed.

More recent reports on the characterisation of the components of the trypanosomatid cAMP signalling pathway (section 1.5.2) have cast serious doubts on the interpretation of the results of this and other studies, which rely on the use of substances that are known to inhibit components of the mammalian cAMP signalling pathway. In short, every component of the cAMP signalling pathway that has been investigated in trypanosomes has been found to be significantly different from its mammalian counterpart (Bieger and Essen, 2001; Naula et al., 2001; Shalaby et al., 2001; Taylor et al., 1999; Zoraghi et al., 2001). Heterotrimeric G-proteins, which regulate the activity of mammalian adenylyl cyclases, are completely absent from the trypanosome genomes. A near complete catalogue of *T. cruzi* genes, as well as the tools for genetic analysis, are now available to re-address the question of whether cAMP signalling plays a role in metacyclogenesis.

An exogenous factor that induces metacyclogenesis? Another controversial series of reports suggested that a factor present in the intestine of recently fed triatomine bugs induces metacyclogenesis by interacting with an epimastigote membrane receptor (de Isola et al., 1986; de Isola et al., 1981). This factor was later identified with a proteolytic cleavage product of α^D-globin from chicken blood, termed GDF (Fraidenraich et al., 1993). Furthermore it was reported that GDF activates the *T. cruzi* adenylyl cyclase and stimulates metacyclogenesis *in vitro* (Fraidenraich et al., 1993) and *in vivo* (Garcia et al., 1995). Attempts to replicate these findings independently have been unsuccessful.

In summary, there are many contradicting reports in the literature as to the conditions that induce metacyclogenesis in *T. cruzi*. The emerging consensus is that stress conditions, including starvation, favour metacyclogenesis, but that sufficient energy sources must be available to sustain differentiation. Clearly, studies with individual strains or isolates of *T. cruzi* may yield results that are not universally applicable. A description of events associated with metacyclogenesis in molecular terms, and experimental manipulation using reverse genetic approaches should allow us to

32

define more clearly which factors are essential for induction of metacyclogenesis, and which are merely contingent. Advances that have been made in both fields are discussed in sections 1.4.3 and 1.7.

1.4.3 Molecular events associated with metacyclogenesis

Well characterised molecular markers for specific *T. cruzi* life-cycle stages are still scarce. Qualitative and quantitative differences between proteins expressed in the various life-cycle stages have been demonstrated on 1D and 2D protein gels and Western blots (Contreras et al., 1985a; Lanar and Manning, 1984; Rangel-Aldao et al., 1986). However the identity of most of these molecules remains to be determined. With the technological advances that have been made since those early studies were conducted 20 years ago, and the advent of high throughput functional genomics approaches (section 1.7.2), it is now feasible to identify such stage-specific gene products on a large scale. A brief overview of the current knowledge of molecular events associated with metacyclogenesis is presented below.

Expression of metacyclic-specific surface glycoproteins gp82 and gp90. The best characterised stage-specific *T. cruzi* genes code for surface proteins, including the epimastigote-specific small mucin genes *TcSMUG* (Di Noia et al., 2000), the amastigote-specific *amastin* (Teixeira et al., 1994) and proteins of the *gp85/sialidase* multi-gene family expressed in cell-derived trypomastigotes. The latter belong to the *trans-sialidase* gene superfamily (Cross and Takle, 1993; Schenkman et al., 1994). Two proteins that belong to this family, gp82 and gp90, have been identified as the major surface glycoproteins specific to metacyclics (Araya et al., 1994; Ruiz et al., 1998). As yet, gp82 represents the only well characterised protein marker for metacyclics in a wide variety of *T. cruzi* strains. Expression of gp90 on the other hand is strain-specific (Beard et al., 1988; Ruiz et al., 1998). The genes coding for gp82 and gp90 have been cloned and shown to belong to multi-gene families, dispersed throughout the *T. cruzi* genome. The protein sequences are 40-60% identical to those of the *gp85/sialidase* family. Steady state levels of *gp82* and *gp90* mRNA are developmentally regulated. Transcript levels are high in metacyclics, very low in

epimastigotes, and absent from other life-cycle stages (Araya et al., 1994; Carmo et al., 1999; Franco et al., 1993).

Gp82 has been implicated in the process of host cell invasion and in the initiation of signalling cascades, both in the parasite and in the target cell. It functions as a parasite surface receptor that binds to an unidentified target cell molecule. This initiates a signalling cascade in the parasite cell, that involves tyrosine phosphorylation of a 175 kDa protein, and leads to invasion of the target cell (Favoreto et al., 1998). The mechanism of transduction from the GPI-anchored gp82 into the cell is unclear. Binding of gp82, and to a lesser extent gp90 induces Ca^{2+} mobilisation in the target cell, and target cell extracts trigger Ca^{2+} mobilisation in metacyclics. Parasite Ca^{2+} mobilisation can be induced by monoclonal antibodies against gp82, but not by antibodies against gp90 (Ruiz et al., 1998). Interestingly, analysis of ten different strains of *T. cruzi* revealed an inverse correlation between the infectivity of a *T. cruzi* strain and the level of gp90 expression. While gp90 was found to be one of the major surface components in the poorly invasive G strain, the antibody failed to detect the protein in the highly invasive CL and Y strains (Ruiz et al., 1998).

Changes in chromatin. A general decrease in transcription rates (Elias et al., 2001b), and phosphorylation of histone H1 (Marques Porto et al., 2002) occurs concomitantly with extensive nuclear reorganisation when replicating forms differentiate to the non-replicating and infective forms (including metacyclics). This is detailed in section 1.2.2.

Differential gene expression during metacyclogenesis. Prior to *in vitro* differentiation in TAU, epimastigotes of the Dm28c strain adhere to the culture flask, and these epimastigotes express specific surface antigens (Bonaldo et al., 1988). Recent discoveries may provide the first clue as to the molecules and control mechanisms involved in this process. To identify genes expressed transiently during metacyclogenesis, a method has been devised recently (Krieger et al., 1999; Krieger and Goldenberg, 1998), called representation of differential expression (RDE). RDE is based on PCR amplification of cDNA sequences unique to a cell population (tester) after subtractive hybridisation with DNA sequences of a related cell population

34

(driver). Using as driver the polysomal RNA fractions from replicating epimastigotes, and as tester the polysomal RNA fractions from epimastigotes induced to differentiate in TAU, several transcripts have been shown to be specific to adhered (and by implication, differentiating) cells. Interestingly, these studies suggest that mRNA mobilisation to the polysomes may be one of the post-transcriptional mechanisms for regulating gene expression in trypanosomes.

The RDE method has so far led to the identification of two types of surface proteins thought to be involved in the differentiation process. This will be discussed in the following two sections. In addition, several novel upregulated transcripts of unknown function await further characterisation.

Chitin-binding-like protein (CBLP). During metacyclogenesis in TAU, transcripts were shown to be mobilised to the polysomal fraction that code for a family of small cysteine-rich proteins with sequence similarity to the chitin-binding domain of wheat germ agglutinin. Maximal transcript levels were detected 24 hours after initiation of differentiation. Consistent with this, the levels of a 7 kDa protein detected by a polyclonal antiserum raised against recombinant CBLP1, are highest in adhered epimastigotes after 24 hours (Dallagiovanna et al., 2001).

Metacyclogenin. A second putative surface protein specifically expressed in differentiating epimastigotes is metacyclogenin (Avila et al., 2001). Metacyclogenin mRNA is associated with the polysomes of 24 hours adhered epimastigotes. The *T. cruzi* genome contains at least 3 copies of the gene arranged in a tandem repeat with an 'associated' gene (*AG*) and a cytosolic tryparedoxin peroxidase gene (*CPX*). The *AG* gene shows an expression pattern similar to *metacyclogenin* (Avila et al., 2001). The 13 kDa metacyclogenin protein is detected exclusively in adhered epimastigotes and it appears to be associated with the parasite cell membrane, as determined by immunocytochemical electron microscopy. The function of metacyclogenin is unknown.

Using reverse genetic approaches it should be feasible to determine the role that these proteins play in differentiation and how commitment to differentiation and attachment of the parasite are linked.

1.5 DISSECTING SIGNAL TRANSDUCTION
IN TRYPANOSOMES

Throughout their life-cycle, trypanosomes are confronted with environments that differ dramatically in their physico-chemical and biological properties (e.g. temperature, pH, osmolarity and Ca^{2+} content, nutrient supply, haemolytic factors, agglutinins, digestive enzymes, components of cellular and immune defence systems from a variety of species). The trypanosome responds to these changes with equally dramatic changes in its cellular architecture and metabolism, and cycles between proliferating and non-proliferating, infective and non-infective, motile and non-motile stages (Vickerman, 1985). The parasite must therefore have mechanisms by which external signals are transduced and integrated to give rise to altered gene expression at the appropriate time. For the parasite cell these adaptive responses are essential in order to survive and proliferate. To the parasitologist they suggest the presence of essential pathways and potential targets for therapeutic interference. It is therefore a major task for the future to identify the components of these signal transduction pathways and discover how changes in gene expression are regulated.

Mechanisms by which extracellular signals are detected and transduced by eukaryotic cells to give rise to altered patterns of gene expression have been studied in great detail. Generally, an extracellular signal (including peptide ligands, ions, electric currents, certain wavelengths of light) alters the conformation of a membrane bound receptor, some of which function as ion-channels or have intrinsic enzymatic activity. They can be associated with intracellular kinases directly, or indirectly via heterotrimeric G-proteins. Transduction of the signal via the receptor protein causes synthesis of second messenger molecules, including cyclic AMP (cAMP), cyclic GMP (cGMP), inositol 1,4,5-triphosphate (IP$_3$), 1,2-diacylglycerol (DAG), the release of Ca^{2+} from intracellular stores, and/or cascades of protein phosphorylation and dephosphorylation. This leads to the modulation of the activity of relevant

enzymes and transcription factors by allosteric regulation. Positive and negative feedback loops and crosstalk between different pathways provides the cell with complex mechanisms to adapt gene expression and metabolism to environmental conditions and developmental requirements (Lodish et al., 1995).

While some components of the cellular signal transduction pathways known from higher eukaryotes are conserved in trypanosomatids (Parsons and Ruben, 2000; Ruben et al., 2002), their exact mechanisms and functions have remained mostly unexplored. Since the predominant mechanism for control of gene expression in kinetoplastids appears to be post-transcriptional (Clayton, 2002; Vanhamme and Pays, 1995), it is expected that signal transduction pathways result in changed RNA turnover or processing rather than changes in transcription (Parsons and Ruben, 2000).

1.5.1 Protein kinases and phosphatases

Reversible protein phosphorylation is a mechanism common to many signal transduction pathways. A considerable number of serine and threonine kinases (but no tyrosine kinases) from trypanosomatids have been purified and biochemically characterised (Boshart and Mottram, 1997). The high degree of sequence conservation found in kinases and phosphatases has allowed isolation of relevant genes, including mitogen activated protein (MAP) kinase homologues, cdc2-related kinases, protein kinase A (PKA), and protein phosphatases type 1, type 2A and 2C. Studies from *Leishmania* (Dell and Engel, 1994) and *T. brucei* (Parsons et al., 1993), using metabolic labelling, specific antibodies or *in vitro* kinase assays, have provided evidence that kinetoplastids have a full set of serine/threonine and tyrosine protein kinases and phosphatases. However, there is no firm evidence for receptor protein kinases and phosphatases, which in multi-cellular organisms represent an important group of signalling molecules. Completion of the trypanosomatid genome projects now allows the assembly of a complete catalogue of the relevant genes.

Several protein kinases and phosphatases show distinct cell-cycle (Gale et al., 1994) or life-cycle specific activity (Aboagye-Kwarteng et al., 1991; Bakalara et al., 1995;

37

Dell and Engel, 1994; Parsons et al., 1993; Parsons et al., 1991). Some kinases have been shown to play a role in parasite development, but factors regulating their activity, and their downstream targets are still largely unknown. For example, a novel zinc finger kinase, ZFK, has recently been shown to play a role in *T. brucei* differentiation. Deletion of ZFK resulted in cells that showed an increased rate of differentiation from slender to stumpy bloodstream forms (Vassella et al., 2001).

The only well characterised *T. cruzi* trypomastigote specific phosphoprotein to date is histone H1, which is phosphorylated in the non-replicating trypomastigote stages, but not in replicating epimastigotes (Marques Porto et al., 2002). The kinase activities and phosphorylation profiles associated with *T. cruzi* metacyclogenesis are unknown.

1.5.2 cAMP signalling

The second messenger cAMP plays a major role in controlling gene expression and metabolism in prokaryotes and eukaryotes. Its functions range from simple environmental responses in lower eukaryotes, such as the photo-avoidance response in *Euglena* (Iseki et al., 2002), and fruiting body formation in the slime mould *Dictyostelium discoideum* in response to starvation (Weeks, 2000), to more complex functions, including the response to ethanol intoxication in *Drosophila* (Moore et al., 1998) and odour perception in humans (Zufall et al., 1994). Its proposed role in *T. cruzi* metacyclogenesis remains controversial (section 1.4.2).

The paradigm of cAMP signalling in eukaryotes. In response to an extracellular signal, cAMP is synthesised by the enzyme adenylyl cyclase (AC). Most AC isoforms of higher eukaryotes are integral membrane proteins with two intracellular pseudo-symmetrical domains and two hydrophobic domains each consisting of six transmembrane helices. The two catalytic domains dimerise to form the active enzyme (Hurley, 1998). The activity of mammalian ACs is controlled by a G-protein-coupled receptor. Binding of an external ligand leads to the dissociation of the G_α from the $G_{\beta\gamma}$ subunits in the associated G-protein and subsequent activation or inhibition of the AC (Bentley and Beavo, 1992; Hurley, 1998; Levin and Reed, 1995).

38

The most extensively studied downstream target for cAMP is cAMP-dependent protein kinase (PKA), a tetrameric protein consisting of two catalytic and two regulatory subunits. Binding of cAMP to the regulatory subunits of PKA releases the catalytic subunits, which then phosphorylate a diverse range of proteins including transcription factors such as the cAMP response element binding protein CREB (Fimia and Sassone-Corsi, 2001). Other downstream targets of cAMP include cyclic nucleotide gated ion channels, drug efflux channels, and the recently discovered cAMP binding protein Epac, which acts as a guanine-nucleotide-exchange factor for Rap1 (de Rooij et al., 1998). The cAMP signal is terminated by the activity of cAMP-dependent phosphodiesterases (PDE) which break down cAMP to AMP by hydrolysis of the 3'-ester bond. The differential sensitivity of various PDEs to specific inhibitors has made them important drug targets.

cAMP signalling in trypanosomes. Production of cAMP was one of the first biochemical reactions described in *T. brucei* (Strickler and Patton, 1975). Meanwhile, abundant evidence suggests that cAMP plays a role in growth and differentiation in trypanosomatids (Naula and Seebeck, 2000; Seebeck et al., 2001). The differentiation of proliferating long slender bloodstream forms of *T. brucei* to the non proliferating short stumpy forms, which are infective to tsetse flies, is thought to be mediated by a density sensing mechanism, in which release of a 'stumpy inducing factor' (SIF) by trypanosomes acts via the cAMP pathway to induce cell cycle arrest and subsequent differentiation (Vassella et al., 1997). Distinct peaks of AC activity were also observed during the *in vitro* differentiation of *T. brucei* bloodstream to procyclic (insect-stage) forms (Rolin et al., 1993). In *T. cruzi*, differentiation of metacyclics to amastigotes appears to be initiated when proteolytic cleavage fragments of fibronectin activate the trypanosomal AC (Ouaissi et al., 1992). Tissue culture-derived *T. cruzi* trypomastigotes contain 4.3 times more cAMP than epimastigotes, and homogenates of trypomastigotes have a 2.6 fold higher cAMP binding activity than epimastigotes (Rangel-Aldao et al., 1987). Evidence suggesting a role for cAMP signalling in the differentiation of *T. cruzi* epimastigotes to metacyclics was discussed in more detail in section 1.4.2.

39

Trypanosomal adenylyl cyclases. In trypanosomatids ACs exist as multi-gene families. *T. cruzi* contains about 30 genes (Taylor et al., 1999), and *T. brucei* has several hundred different AC genes (Alexandre et al., 1996; Naula et al., 2001; Seebeck et al., 2001). All trypanosomatid ACs analysed so far are structurally distinct from mammalian-type ACs. They consist of a large variable, presumably extracellular N-terminal domain, a single transmembrane segment and a conserved C-terminal catalytic domain, which is active as a dimer of two AC molecules (Naula et al., 2001; Taylor et al., 1999). Some isoforms of *T. brucei* and *T. cruzi* AC appear to be localised to the flagellum (Paindavoine et al.; 1992, Martin Taylor, unpublished).

The structure of trypanosomal ACs is reminiscent of receptor form mammalian guanylyl cyclases which are activated directly by external ligands. It has therefore been suggested that trypanosomatid ACs may act directly as receptors and, unlike mammalian ACs, trypanosomal ACs are not regulated by G-protein coupled receptors. Putative AC activating ligands have been reported (Fraidenraich et al., 1993; Ouaissi et al., 1992) but the linkage between ligand binding and cAMP production has not been unequivocally demonstrated. Consistent with the proposed receptor function of trypanosomal ACs, heterotrimeric G protein subunit genes homologues have not been identified by any of the three trypanosomatid genome projects. Sequence analysis of *T. cruzi* ACs (Taylor et al., 1999) and the recently solved crystal structure for two *T. brucei* ACs (Bieger and Essen, 2001) clearly show that the regions known to be involved in binding to the $G_{\beta\gamma}$ subunit of heterotrimeric G-proteins in the mammalian ACs, as well as the binding site for the AC activator forskolin, are not present.

Phosphodiesterases. Recently the cloning and biochemical characterisation of different *T. brucei* and *T. cruzi* cyclic AMP dependent phosphodiesterases has been reported (D'Angelo et al., 2004; Gong et al., 2001; Kunz et al., 2004; Zoraghi et al., 2001). Inhibitor studies revealed that these trypanosomal phosphodiesterases display a pharmacology quite distinct from previously characterised PDEs. Importantly, the enzyme activity was not inhibited by 3-isobutyl-1-methylxanthine (IBMX), which is commonly used to inhibit mammalian PDEs. Interestingly, substances that inhibited the *T. brucei* protein TbPDE2A *in vitro*, such as trequinsin, dipyridamole, sildenafil,

and ethaverine inhibited growth of *T. brucei* in culture (Zoraghi et al., 2001). Deletion of the gene for another *T. brucei* phosphodiesterase, *TbPDE1*, has not resulted in a discernible phenotype (Gong et al., 2001), and the function of the different PDE isoforms in *T. brucei* remains unclear. The subcellular location of a *T. cruzi* phosphodiesterase, TcPDE1, was investigated by subcellular fractionation of epimastigotes and immunofluorescence microscopy, using an antibody raised against the recombinant protein. TcPDE1 was found to be membrane associated and concentrated on the flagellum of epimastigotes (D'Angelo et al., 2004).

Downstream effectors of cAMP. The mechanism by which cAMP exerts its effects in trypanosomes is unknown. It has consistently proven difficult to correlate cAMP binding activities with kinase activities. One line of evidence suggested that *T. cruzi* contains a single cAMP binding protein, CARPT, which is not associated with phosphotransferase activity (Rangel-Aldao et al., 1985; Rangel-Aldao et al., 1983). A short cDNA fragment of CARPT has been cloned (GenBank:S79626), but it shows no similarity to any other known cAMP binding protein and further characterisation has not been reported. A PKA-like kinase activity which is specifically activated by cAMP was later purified from *T. cruzi* epimastigotes, and the purified holoenzyme is recognised by antibodies against bovine PKA (Ochatt et al., 1993; Ulloa et al., 1988). The *T. cruzi* PKA is the only kinase activity from trypanosomatids described to date, which was shown to be stimulated by cAMP. A recently characterised PKA regulatory subunit from *T. brucei* shares structural similarity with the mammalian homologues, but, unexpectedly, the *T. brucei* PKA is activated by cGMP but not by cAMP (Shalaby et al., 2001). The physiological relevance of this cGMP activation is unclear, as there is no evidence for guanylate cyclase genes or enzyme activity in trypanosomes. Thus, a clear picture of the downstream effectors of cAMP in trypanosomes has not yet emerged.

cAMP induction of Tc26. A single cAMP-inducible RNA transcript termed *Tc26* has been identified in *T. cruzi* (Heath et al., 1990). *Tc26* transcripts are specific to metacyclics, and not found in epimastigotes, unless membrane permeable cAMP analogues (dibutyryl cAMP or 8-BrcAMP) are added to the medium. *Tc26* was later found to be part of a non-LTR retrotransposon (*L1Tc*), widely dispersed in the

41

T. cruzi genome (Allen and Kelly, 2001), and it appears unlikely that it plays a direct role in the cAMP response.

1.5.3 Ca^{2+} signalling

Eukaryotic cells maintain a low Ca^{2+} concentration in the cytoplasm. This is mediated by Ca^{2+} transport out of the cell and sequestration in internal compartments, the most important of which are the mitochondrion and ER. A signal is generated by a rise in cytosolic $[Ca^{2+}]$, and activation of specific Ca^{2+} -binding proteins. In trypanosomes similar mechanisms are thought to function, and several of the Ca^{2+} regulators and effectors such as Ca^{2+} pumps and Ca^{2+} -binding proteins have been identified (Ruben et al., 2002). In addition to Ca^{2+} sequestration in the mitochondrion, kinetoplastids and apicomplexans store Ca^{2+} in a specialised acidic compartment, the acidocalcisome (de Souza et al., 2000). The exact role of the acidocalcisome is unclear. Ca^{2+} binding proteins in the ER (e.g. BiP, calreticulin) have been shown to have a function in translation, protein folding and protein trafficking. A wide variety of Ca^{2+} binding proteins containing the EF-hand motif have been identified and cloned from *T. brucei* and *T. cruzi*. These include calmodulin (Chung and Swindle, 1990), which modulates the activity of target proteins in response to Ca^{2+}, and EF-hand proteins associated with the flagellum, such as the *T. cruzi* flagellar calcium binding protein FCaBP (Maldonado et al., 1997), whose function is unknown.

In higher eukaryotes the inositol-lipid signalling pathway is activated by interaction of a ligand with a G-protein coupled receptor. Subsequent activation of phospholipase C (PLC) leads to cleavage of phosphatidyl inositol 4,5-biphosphate (PIP_2) to IP_3 and DAG. IP_3 causes Ca^{2+} release from ER into the cytoplasm through binding to specific channels in the membrane of the ER. DAG and elevated cytosolic Ca^{2+} in turn activate PKC (Nishizuka, 1992). Several components of this pathway are found in trypanosomes. A PLC orthologue cloned from *T. cruzi* shows some unique features, such as lack of an apparent pleckstrin homology (PH) domain. Its expression is upregulated under conditions that stimulate the *in vitro* differentiation of *T. cruzi* trypomastigotes to amastigotes (Furuya et al., 2000). PKC-like activities have been

reported (Gomez et al., 1989; Keith et al., 1990), but no gene from the PKC family has been identified in the kinetoplastids.

Ca^{2+} signalling plays an important part in the invasion of mammalian cells by trypo-mastigotes (Burleigh and Andrews, 1998; Tan and Andrews, 2002). During invasion, Ca^{2+} dependent signalling cascades, are triggered both in the parasite and in the host cell. These are mediated by the metacyclics specific surface glycoprotein gp82 (section 1.4.3) and depend on the activity of the *T. cruzi* serine hydrolase oligopeptidase B (Burleigh and Andrews, 1995; Caler et al., 1998).

Although some signalling effectors have been identified in trypanosomes, their up-stream receptors and downstream targets remain mostly unclear. The discovery that parasite enzymes often do not respond to inhibitors or activators in the same way as their counterparts in higher eukaryotes (Zoraghi et al., 2001) bodes well for potential therapeutic interference. However, those early studies on signal transduction in trypanosomes that were based on the assumption that inhibitors and activators of mammalian enzymes would equally affect the trypanosome enzyme (for example Gonzales-Perdomo et al., 1988; Heath et al., 1990) must be reviewed and probably revised. One of the challenges of post-genomics research will be to link individual signalling molecules, and to determine their function in the parasite life-cycle.

1.6 UNUSUAL MECHANISMS OF GENE EXPRESSION IN TRYPANOSOMES

Most organisms that have been studied, from *E. coli* to humans, regulate the expression of genes predominantly by controlling transcription initiation. As detailed in every molecular biology textbook (Lodish et al., 1995), transcription factors bind to DNA elements (promoters) upstream of individual protein coding genes which leads to recruitment of RNA polymerase II and assembly of the transcription initiation complex. All evidence to date indicates that this is not the case in trypanosomatids. In these organisms, constitutive transcription is the norm and the expression of individual genes is controlled at various stages *post* transcription (Clayton, 2002; Vanhamme and Pays, 1995). The mechanisms by which trypanosomes regulate

43

stage-specific gene expression patterns are thus distinct from those in most other organisms that have been studied.

1.6.1 Organisation of genes in polycistronic transcription units

Gene organisation and expression in kinetoplastids differs that of higher eukaryotes in several respects. Genes on kinetoplastid chromosomes are arranged in large directional clusters (Andersson et al., 1998; Hall et al., 2003). Typically, genes lack introns, with the first reported exception to this rule being the genes for poly(A) polymerase (Mair et al., 2000). Detailed analysis of the transcripts produced from a number of genes, including the calmodulin (Tschudi and Ullu, 1988) and phosphoglycerate kinase loci (Gibson et al., 1988) revealed the presence of large RNA transcripts that contain the coding sequence of several neighbouring genes as well as intergenic sequences not found in the mature mRNA. This research established that genes are transcribed into large polycistronic pre-mRNA, from which mature monocistronic mRNAs are subsequently produced by RNA processing (see below). The selective advantage of polycistronic transcription is not clear. There is no evidence to date to suggest that functionally linked genes are routinely organised in common transcription units. However, given the large proportion of genes of unknown function in the kinetoplastid genomes (Hall et al., 2003), the existence of such co-regulated units cannot yet be ruled out.

1.6.2 RNA polymerases, promoters and transcription

Genes for the core subunits of trypanosomal RNA polymerase I (pol I) (Smith et al., 1989), RNA polymerase II (pol II) (Evers et al., 1989) and RNA polymerase III (pol III) (Kock et al., 1988; Smith et al., 1989) have been cloned, biochemically characterised and classified according to their differential sensitivity to the fungal toxin α-amanitin (Clayton, 2002). The trypanosome pol II enzyme is structurally different form that of other eukaryotes (Evers et al., 1989). Specifically, it lacks the heptapeptide repeats in the C-terminal domain, which in higher eukaryotes is critical for the coupling of transcription, splicing and poly-adenylation.

44

Most protein coding genes in trypanosomes are transcribed by pol II. However, pol II promoters have remained elusive, and the nature of the pol II complex remains to be characterised. Transfection experiments in *T. cruzi*, using episomal vectors suggest that pol II can initiate transcription in the absence of a specific promoter (Teixeira, 1998; Teixeira et al., 1995). Transcription from chromosomes appears to be predominantly unidirectional, however, and it is thought that each contains just a few transcription initiation sites. Nuclear run-on experiments used to study transcription on *L. major* chromosomes I and III indicate that transcription initiation occurs in the strand switch region, between gene clusters (Martinez-Calvillo et al., 2004; Martinez-Calvillo et al., 2003). The complete sequence of a number of chromosomes with and without such strand switch regions are available (or will be shortly) as a result of the trypanosomatid genome sequencing projects. This provides ample material to identify the nature of such transcription initiation sites.

Kinetoplastid ribosomal RNA genes, like those in other eukaryotes, are transcribed in the nucleolus by pol I. An unusual feature of trypanosome gene expression is the fact that pol I mediated transcription of protein coding genes can result in functional mRNA (Zomerdijk et al., 1991a). In *T. brucei*, the VSG and procyclin genes are transcribed by pol I. The respective promoters have been characterised in detail (Janz and Clayton, 1994; Zomerdijk et al., 1991a; Zomerdijk et al., 1991b; Zomerdijk et al., 1990), and they resemble other eukaryotic rRNA promoters. Even though they share little sequence similarity, the procyclin and rRNA promoter have been found to be functionally interchangeable (Janz and Clayton, 1994). The VSG and procyclin promoters represent the only promoters for protein coding genes of trypanosomes identified to date (Clayton, 2002; Vanhamme and Pays, 1995). Transfection experiments in *T. brucei* have established that pol I mediated transcription can also drive the expression of reporter genes under the control of a rRNA promoter, and that the mRNA is correctly processed. Transcription of these reporter genes was shown to occur in the nucleolus (Rudenko et al., 1991). Interestingly, pol I mediated transcription of VSG occurs outside the nucleolus (Chaves et al., 1998; Navarro and Gull, 2001) in a distinct locus, termed the expression site body (ESB), which is present only in the bloodstream form of the parasite (Navarro and Gull, 2001). Recruitment of a single VSG gene to a privileged site such as the ESB could provide

a mechanism by which monoallelic expression from a large repertoire of VSG genes is achieved (Borst, 2002).

The precursor RNA for the spliced leader sequence (SL, see below) is transcribed by pol II, and the SL genes represent the first instance where a pol II promoter has been characterised in trypanosomatids (Gilinger and Bellofatto, 2001; Gunzl et al., 1997; Nunes et al., 1997). Interestingly, recent evidence suggests that pol II mediated transcription of the SL genes in *T. cruzi* also occurs in a distinct nuclear location, close to the nucleolus (Schenkman et al., 2004). Advances in imaging technology, has had a major impact on the perception of the eukaryotic nucleus. It has become clear that there are distinct compartments within the nucleus, despite the lack of internal membrane boundaries (Dundr and Misteli, 2001; Spector, 2003). The question to what extent nuclear architecture contributes to the control of gene expression is an area of intense investigation (Misteli, 2000). It will therefore be of interest to identify the factors that direct specific genes to distinct subnuclear compartments in trypanosomes, and to determine to what extent the positioning of genes within the nucleus affects their expression.

Small RNAs including tRNAs are transcribed by pol III, and a pol III promoter has been described (Fantoni et al., 1994). Small nucleolar RNAs are transcribed as part of polycistronic pol II transcription units (Clayton, 2002; Uliel et al., 2004).

There are still considerable gaps in our understanding of the steps from the polycistronic pre-mRNA to the translated protein (D'Orso et al., 2003). While trypanosomes share some of the mechanisms known from higher eukaryotes, there are clearly many interesting differences, as outlined below.

1.6.3 Trans-splicing and polyadenylation

The primary polycistronic transcript is processed to monocistronic mature mRNAs by RNA processing complexes (Figure 1.4). Unusually, in trypanosomes pol II-mediated transcription and mRNA capping are uncoupled. The 5' end maturation of the mRNA occurs through addition of a common 39 nucleotide spliced leader (SL)

sequence derived from a separate transcript, in a *trans*-splicing reaction. *Trans*-splicing is mechanistically similar to *cis*-splicing in other eukaryotes (Davis, 1996; Nilsen, 1995). Branched intermediates are formed and utilise a 5-GpU dinucleotide splice donor and an ApG dinucleotide 3' splice acceptor site. The SL is added to the 5' UTR sequence and does not affect the reading frame of the gene.

Originally discovered in trypanosomes (Sutton and Boothroyd, 1986), *trans*-splicing was later found to occur also in *Euglena* and in a number of metazoan phyla, including nematodes (e.g. *C. elegans* and *Ascaris*), platyhelminthes (e.g. *Schistosoma mansoni*), cnidaria (*Hydra*) and chordata (*Ciona intestinalis*, a sea squirt) (Davis, 1996; Stover and Steele, 2001; Vandenberghe et al., 2001). Unlike in trypanosomes, only a minority of mRNAs acquire a spliced leader sequence in these organisms. *Trans*-splicing has not been shown to occur in vertebrates. As gene expression has been analysed in only a very small proportion of lower eukaryotes, it is not clear how widespread *trans*-splicing is and what function it serves.
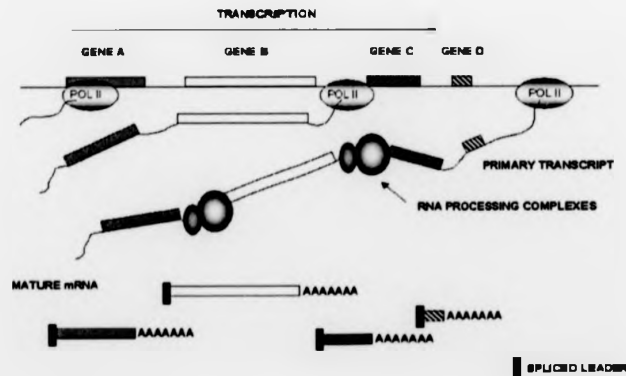
Figure 1.4: Polycistronic transcription in trypanosomes
(from Taylor and Kelly, 2002)

47

In a process coupled to the *trans*-splicing reaction, a 3'-poly-A tail is added to the mRNA (poly-adenylation). In higher eukaryotes, the hexanucleotide sequence AAUAAA near the 3' end of the RNA directs poly-adenylation. Trypanosomes, by contrast, contain no consensus polyadenylation signal. It occurs at a fixed distance upstream of the splice acceptor site and depends on the presence of pyrimidine-rich sequence elements in the intergenic region. (Matthews et al., 1994; Schurch et al., 1994). A poly(A) binding protein has been characterised, which is thought to be involved in mRNA stability and translation (Batista et al., 1994; Hotchkiss et al., 1999).

How the trypanosome mRNA is exported from the nucleus to the cytoplasm, and the degree of regulation involved, is presently unknown.

1.6.4 Differential gene regulation

The steady state levels of mRNA derived from different genes along the same transcription unit can differ radically (Abuin et al., 1999; Gibson et al., 1988; Teixeira et al., 1995). This implies that trypanosome gene expression must be controlled post-transcriptionally. The best-characterised polycistronic transcription units in trypanosomes are those that contain the genes for the *T. brucei* surface proteins VSG and procyclin. A combination of transcriptional and post-transcriptional controls determines the absolute stage-specificity of these genes. This includes differential transcription elongation (Vanhamme et al., 1995), mRNA stability (Furger et al., 1997; Vanhamme et al., 1995) and translation efficiency (Schurch et al., 1997). Even though these genes differ from other protein coding genes in that they are transcribed by pol I and not pol II, their study has uncovered some general mechanisms of trypanosome gene regulation (Vanhamme and Pays, 1995).

A central mechanism in the regulation of kinetoplastid gene expression is differential stability of mRNA molecules (D'Orso et al., 2003). *Cis* regulatory elements that determine the stability of mRNAs are typically localised in the 3' untranslated region (UTR) (Berberof et al., 1995; Blattner and Clayton, 1995; Furger et al., 1997; Nozaki

and Cross, 1995; Teixeira et al., 1995), or less commonly in the 5' UTR (Pasion et al., 1996). Treatment of trypanosomes with protein synthesis inhibitors such as cycloheximide can lead to an increase or decrease of specific transcripts (Abuin et al., 1999; Di Noia et al., 2000; Teixeira et al., 1995). This suggests that labile *trans*-acting protein factors play a role in selectively stabilising or destabilising certain transcripts, presumably by binding to the *cis*-acting element.

Few of the *cis* regulatory elements and even fewer of the *trans* acting factors that regulate gene expression have been characterised. Given the importance of RNA processing in the control of trypanosome gene expression, surprisingly few protein factors that interact with RNA have been functionally characterised (D'Orso et al., 2003). Since many RNA binding proteins can be easily recognised by the presence of conserved domains, that list is expected to grow as researchers continue to mine the trypanosomatid genome databases.

A small number of developmentally regulated RNA binding proteins have been studied in detail. These include the nuclear proteins p34/37 (Zhang et al., 1998; Zhang and Williams, 1997) and the nucleolar protein Nopp44/46 (Das et al., 1998), which have a role in ribosome biogenesis in *T. brucei* (Pitula et al., 2002a; Pitula et al., 2002b). Two RNA binding proteins from *T. brucei* called TbZFP1 and TbZFP2 have recently been demonstrated to play a direct role in regulating differentiation from bloodstream to procyclic forms (Hendriks et al., 2001). A knock down of TbZFP2 by RNAi inhibited differentiation to the procyclic form. Overexpression of TbZFP2 on the other hand resulted in a stage-specific remodelling of the cyto-skeleton, such as normally occurs during procyclic development (Hendriks et al., 2001). Homologues of the *ZFP1* and *ZFP2* genes have recently been cloned from *T. cruzi* (Morking et al., 2004), and it will be extremely interesting to determine their role in *T. cruzi* differentiation.

In *T. cruzi*, mechanistic insight into differential gene expression has been gained from analysis of stage-specific surface proteins.

Amastin/tuzin. Multiple copies of the *amastin* gene, which codes for an amastigote specific surface glycoprotein, are arranged in tandem repeats, alternating with genes of unknown function called *tuzin*. The *amastin/tuzin* gene cluster is equally transcribed in all life-cycle stages, but the steady state RNA levels are differentially regulated. While *amastin* mRNA levels are at least 50 times higher in amastigotes than in other stages (Teixeira et al., 1994), the mRNA levels of the *tuzin* genes are not differentially regulated (Teixeira et al., 1995). Reporter gene assays demonstrated a role for the *amastin* 3' UTR in developmental regulation of mRNA levels (Nozaki and Cross, 1995; Teixeira et al., 1995). Further dissection of the 3'UTR sequence has shown that this process is mediated by a position and orientation dependent element, which specifically binds a 36 kDa protein factor that is exclusively expressed in amastigotes (Coughlin et al., 2000; Teixeira et al., 1995).

Small mucin gene family. A multi-copy family of small mucin genes, called *TcSMUG*, exhibits stage-specific differences in mRNA levels. The highest levels are found in epimastigotes (Di Noia et al., 2000). Two functionally different *cis*-acting elements were identified in the 3' UTR of *TcSMUG* genes in reporter gene assays. A destabilising effect in metacyclics is mediated by a 44 nucleotide long AU-rich element (ARE), which also affects translation efficiency. Interestingly, this element is similar to ARE elements from higher eukaryotes (Di Noia et al., 2000). In addition, a novel G-rich element, termed GRE, selectively stabilises *TcSMUG* in epimastigotes. There is evidence for a protein complex that recognises the GRE element (D'Orso and Frasch, 2001a). The RNA binding protein TcUBP-1, which specifically recognises AU-rich sequences, is involved in stage-regulated expression of *TcSMUG* genes (D'Orso and Frasch, 2001b). TcUBP-1 itself is developmentally regulated, and forms part of a large (~450k Da), ribonucleoprotein complex, suggesting the involvement of sophisticated processing machinery. This complex also contains the poly(A) binding protein and at least one other RNA binding protein (D'Orso and Frasch, 2002).

TcUBP-1 is part of a family of RNA binding proteins which contain the RRM type RNA binding domain (Burd and Dreyfuss, 1994). A query of the *T. cruzi* genome database using the RRM domain resulted in the identification and cloning of five

other members of this family. These proteins show distinct stage-specific expression patterns, bind to RNA *in vitro* and are associated with different *T. cruzi* transcripts *in vivo* (De Gaudenzi et al., 2003). It will be interesting to determine the precise function of different RNP complexes in different life-cycle stages, and identify their specific RNA targets. A detailed characterisation of the mechanisms of *TcSMUG* regulation and further characterisation of these additional members is likely to add considerably to our understanding of the mechanisms by which stage-specific genes are regulated in *T. cruzi*.

Differential translation of mRNA transcripts and post-translational mechanisms add additional layers of control. For example, in *T. cruzi* the activity of the major cysteine proteinase cruzipain is upregulated in the insect stages, despite constant levels of mRNA throughout the life-cycle (Tomas and Kelly, 1996). In *T. brucei* stage-specific expression of the RNA-binding protein p37 is regulated both at the level of translation and post-translationally, by selective protein degradation (Li et al., 2003). A comparison of polysomal RNA fractions from epimastigotes and metacyclics identified a number of transcripts that showed stage-specific association with the polysomes (section 1.4.3). Differential mobilisation of transcripts to polysomes may be a mechanism by which stage-specific gene expression is regulated (Avila et al., 2001). The *trans*-acting factors that mediate differential mobilisation to polysomes and the factors that regulate translation have yet to be characterised.

Another level at which gene expression may be modulated as *T. cruzi* alternates between cell proliferation and cell differentiation is suggested by recent work comparing replicating and non-replicating forms of *T. cruzi* (Elias et al., 2001b). Global pol I and pol II transcription in non-replicating forms (metacyclics, cell-derived trypomastigotes, and epimastigotes in stationary phase) was significantly down regulated compared to replicating epimastigotes. Specifically, non-replicating trypomastigotes were found to incorporate 10 times less ^{3}H-uridine than replicating epimastigotes. Furthermore, total RNA extracted after *in vitro* incorporation of ^{32}P-UTP in permeabilised trypanosomes was hybridised with several pol II transcribed genes and satellite DNA. Signals with both types of probes were stronger in replicating than in non-replicating cells, while the rate of RNA decay after addition of

transcription inhibitor actinomycin D was shown to be similar in both forms. This down-regulation of transcriptional activity occurred in parallel with extensive remodelling of the nuclear structure (Elias et al., 2001b). The authors suggest two levels of control: (i) a global, non-specific down-regulation of transcription in non-proliferating stages, and (ii) post-transcriptional control of stage-specific genes, which is expected to be directed by environmental signals (Elias et al., 2001a).

No link has yet been made from a membrane surface receptor to a factor that mediates RNA stability. Elucidation of the mechanisms of differential mRNA stabilisation and translation is central to understanding how parasite development is controlled and remains therefore an active area of research.

1.7 EXPERIMENTAL MANIPULATION OF *T. CRUZI*

1.7.1 The genetic 'toolbox'

The first report of successful transformation and expression of a reporter gene in a trypanosomatid (Bellofatto and Cross, 1989) initiated a new era of research. Rapid evolution of techniques that allowed stable transformation, and manipulation of gene expression led to important discoveries in many areas of trypanosome biology (Clayton, 1999; Taylor and Kelly, 2002). The current 'genetic toolbox' for kinetoplastid research is reviewed by Beverley (2003). Individual techniques, with emphasis on *T. cruzi* research are discussed below.

DNA is delivered into trypanosomes by electroporation. Selection of transfected cells typically depends on the presence of a drug selectable marker gene which is flanked by processing signals, usually derived from a trypanosome housekeeping gene, such as glyceraldchyde-3 phosphate dehydrogenase (*GAPDH*). A list of commonly used marker genes for trypanosome transfections is presented in the review by Clayton (1999).

Episomal vectors. Several vectors have been devised for episomal expression of genes, including cosmid shuttle vectors that facilitated transfer of large (30-45 kb)

inserts (Kelly et al., 1994). The first shuttle vector that allowed expression of trans-fected genes in *T. cruzi* was pTEX (Kelly et al., 1992), which contains the neomycin phosphotransferase (*NEO*) gene flanked by *GAPDH* intergenic sequences, and a multicloning site for the insertion of the gene of interest. In the cell, pTEX is repli-cated as an episome, usually in the form of concatemers, which segregate randomly at cell division and are not always stable. The mechanisms of replication and tran-scription initiation have not been determined. In *T. cruzi* and *Leishmania*, no pro-moter is necessary for the transcription of episomal genes (Kelly et al., 1992). pTEX is routinely used to express genes in *T. cruzi*, for a number of different applications. These include over-expression studies, expression of tagged proteins (e.g. expression of epitope-tagged proteins for localisation studies) or expression of genes at a de-velopmental stage when they would normally be suppressed. A derivative of pTEX, termed pRIBOTEX was constructed by inserting a ribosomal promoter (Martinez-Calvillo et al., 1997). This increased expression levels of a chloramphenicol acetyl transferase (CAT) reporter gene several thousand fold. In contrast to pTEX, which replicates episomally, pRIBOTEX was found to integrate into the rDNA locus.

Integrative transfection. Often it is desirable to stably integrate DNA into the genome. To achieve this, trypanosomes can be transfected with a piece of DNA, containing a selectable marker gene with 5' and 3' flanking sequences that facilitate transcript processing and a few hundred base pairs homologous to the target locus. Integration into the genome occurs by homologous recombination (Hariharan et al., 1993). This method can be used to delete a gene from the genome (gene knockout). Null mutants are generated by deleting both allelic copies from the genome in two consecutive transfection steps, using two different marker genes. However, conven-tional gene knockout has limited applications in *T. cruzi* for two reasons. This process is very slow, requiring six to nine months to generate a double knockout cell line (Taylor and Kelly, 2002). More significantly, many genes of interest, such as cruzipain, adenylyl cyclase, and many interesting stage-specific surface proteins are present in multiple copies, which makes it impractical or impossible to delete them. This may explain why only a small number of *T. cruzi* null mutants have been reported in the literature (Caler et al., 1998; de Jesus et al., 1993; Manning-Cela et al., 2001).

Inducible expression. When a construct is integrated into a polycistronic transcription unit, its expression depends on pol II run-through transcription. Often it is desirable to have conditional mutants with the gene of interest under the control of an inducible promoter. Based on the tetracycline (Tet) repressor system from *E. coli*, inducible systems have been developed for application in trypanosomes (Wirtz and Clayton, 1995; Wirtz et al., 1998). Both the endogenous *T. brucei* pol I promoter from the procyclin locus (Wirtz and Clayton, 1995) or, more frequently, the bacteriophage T7 RNA polymerase and corresponding promoter have been used (DaRocha et al., 2004; Martin Taylor, unpublished; Wang et al., 2000; Wirtz et al., 1994; Wirtz et al., 1998). A vector for inducible expression of genes in *T. cruzi* is shown in Figure 1.5. The Tet-inducible expression system can also be used for *in vivo* studies. When mice are infected with trypanosomes carrying a Tet-inducible gene, administration of the Tet-derivative doxycyclin via drinking water will induce transcription of that gene in the trypanosomes.
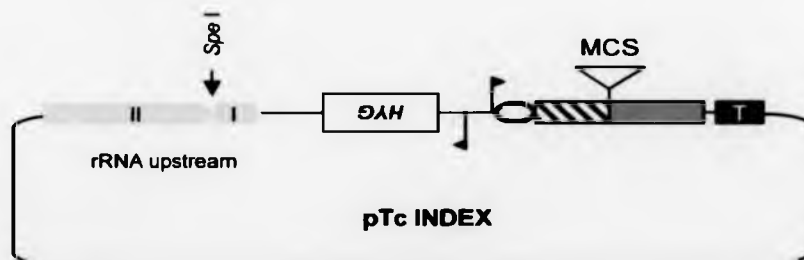
Figure 1.5: A vector for inducible protein expression in T. cruzi
Vector pTcINDEX was constructed by Martin Taylor (unpublished) for Tet-inducible expression of proteins in *T. cruzi*. The gene of interest is cloned into the multiple cloning site (MCS), flanked by the HX1 splice acceptor site from the *T. cruzi* ribosomal protein P2β H1.8 gene (hatched box; Vazquez and Levin, 1999) and actin intergenic sequence (green box). The gene of interest and the hygromycin-selectable marker gene (*HYG*) are under the control of T7 promoters (black arrows). The open circle represents the Tet- repressor binding site that allows transcription of the gene of interest to be regulated by tetracycline. T: transcription terminator. The vector is linearised with *Spe* I, and transfected into cells that carry the genes for the Tet repressor protein and T7 RNA polymerase. Sequences corresponding to the upstream region of the 18S rRNA locus (I and II) target the vector for integration into the genome.

RNAi. The discovery of RNA interference (RNAi) has added a powerful new method to the repertoire of genetic tools. This has had a major impact on *T. brucei* research (Ullu et al., 2004). It has also opened up a completely new field of research into the newly discovered role of small non-coding RNAs in gene silencing.

The initial impetus came from a landmark paper that demonstrated that expression of double-stranded RNA (dsRNA) in *C. elegans* causes rapid and sequence-specific degradation of mRNA transcripts, while leaving the endogenous genes unaltered (Fire et al., 1998). Soon it became apparent that this puzzling phenomenon functions in a wide variety of organisms including *T. brucei*, where it was first used to inhibit tubulin gene expression (Ngo et al., 1998). Subsequent research into the mechanism and function of RNAi in a number of organisms clearly demonstrated that small RNA molecules play important regulatory roles in gene silencing phenomena, ranging from viral defence mechanisms to developmental control. The hallmark of these processes is the generation of small (21-28 nt) RNAs from a number of different types of dsRNA precursors. These small RNAs are classified into three groups depending on their origin and function: 'siRNAs' guide RNA degradation, 'miRNAs' guide translation inhibition (Meister and Tuschl, 2004) and 'rasiRNAs' direct heterochromatic modifications (Lippman and Martienssen, 2004).

Components of the RNAi pathways have been identified in a number of organisms, and their study has already provided considerable insight into the mechanisms involved (Meister and Tuschl, 2004). Two proteins are central, and common to all organisms that are known to perform RNAi: DICER and argonaute (AGO). The aptly named DICER protein is a dsRNA specific RNase-III type endonuclease, which produces the small RNA duplexes (i.e. siRNAs) from dsRNA precursors. The protein consists of an RNaseIII domain, and a dsRNA binding domain. The siRNAs are subsequently incorporated into effector complexes. The so-called RNA-induced silencing complex (RISC) targets the sequence-specific degradation of target RNAs. The RISC complex always contains a member of the AGO family, which is thought to bind the siRNA directly. AGO proteins contain two conserved domains, PAZ and PIWI, which appear to function in protein-RNA and protein-protein interactions respectively.

The mechanism of RNAi in *T. brucei* has been extensively studied (Ullu et al., 2004). A DICER-like enzyme, whose activity can be detected in cell free extracts processes dsRNA to siRNAs. The corresponding 'DICER' gene has not been identified, and the genome database contains no genes similar to the DICER consensus sequence. This could indicate that the *T. brucei* DICER is considerably divergent. The argonaute family member TbAGO1 is the only component of the *T. brucei* RNAi machinery that has been characterised. It is thought to form part of the RISC complex where it binds the siRNA directly (Ullu et al., 2004). Generation of two independently derived AGO1 knockouts has provided important clues as to the function of RNAi in *T. brucei* (Durand-Dubief and Bastin, 2003; Shi et al., 2004a). Both studies demonstrated that AGO1 is essential for RNAi to occur in procyclic forms, but the gene (and the ability to perform RNAi) is not essential for the survival and proliferation of the cells. Mutational analysis of TbAGO1 has subsequently identified a single arginine residue in the PIWI domain of the protein to be essential for RNAi (Shi et al., 2004b). Consistent with the proposal that RNAi has evolved as a mechanism for silencing retroposon transcripts (Tabara et al., 1999) and anti-viral defence, TbAGO1 knockout cells were found to be defective in retrotransposon silencing (Shi et al., 2004a). The cells failed to accumulate siRNAs and the steady state levels of retroposon-transcripts increased fivefold in the *T. brucei* AGO1 knockouts, an effect both of increased transcript stability and higher transcription rates (Shi et al., 2004a). As RNAi plays a central role in heterochromatic silencing in other organisms (Lippman and Martienssen, 2004; Volpe et al., 2002), the latter finding could indicate a similar function in *T. brucei*. The AGO1 knockouts described by Durand-Dubief and Bastin (2003), displayed a slow growing phenotype, and evidence of aberrant chromosome segregation. In the fission yeast *S. pombe*, the RNAi machinery is required for centromere cohesion and proper chromosome segregation in mitosis and meiosis (Volpe et al., 2002). Future studies on *T. brucei* RNAi mutants will certainly determine to what extent similar mechanisms of RNA guided heterochromatic modifications operate in *T. brucei*. The interesting possibility that small RNAs could be involved in the regulation of specific genes opens up new avenues of research. These could add additional complexity to the mechanisms by which kinetoplastid genes are regulated. Investigation of RNAi thus promises to shed new light on the biology of *T. brucei*.

The most immediate impact of the discovery of RNAi, however, was its immense usefulness as a tool to knock down genes. It is now the method of choice to down-regulate gene expression in *T. brucei*. Several labs have developed transfection vectors for expression of dsRNA in *T. brucei*, using two principal methods to produce dsRNA: (i) insertion of the sequence of interest in between two opposing promoters. Transcription of the sense and antisense strand leads to formation of dsRNA which acts as a substrate for DICER (LaCount et al., 2000; Wang et al., 2000); (ii) insertion of inverted copies of the sequence of interest under the control of a single promoter. The transcript will then form a hairpin structure of dsRNA (Bastin et al., 2000a). A major advantage of RNAi technology is its suitability for large-scale high-throughput analysis, which could not easily be achieved by conventional gene knockout strategies. For instance, simultaneous down-regulation of multiple genes can be achieved by targeting a sequence common to all these genes. This will no doubt greatly facilitate the study of functionally redundant genes. Genome-wide RNAi screens have been used to systematically knock down (almost) all annotated coding sequences in *C. elegans*, resulting in the identification of a large number of genes not previously associated with a phenotype (Kamath and Ahringer, 2003). An ongoing *T. brucei* functional genomics project aims to create a systematic collection of RNAi mutants (www.trypanofan.org).

RNAi was expected to prove especially useful to *T. cruzi* researchers, as it would circumvent the two major impediments to conventional gene knockout. An RNAi approach significantly reduces the time to generate knockdown mutants, and multiple gene copies can be targeted simultaneously. Attempts were therefore made to knock down genes by RNAi using the strategies that have been successful in *T. brucei*. For example in this lab, a vector was constructed for Tet-inducible RNAi (Martin Taylor, unpublished) which allows cloning of the sequence of interest between two opposing inducible T7 promoters, and integration into the 18S rRNA spacer region in the *T. cruzi* genome. Thorough attempts knock down genes by RNAi in *T. cruzi* (DaRocha et al., 2004, Taylor, unpublished) and in *Leishmania* (Robinson and Beverley, 2003), have all yielded negative results. Attempts to knock down trypanothione reductase in *T. cruzi* by expressing antisense RNA from an episomal vector equally failed to down regulate this gene (Tovar and Fairlamb,

57

1996). Thus, *T. cruzi* researchers are left with limited options for generating gene knockdown mutants.

A likely explanation for the inability of these organisms to perform RNAi is this: no *AGO1* homologue or any other gene involved in RNAi has been identified in the *T. cruzi* or *Leishmania* genomes (Martin Taylor, personal communication Ullu et al., 2004). Thus, at least the genome sequencing strains lack the genes for essential components of the RNAi machinery. Whether this is true for all *T. cruzi* strains and *Leishmania* species remains to be determined. How *T. cruzi*, whose genome is rich in retroposon-like elements keeps these in check is a matter of speculation. Unlike *T. brucei* and *T. cruzi*, the *Leishmania* genome is largely deficient in transposable elements, and may therefore not require this defence mechanism.

1.7.2 Functional genomics approaches

Whole genome sequence information allows to ask completely new questions (Lander, 1996). Novel high throughput methods (below) make it possible to monitor gene expression patterns or the levels and modification state of proteins in response to a stimulus on a genome wide level. Bioinformatic analysis already plays a pivotal role in molecular biology research, and computational power will become even more essential in the future to process the vast amount of data generated by these studies.

Microarray expression profiling. Microarrays are an increasingly popular method to monitor changing mRNA levels in response to experimental conditions or to compare gene expression between different cell types. It is thus a potentially very powerful method to identify developmentally regulated genes, despite the technical difficulties associated with it, and the advanced analysis skills required to detect meaningful patterns in the data produced. Microarray technology has for example been extremely successful in identifying novel stage-regulated transcripts in the malaria parasite *Plasmodium*. Random shotgun DNA microarrays were probed with RNA from blood stage trophozoites and sexual stage gametocytes and a large number differentially expressed transcripts were identified. This approach also identified previously non-annotated genes (Hayward et al., 2000). A similar approach has

58

been used to compare transcripts of *T. brucei* insect and bloodstream forms (Diehl et al., 2002) or *L. major* promastigotes, metacyclics and amastigotes (Beverley et al., 2002). Both studies successfully identified a number of known stage-specific transcripts, as well as a number of novel stage-regulated genes. However, the proportion of genes that showed more than two-fold difference in expression levels was very small. In contrast to *Plasmodium*, where genes are transcriptionally regulated, trypanosomatids control gene expression mainly post-transcriptionally (section 1.6). Monitoring their gene expression by microarray technology is therefore expected to provide only limited information.

Proteomics. High-throughput proteomics approaches (Pandey and Mann, 2000) on the other hand may prove more useful to dissect developmental regulation in kinetoplastids. Protein extracts from cells cultured under different conditions can be separated by 2D-gel electrophoresis, and protein spots which reproducibly correspond to proteins that are up- or down-regulated or modified in response to experimental stimuli, or in different stages of the life-cycle, can be excised and analysed by mass spectrometry. The genes corresponding to the selected peptide spots can easily be identified if the sequence database is sufficiently complete.

For example, proteomic methods have been successfully used to study gene expression and changes in protein phosphorylation in response to a stimulus in fibroblasts: Newly synthesised proteins were monitored by using pulsed [^{35}S] methionine labelling followed by autoradiography for spot detection on 2D-gels, and protein identification by high-sensitivity MALDI mass spectrometry (Predic et al., 2002). To detect changes in phosphorylation patterns, protein extracts separated by 2D-gel electrophoresis were electro-blotted, and phosphorylated proteins detected with anti-phosphotyrosine and anti-phosphoserine antibodies and identified by MALDI-TOF mass fingerprinting and ESI peptide sequencing. Analysis of peptide fragments by tandem mass spectrometry identified the precise site of phosphorylation (Soskic et al., 1999).

One difficulty associated with 2D gel electrophoresis is that the results have not always been reproducible. The isolation procedure of the protein sample also affects

59

the results significantly. Improved equipment and commercially available reagents have allowed for better standardisation. A detailed protocol for the analysis of trypanosome proteins by 2D gel electrophoresis and mass spectrometry has recently been made available (van Deursen et al., 2003). Using this method, around 1000 protein spots could be resolved. An initial comparison of the soluble proteins of two distinct life-cycle-stages of *T. brucei*, bloodstream and procyclic forms revealed a remarkably similar profile for both stages (van Deursen et al., 2003).

The tools for genetic manipulation of trypanosomes make it feasible to generate specific mutants for interesting genes discovered by high-throughput technologies, and eventually determine their function.

1.8 THE AIM OF THIS STUDY

In response to stress conditions, *T. cruzi* epimastigotes differentiate to metacyclics, which are pre-adapted for transmission to a mammalian host. The process of metacyclogenesis is poorly understood. In order to get a handle on the molecular mechanisms that control metacyclogenesis, it was necessary to establish an *in vitro* differentiation system suitable for experimental manipulation. The main part of the work then focused on the functional characterisation of two novel genes, *MET2* and *MET3*, which were identified in a screen for transcripts that are upregulated during metacyclogenesis in the Dm28c strain of *T. cruzi* (Yamada-Ogatta et al., 2004).

The specific objectives of the project were to:

- Determine the conditions for *in vitro* differentiation of our CL Brener laboratory stock and develop a system suited for experimental manipulation
- Identify molecular markers associated with metacyclogenesis in CL Brener
- Identify the CL Brener homologues of the *MET2* and *MET3* genes
- Determine the stage-specificity of these genes in the CL Brener strain
- Determine the subcellular localisation of the MET2 and MET3 proteins
- Functionally characterise *MET2* and *MET3* by generating knockout mutants

60

2. MATERIALS AND METHODS

2.1 MATERIALS

General chemicals were from Sigma, restriction enzymes from Promega, and oligo-nucleotides were synthesised by Sigma Genosys.

Commonly used buffers were phosphate buffered saline (PBS): 10 mM phosphate buffer, 2.7 mM KCl and 137 mM NaCl, pH 7.4; Tris buffered saline (TBS): 100 mM Tris, 150 mM NaCl, pH 7.4; 20X SSC: 3 M NaCl, 0.3 M sodium citrate.

2.2 CELL CULTURE

2.2.1 *Trypanosoma cruzi* strain

T. cruzi CL Brener (Zingales, 1997a) was used for all experiments. Epimastigotes were grown at 27°C in RPMI-1640 medium (Sigma) supplemented with 10% (v/v) heat-inactivated foetal calf serum (Sigma), 0.5% (w/v) trypticase, 0.5% (w/v) HEPES, 0.03 M haemin, 2 mM sodium glutamate, 2 mM sodium pyruvate, 2.5U ml^{-1} penicillin, 2.5 µg ml^{-1} streptomycin (referred to in the text as 'RPMI-1640').

2.2.2 Induction of metacyclogenesis

Triatomine Artificial Urine (TAU). The method of Contreras, (1985b) was used. Briefly, late log phase epimastigotes were pre-incubated for 2 hours at a density of 1-5x10^8 cells ml^{-1} in TAU (190 mM NaCl, 17 mM KCl, 2 mM MgCl$_2$, 8 mM phosphate buffer pH 6.0), followed by dilution to 3-5x10^6 cells ml^{-1} into 10 ml of TAU supplemented with 10 mM L-proline or 2 mM D-glucose, or as required, and the cells incubated at 27°C.

Grace's Insect Medium. Late log phase epimastigote cultures (between 5x10^6 and 1x10^7 cells ml^{-1}, containing less than 1% metacyclics) were collected by centri-

fugation at 1,640 g for 5 min. The old RPMI-1640 medium was removed and replaced with an equal volume of 80% RPMI-1640/20% Grace's Insect Medium (Gibco BRL) (v/v) and the cells were incubated at 27°C.

Giemsa staining. To determine the proportion of metacyclics, cells were collected by centrifugation and fixed in 2% paraformaldehyde for at least 20 min at r/t. The fixed cells were washed in PBS and suspended in PBS. 20 µl of appropriate dilutions were spotted onto glass slides and air-dried. The slides were submerged for a few seconds in methanol and stained with Giemsa for 10 min. Slides were washed with H_2O and air-dried. The percentage of metacyclics (as assessed by the location of the kineto-plast posterior to the nucleus) was determined from at least 200 cells, using a light microscope at 1,000x magnification.

2.2.3 Infection of macrophages with *T. cruzi*

Mouse macrophages (Raw 264 cell line) were used as host cells for the mammalian stages of the *T. cruzi* life-cycle. Macrophages were grown in RPMI-1640 medium (supplemented as for *T. cruzi* epimastigotes, but without haemin) at 33°C in a hu-midified atmosphere containing 5% CO_2 in tissue culture flasks or 24 well microtiter plates. To obtain *T. cruzi* amastigotes for microscopic examination, macrophages were placed on sterile glass cover slips in 24 well microtiter plates with 0.5-1 ml medium per well. Macrophages were left to adhere for at least 2 hours before infec-tion with 10 µl of a stationary *T. cruzi* culture (protocol adapted from Allaoui et al., 1999). Cell-derived trypomastigotes were obtained by infecting 10 ml cultures of approximately 50% confluent macrophage monolayers in tissue culture flasks with a 100 µl inoculum of a stationary *T. cruzi* culture. After overnight incubation, the infected macrophages were washed twice with medium to remove any residual epi-mastigotes. Cover slips were processed for immuno-staining 24 hours after infection. The culture medium in the flasks was replenished every 2 to 3 days. The first trypo-mastigotes emerged from macrophages 6 days after infection.

2.2.4 Bacteria strains and culture

E. coli XL1-blue and *E. coli* BL21 were grown at 37°C in NZCYM medium (Sigma) with agitation or on NZCYM-agarose plates with appropriate antibiotics. For blue/white selection, 0.4% X-Gal and 6 mM IPTG were added to each plate.

2.3 PREPARATION, MANIPULATION AND ANALYSIS
OF NUCLEIC ACIDS

2.3.1 Nucleic acid preparation from cells

Preparation of plasmid DNA. Plasmid DNA was prepared from *E. coli* using the standard alkaline lysis procedure (Sambrook et al., 1989), or the QIAprep® Miniprep or HiSpeed™ Plasmid Midi Kits (Qiagen), according to the manufacturer's instructions.

Isolation of genomic DNA from T. cruzi. DNA was prepared according to Kelly (1993). Parasite cells were pelleted by centrifugation at 1,640 g for 5 min, and washed once in PBS. The cells were re-suspended in TEN cell lysis buffer (50 mM NaCl, 50 mM EDTA, 1% SDS, 50 mM Tris-HCl, pH 8.0) and 100 μg ml^{-1} Proteinase K (Roche) was added. The mixture was incubated at 37°C overnight. The DNA was then purified by extraction with equal volumes of phenol:chloroform, and precipitated by addition of 1 volume of isopropanol. The DNA aggregate was transferred to a new tube, and RNA was degraded by addition of 10 μg ml^{-1} heat-treated RNase A (Sigma).

Isolation of T. cruzi chromosomes for Clamped Homogenous Electric Fields Electrophoresis (CHEFE). Parasites from late log phase cultures (10^7 cells ml^{-1}) were collected by centrifugation at 1,640 g for 5 minutes at r/t. The resulting pellet was washed once in PBS and re-suspended in PBS at approximately 3×10^8 parasites ml^{-1}. This suspension was mixed 1:1 with 1.5% low melting point agarose (Bio Rad) maintained at 42°C, and immediately pipetted into a block/plug mould (Bio Rad) and left on ice to solidify for 20 min. The blocks were removed from the casts and incu-

bated in digestive buffer (0.5M EDTA, pH 8.0; 3% sarcosyl, 2 mg ml^{-1} of Proteinase K) for 48 hours at 50°C. Blocks were washed several times with L-buffer (100 mM EDTA, 10 mM Tris pH 8.0, 20 mM NaCl) and stored in L-buffer at 4°C.

Isolation of RNA from T. cruzi. RNA was prepared according to Kelly (1993). Parasite cells were pelleted by centrifugation, washed once in PBS, and re-suspended in 4 M guanidinium thiocyanate solution (GTC) to disrupt cells and inactivate RNases. After shearing DNA by passage of the lysates through a narrow gauge syringe needle, the lysates were layered onto a 1.5 ml 5.7 M caesium chloride cushion and centrifuged at 36,000 rpm at 20°C for 20 hours in a Beckman L8/80 Ultracentrifuge using a SW55 rotor. The RNA pellet was washed in 70% ethanol and suspended in 400 μl RNA suspension buffer (0.3 M sodium acetate pH 5.5, 0.1 mM EDTA). The RNA was precipitated by addition of 1 ml ice-cold 70% and incubation at -20°C for 2 hours. The RNA pellet was dissolved in sterile ddH$_2$O, and RNA concentration was determined by measuring the absorbance at 260 nm (an A$_{260}$ of 1.0 corresponds to 40 μg RNA ml^{-1}). Alternatively, the RNeasy kit (Qiagen) was used, according to the manufacturer's instructions.

2.3.2 Polymerase chain reaction (PCR)

General PCR conditions. PCR was used to generate DNA fragments for the construction of expression vectors, knockout targeting constructs and probes for hybridisation experiments. Reaction mixes contained approximately 100 ng of plasmid or genomic template DNA, 20 pmol of each relevant primer (listed in appendix), 2 mmol dNTP mix, 1x PCR buffer (Bioline), 1-5 mM MgCl$_2$ (optimal concentration determined for each primer pair), and 1 unit Taq polymerase (Bioline) or TaqPlus Precision Polymerase mixture (Stratagene) in a total volume of 50 μl. Reactions were run on a Hybaid Omni-Gene thermal cycler, and typical cycling conditions were 30 cycles of 30 seconds denaturing at 94°C/30 seconds annealing at 50°C/1 minute elongation at 72°C. The annealing temperature and elongation time were adjusted according to the primer pair used.

PCR Primers. Restriction sites introduced to facilitate cloning are underlined.


```
SL-PRIMER:       pGGGGGATCCACAGTTTCTGTACTAOH

met2-lc1r:       pCTCCATAGCGTTGACGAACAOH

met2-lc1r:       pATGATGGGAACGACAATTCAOH


TRCMET2F:        pAGAGGATCCCCTGCCGCTCCAACCAAOH

TRCMET2R:        pAGAAAGCTTCCGACTGCGTATGCTCATCOH

TRCMET3F:        pAGAGGATCCGCAAGGATAACAAATCCCACOH

TRCMET3R:        pCACAAGTCCGGCGCAGAAACTGCTAOH


met2ORF-1f:      pCCCACTAGTATGATTCCCTGCCGCTCCAACCOH

met2ORF-1r:      pCCCGCCGGCCTTAGATGAATCATCTATAATCCTGCOH

met3ORF-1f:      pCCCACTAGTATGTCGCGACGCAAAGATAACOH

met3ORF-1br:     pAAAGCCGGCGGGCCCACCTGTCCTCOH


MET3-GFPf(2):    pCCCGATATCATGTCGCGACGCAAAGATAACOH

MET3-GFPf(5):    pCCCGAATTCATGCGTCGCCACCGTGACGOH

MET3-GFPf(7):    pCCCGAATTCATGGGCCGTCAGGGAACGCCTCOH

MET3-GFPf(11):   pCCCGAATTCATGTCTACGGAAAGCGACCOH

MET3-GFPf(14):   pCCCGAATTCATGACAGGGAGCCGAGCAAGOH


MET3-GFPr(3):    pCCCGGATCCTGATCGCGAAGTCGACGOH

MET3-GFPr(4):    pCCCGGATCCTGATGCTCCCGTCCTTGATOH

MET3-GFPr(6):    pAAAGGATCCGGCCCACCTGTCCTCGOH

MET3-GFPr(9):    pCCCGGATCCCCTTCACCCGTGCGTGCOH

MET3-GFPr(10):   pCCCGGATCCGAGCGTGGGAGATAGGOH

MET3-GFPr(12):   pCCCGGATCCCCTATTACACGTCCCGOH

MET3-GFPr(13):   pCCCGGATCCGATGTGGACGGGAGGAOH


MET3-9E10f(7):   pAAAACTAGTATGGGCCGTCAGGGAACGCCTCOH

MET3-9E10r(8):   pGGGAAGCTTACAAGTCCTCTTCAGOH
```

65

```
MET2-5FL-F:    pCCCGAGCTCTGTGCATGTATCCTTC OH
MET2-5FL-R:    pCCCGATATCAACGCACCAGAGACAATC OH
MET2-3FL-F:    pCCGCTCGAGAAGCTTTTAAAGGACATTG OH
MET2-3FL-R:    pCCCGGTACCTCTTCTTCTCGCATGC OH
MET3-5FL-F:    pCCCGAGCTCTATGTGAGACTTGAACG OH
MET3-5FL-R:    pCCCGATATCTATTATTGTGTGACGCAGC OH
MET3-3FL-F:    pCCCGGTACCGAGTGCAATATATACAC OH
MET3-3FL-R:    pCCCCTCGAGTTTCTGCGCCAATTGAG OH
```

2.3.3 RT-PCR amplification of *MET2* and *MET3* cDNA

Reverse transcriptase PCR (RT-PCR) was used to identify the splice acceptor site the
5'UTR of the *MET2* and *MET3* genes. The Access RT-PCR system (Promega) was
used according to the manufacturer's recommendations. For first strand cDNA syn-
thesis, 1 µg of Oligo(dT)$_{15}$ primer and 1 µg total RNA from *T. cruzi* stationary cells
were used. Reverse transcription reactions were carried out in 50 µl, at 37°C for 60
min, followed by 2 min at 94°C. For the second strand cDNA synthesis and PCR
amplification a forward primer was used that recognised the spliced leader sequence
(SL-PRIMER), and the reverse primer was specific to the coding sequence of *MET2*
(primer met2-lc1r) or *MET3* (primer met3-lc1r). Cycling conditions were 30 seconds
denaturation at 94°C/30 seconds annealing at 55°C/1 minute elongation at 72°C (30
cycles). The PCR fragments were cloned in pGEM-T (Promega) and sequenced.

2.3.4 Cloning techniques and vector construction

DNA cloning techniques. Standard techniques were employed for restriction enzyme
digestion of DNA, ligation reactions and propagation of plasmids in *E. coli*.

Isolation of DNA fragments from agarose gels. For purification of DNA from aga-
rose gels the QIAquick® Gel Extraction Kit (Qiagen) was used according to the
manufacturer's instructions.

Vectors for expression of MET2 and MET3 in E. coli. For in-frame cloning into the expression vector pTrcHis C (Invitrogen), the respective coding sequences were PCR-amplified from *T. cruzi* CL Brener genomic DNA as follows: fragment met2C was amplified with the primers TRCMET2F and TRCMET2R, and fragment met3C was amplified with the primers TRCMET3F and TRCMET3R. pTrcHis-met2 and pTrcHis-met3 were generated by digesting the pTrcHis C vector and the met2C and met3C fragments with *Bam* HI and *Hind* III, and subsequent ligation of the relevant PCR fragment with the vector. This resulted in fusion genes encoding proteins with an N-terminal 6x histidine tag for protein purification, and an epitope tag for detection with the anti Xpress antibody. Plasmid clones were identified by restriction digest analysis, and the sequences were verified by sequencing both strands of the inserts.

Vectors for expression of c-myc tagged protein in T. cruzi. Episomal expression vectors pTEX-MET2-9E10 and pTEX-MET3-9E10 (Figure 2.1) were constructed by replacing the *APX* coding sequence in pTEX-APX-9E10 (Wilkinson et al., 2002) with the gene of interest. The resulting proteins carried at their C-terminus the amino acid sequence of the c-myc(9E10) epitope (EQKLISEEDL). The fragments corresponding to the *MET2* and *MET3* coding sequences were amplified by PCR from *T. cruzi* CL Brener genomic DNA as follows. The *MET2* coding sequence was amplified using primers met2ORF-1f and met2ORF-1r. The *MET3* coding sequence was amplified using primers met3ORF-1f and met3ORF-1br. The resulting PCR fragments were cloned in pGEM-T (Promega) and sequenced. The vector pTEX-APX-9E10 was digested with *Spe* I / *Eco* RV to remove the 1 kb *APX* fragment. The linearised vector fragment was ligated with the *MET2* or *MET3* coding sequence, which had been excised from the relevant pGEM-T construct with *Spe* I / *Nae* I. The resulting plasmids pTEX-MET2-9E10 and pTEX-MET3-9E10 were sequenced, and the correct in-frame fusion of the c-myc sequence with the 3' end of the coding sequence was confirmed.
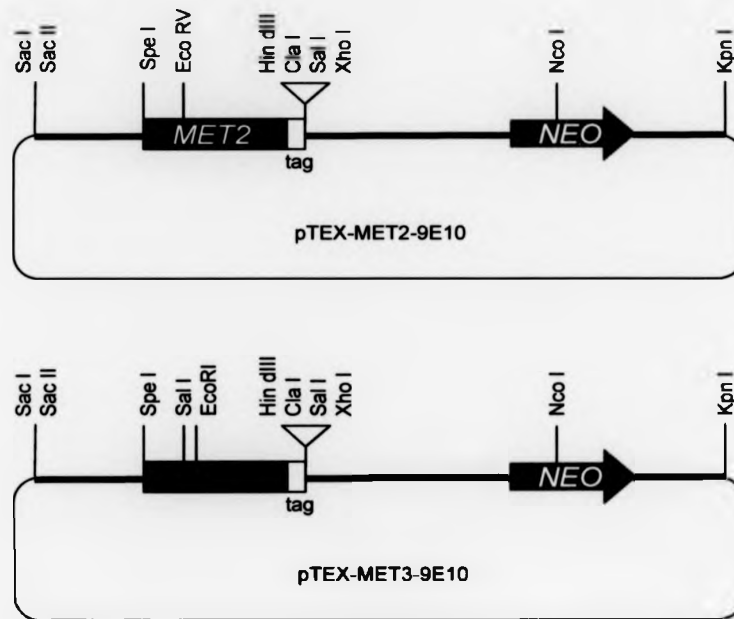
Figure 2.1: pTEX-MET2-9E10 and pTEX-MET3-9E10
Episomal vectors for expression of MET2 and MET3 proteins with a C-terminal c-myc epitope tag.
Strong black lines indicate *GAPDH* intergenic sequences that facilitate transcript processing.

Vectors for expression of GFP-fusion proteins in T. cruzi. pTEXeGFP was constructed by Martin Taylor. The eGFP gene was excised from pEGFP (Clontech) using *Eco* RI (overhang blunted) and *Hin* dIII and ligated into pTEX (Kelly et al., 1992) digested with *Xho* I (overhang blunted) and *Hind* III. Vectors to express MET3-GFP fusion proteins were constructed by replacing the BPP-1-N fragment in pTEXeGFP-N (Bromley et al., 2004) with fragments of the *MET3* coding sequence. The *MET3* fragments were generated by PCR, using the vector pTEX-MET3-9E10 as DNA template. Forward primers were MET3-GFPf(2), -(5), -(7), -(11) and -(14). Reverse primers were MET3-GFPr(3), -(4), -(6), -(8), -(9), -(10), -(12) and -(13). The resulting PCR fragments were digested either with *Eco* RI / *Bam* HI or *Eco* RV / *Bam* HI, depending on which site was engineered into the primer. pTEXeGFP-N was

68

cut with *Eco* RI / *Bam* HI to remove the BPP-1-N fragment, and then ligated with the *Eco* RI / *Bam* HI digested PCR fragments. For ligation with the *Eco* RV / *Bam* HI digested PCR fragments, pTEXeGFP-N was linearised with *Eco* RI and the ends made blunt. The BPP-1-N fragment was then excised by digesting with *Bam* HI, and the vector fragment ligated with the relevant PCR fragment. To build fusion construct 7/8 (encoding N-terminally truncated MET3 with a C-terminal c-myc tag) primers MET3-9E10f(7) and MET3-9E10r(8) were used to amplify the relevant fragment from pTEX-MET3-9E10. This PCR fragment was *Spe* I / *Hin* dIII digested and cloned into *Spe* I / *Hin* dIII digested pTEX. The MET3-GFP and MET3-9E10 fusion constructs were sequenced. A schematic diagram of the constructs is shown in Figure 4.27. The names of the constructs indicate the primer pairs used for the amplification of the *MET3* fragments.

MET2 and MET3 knockout targeting constructs. Fragments corresponding to the sequences flanking the *MET2* and *MET3* coding sequences were amplified by PCR from *T. cruzi* CL Brener genomic DNA as follows. The *MET2* 5' flanks were amplified using primers MET2-5FL-F and MET2-5FL-R. The resulting fragment derived from allele 39p3 is 921 bp and contains a diagnostic *Xho* I restriction site. The fragment derived from allele 1o17 is 639 bp. The *MET2* 3' flanks (2.3 kb fragments) were amplified using primers MET2-3FL-F and MET2-3FL-R. The *MET3* 5' flank (416 bp fragment) was amplified using primers MET3-5FL-F and MET3-5FL-R. The *MET3* 3' flank (1.3 kb fragment) was amplified using primers MET3-3FL-F and MET3-3FL-R. The PCR fragments were cloned in vector pGEM-T (Promega). The resulting plasmids pGEM-T-5'MET2(1o17), pGEM-T-5'MET2(39p3), pGEM-T-3'MET2(1o17), pGEM-T-5'MET3 and pGEM-T-3'MET3 were sequenced.

Construction of pko-MET2(39p3)-NEO. The 5' *MET2*(39p3) *Xho* I / *Kpn* I fragment excised from pGEM-T-5'MET2(39p3) was ligated into pko2-MPX-NEO (Shane Wilkinson, unpublished) from which the 5' *MPX* fragment had been excised with *Sac* I / *Spe* I. The resulting plasmid pko-5'MET2-NEO was digested with *Hin* dIII / *Kpn* I to remove the *3'MPX* fragment. The 3'MET2(39p3) *Hin* dIII / *Kpn* I fragment excised from pGEM-T-3'MET2(39p3) was ligated into the linearised pko-5'MET2-NEO vector, resulting in plasmid pko-MET2-NEO (Figure 2.2A).

Construction of pko-MET2(1o17)-HYG. The 5' *MET2*(1o17) *Sac* I / *Spe* I fragment excised from pGEM-T-5'MET2(1o17) was ligated into pSHYGK (courtesy of Martin Taylor) which had been linearised with *Sac* I / *Spe* I. The resulting plasmid pko-5'MET2-HYG was linearised with *Hin* dIII / *Kpn* I. The 3'MET2(1o17) *Hin* dIII / *Kpn* I fragment excised from pGEM-T-3'MET2(1o17) was ligated into linearised pko-5'MET2-HYG resulting in plasmid pko-MET2(1o17)HYG (Figure 2.2B).

Construction of pko-MET2(1o17)-PAC. The *HYG* gene in pko-MET2(1o17)HYG was replaced with the *PAC* gene as follows. pko-MET2(1o17)HYG was digested with *Eco* RI / *Sac* I and the pko-MET2(1o17) vector fragment was isolated from an agarose gel (thereby removing the *HYG* ORF and the 5' *GAPDH* flank). Digestion of pko-MPX-PAC (Shane Wilkinson, unpublished) with *Eco* RI / *Sac* I, allowed isolation of a *Eco* RI / *Sac* I fragment containing the *PAC* ORF, and an *Eco* RI fragment containing the 5' *GAPDH* flank. First, the *PAC* ORF was ligated into the pko-MET2(1o17) vector fragment, resulting in pko-MET2-PAC1. Then the 5' *GAPDH* flank was ligated into pko-MET2-PAC1, linearised with *Eco* RI, resulting in plasmid pko-MET2(1o17)-PAC (Figure 2.2C). The correct orientation of the 5' *GAPDH* flank was verified by a diagnostic *Sac* II digest.

Construction of pko-MET3-HYG. The 5'*MET3* *Sac* I / *Spe* I fragment excised from pGEM-T-5'MET3 was ligated into pSHYGK (Martin Taylor, unpublished) which had been linearised with *Sac* I / *Spe* I. The resulting plasmid pko-5'MET3-HYG was first linearised with *Hin* dIII, the ends blunted, and then digested with *Kpn* I. pGEM-T-3'MET3 was digested with *Xho* I and the end made blunt. Then the plasmid was digested with *Kpn* I to excise the 3'MET3 *Xho* I / *Kpn* I fragment, which was ligated into linearised pko-5'MET3-HYG, resulting in plasmid pko-MET3-HYG (Figure 2.3A).

Construction of pko-MET3-PAC. The *HYG* gene in pko-MET3-HYG was replaced with the *PAC* gene as follows. pko-MET3-HYG was digested with *Spe* I / *Sal* I to obtain vector fragment pko-MET3, containing the MET3 flanks and the 3' *GAPDH* flank. A fragment containing the 5' GAPDH flank and the *PAC* ORF was excised

70

from pko-MET2(1o17)PAC (see above) and ligated into pko-MET3, resulting in pko-MET3-PAC (Figure 2.3B).

The integrity of all knockout constructs was checked by analysis of restriction patterns. For transfection of *T. cruzi*, the fragments containing the *MET2* or *MET3* flanks and drug selectable marker gene were excised from the relevant vector with *Sac* 1 / *Kpn* 1.
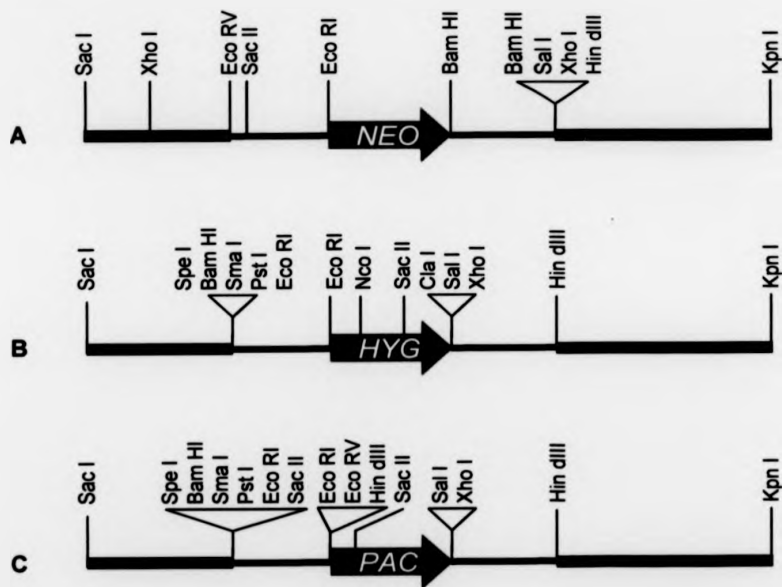
Figure 2.2: Targeting constructs for MET2 *knockouts*

Linear targeting fragment from **(A)** pko-MET2(39p3)-NEO, **(B)** pko-MET2(1o17)-HYG and **(C)** pko-MET2(1o17)-PAC. These fragments were used to replace the *MET2* gene with the neomycin phosphotransferase gene (*NEO*), the hygromycin phosphotransferase gene (*HYG*) or the puromycin acetyl transferase gene (*PAC*). Strong black lines indicate *GAPDH* intergenic sequences that facilitate transcript processing. Blue lines indicate sequences flanking the *MET2* gene, used to target homologous recombination.
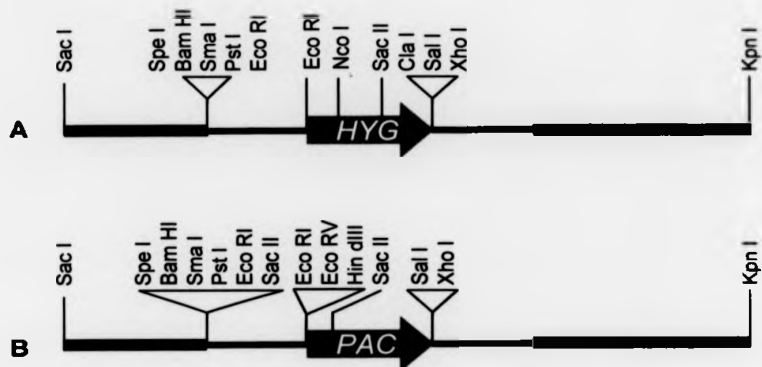
Figure 2.3: Targeting constructs for MET3 *knockouts*
Linear targeting fragment from **(A)** pko-MET3-HYG and **(B)** pko-MET3-PAC. Strong black lines indicate *GAPDH* intergenic sequences that facilitate transcript processing. Red lines indicate sequences flanking the *MET3* gene.

2.3.5 Radio-labelling of DNA fragments

The Rediprime™ II Random Prime Labelling System (Amersham Pharmacia Biotech) was used according to the manufacturer's instructions to label DNA fragments with [^{32}P] dCTP for use in hybridisation protocols.

2.3.6 Electrophoresis and blotting of nucleic acids

Southern blotting. Genomic DNA was incubated with the appropriate restriction enzymes for >4h at 37°C and size-fractionated on 0.8% agarose gels in 1x TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0), at 20 V over-night. DNA was transferred to Nylon membranes (Osmonics) in 10x SSC by standard Southern blot procedures, using capillary transfer over-night. After transfer, the DNA was cross-linked to the membrane in a UV Stratalinker (Stratagene).

Separation of T. cruzi chromosomes. 1% agarose gels for Clamped Homogenous Electric Fields Electrophoresis (CHEFE) were made with 0.5 x TBE buffer (0.89M Tris, 0.89M sodium borate, 0.02M EDTA) and low-melt agarose (Bio Rad). *T. cruzi* chromosomes embedded in agarose blocks (section 2.3.1) were inserted into the wells of the gel and sealed with agarose. *Saccharomyces cerevisiae* chromosomal DNA (Bio Rad) was used as size markers. Gels were run in the CHEF Mapper® XA System (Bio Rad), at 14°C, using the auto-algorithm settings. A separation range of 0.2 Mb to 1 Mb was chosen for *MET2* analysis, and 0.5 Mb to 2 Mb for *MET3* analysis. To visualise chromosomes the gel was stained after the run with ethidium bromide, and de-stained in distilled water. The DNA transferred to nylon membranes by Southern blotting.

Agarose electrophoresis of RNA. To denature, 10 μg RNA in 5.4 μl ddH$_2$O were mixed with 5.4 μl 6 M glyoxal, 16 μl DMSO, 3 μl 1x TAE buffer, and incubated for 1 hour at 50°C. Denatured RNA was size-fractionated on 1.2% agarose gels, run in 1x TAE buffer at 100 V. RNA was visualised by staining the gel with ethidium bromide after fractionation.

Northern blotting. RNA was transferred to Nylon membranes in 20x SSC by capillary transfer overnight, using the same setup as for Southern bots.

Hybridisation of membranes. Non-specific binding sites on the membranes were blocked by pre-hybridisation for >1 hour at 65°C in hybridisation solution (1x Denhardt's solution, 6x SSC, 0.5% SDS, 0.2 mg ml^{-1} sheared and denatured herring sperm DNA). Labelled DNA probes were denatured by boiling for 5 min, chilled on ice and added to the hybridisation. Hybridisation was overnight at 65°C. To remove non-specific bound probe, membranes were washed in 2x SSC for 5 min at r/t, 20 min in 0.5x SSC / 0.1% SDS at 65°C, and 20 min in 0.1x SSC / 0.1% SDS at 65°C with agitation. [^{12}P]-labelled probes were detected by exposing the membrane to X-ray film (Kodak XAR-5) and intensifying screens at -80°C for the required time.

2.3.7 DNA sequencing

DNA sequencing was performed by a dideoxynucleotide chain termination method using a DNA Sequencing Kit (Applied Biosystems). Briefly, DNA was mixed with 1.6 pmol of oligo-nucleotide primer and the BigDye™ Terminator Cycle Sequencing Ready reaction mix, in a total volume of 10 μl. Cycling conditions were 25 cycles of 30 sec at 96°C / 15 sec at 50°C / 4 min at 60°C in a Hybaid OmniGene thermal cycler. Unincorporated nucleotides were removed by passing the sample through a Sephadex G-50 column before loading onto an ABI PRISM 377 automated sequencer.

2.3.8 Transfection of *T. cruzi*

Electroporation. DNA was introduced into *T. cruzi* by electroporation. Briefly, epi-mastigotes from a logarithmically growing culture were collected by centrifugation and washed once in cytomix (2 mM EGTA, 120 mM KCl, 0.15 mM $CaCl_2$, 10 mM K_2HPO_4/KH_2PO_4, 25 mM HEPES, 5 mM $MgCl_2.6H_2O$, 0.5% glucose, 100 μg ml^{-1} BSA, 1mM hypoxanthine; pH 7.6). The cells were then suspended in 1x cytomix at approximately 10^8 cells ml^{-1}. 0.5 ml of this cell suspension was mixed with approximately 10 μg DNA (in 50 μl) in a Bio Rad Gene Pulser Cuvette. Cells were electroporated in a Bio Rad Gene Pulser apparatus set at 1.5 kV, using two pulses.

Selection of episomes. Electroporated cells were transferred to 10 ml RPMI-1640 medium. After 48 hours incubation at 27°C, the culture was split in three aliquots. Fresh RPMI-1640 was added to each aliquot to a final volume of 10 ml, and drug was added as required. After one week, these cultures were diluted 1:10 into fresh selection medium.

Selection of integration events. To obtain clonal lines, electroporated cells were transferred to 12 ml supplemented RPMI-1640 medium, and transferred to a 24 well microtiter plate (0.5 ml per well). After 48 hours incubation at 27°C / 5% CO_2, 0.5 ml of fresh medium was added, containing twice the required drug concentration. Plates were monitored at weekly intervals. If the cell density in the wells was too

74

high, 0.25 ml aliquots were transferred to a fresh plate, and fresh selection medium added to 1 ml.

Drugs. To select for resistant epimastigotes, Hygromycin B (Life Technologies) and G418 were added to final concentrations of 80-100 µg ml^{-1}. Puromycin was added to a final concentration of 3-5 µg ml^{-1}. Drug resistant populations typically started to grow out 4-6 weeks after transfection.

2.4 PREPARATION AND ANALYSIS OF PROTEIN

2.4.1 Preparation of protein extracts

Preparation of T. cruzi protein. To obtain whole cell lysates, 2 x 10^7 *T. cruzi* cells were lysed with 200 µl of protein sample buffer (0.4 M Glycine, 2% (w/v) SDS, 5% (w/v) 2-mercaptoethanol, 0.1% bromphenol blue), vortexed and boiled for 5 min. Samples were stored at –20°C. Alternatively, *T. cruzi* cells were lysed in protein lysis buffer N (1% (v/v) Nonidet P-40, 50 mM Tris pH 8.0, 150 mM NaCl, Roche Mini-Complete™ protease inhibitors), 30 min on ice. Lysates were centrifuged in a micro-fuge at 13,000 rpm for 5 min to separate the soluble and insoluble fractions. To assess the effect of salt concentration on MET3 protein solubility, the NaCl concentration in protein lysis buffer N was adjusted as required.

Preparation of extracts for detection of phosphorylated proteins (adapted from Mynott et al., 1999). 1.5x10^8 *T. cruzi* cells were collected by centrifugation at 1640 g for 5 min. The cell pellet was washed once in PBS, suspended in 700 µl protein lysis buffer T, containing phosphatase inhibitors and protease inhibitors (25 mM Tris pH 7.4, 75 mM NaCl, 0.4 mM EDTA, 0.5% Triton X-100, 0.4 mM sodium ortho-vanadate, 10 mM Na F, 10 µg ml^{-1} leupeptin, 740 µm PMSF, 0.7 µg ml^{-1} pepstatin A, 1 µg ml^{-1} E64, 100 µM TLCK) and incubated on ice for 30 min until lysis was complete. The suspension was centrifuged in a microfuge at 13,000 rpm for 20 min at 4°C. The resulting supernatant (soluble fraction) was rapidly frozen in liquid nitrogen and stored at -80°C. The pellets (insoluble fraction) were stored at -20°C.

Preparation of protein extracts for detection of cruzipain activity. Parasites were collected by centrifugation at 1640 g at 4°C for 10 min. The cell pellet was washed once in PBS, solubilised at 1×10^8 cells ml^{-1} in TBS/ 1% (v/v) Nonidet P-40 and vortexed rapidly for 15 sec. The suspension was centrifuged in a microfuge at 13,000 rpm for 3 min at r/t and the resulting supernatant was rapidly frozen in liquid nitrogen and stored at -20°C (Greig and Ashall, 1990).

Determination of protein concentration. The protein content of parasite extracts was estimated using the BCA protein assay system (Pierce), using bovine serum albumin (BSA) as a standard and performed according to the manufacturer's instructions.

2.4.2 Expression of recombinant MET2 and MET3 in *E. coli*

100 ml of an overnight culture of *E. coli* BL21 pTrcHis-met2 or *E. coli* BL21 pTrcHis-met3 in NZCYM containing 100 μg ml^{-1} ampicillin (NZCYM-amp) was diluted 1:10 into 1 litre NZCYM-amp, and incubated at 37°C with shaking for 1-3 hours. Expression was induced by adding IPTG to a final concentration of 1 mM and the culture was incubated for an additional 2 hours (for MET2 expression) or overnight (for MET3 expression). The cells were collected by centrifugation in a Beckman J2-21 centrifuge at 4,000 rpm for 20 min. The cell pellets were frozen at -80°C for at least 30 min prior to lysis.

2.4.3 Purification of recombinant MET2 and MET3 from *E. coli*

For preparation of cleared lysates under native conditions, cell pellets were lysed for 30 min on ice in lysis buffer (50 mM NaH_2PO_4 pH 8.0, 300 mM NaCl, 10 mM imidazole, 1 μg ml^{-1} lysozyme, 1% Triton X-100, Roche Mini Complete™ protease inhibitors, EDTA-free), before sonication on ice for six 10-second bursts, with a 10-second cooling period between each burst. The lysate was centrifuged in a Beckman J2-21 centrifuge at 12,000 rpm for 20 min at 4°C. 2 ml of 50% Ni-NTA slurry (Qiagen) was added to the cleared lysate and mixed by shaking at 4°C for at least one hour. The lysate-Ni-NTA mixture was loaded into a column and the flow-

through collected. The column was washed with 4x 5 ml wash buffer (50 mM NaH$_2$PO$_4$, pH 8.0, 300 mM NaCl, 10 mM imidazole, Mini Complete™). The protein was eluted by increasing the amounts of imidazole in the buffer to 20 mM, 50 mM, 100 mM, 200 mM and 500 mM, and using 4 ml of each elution buffer. All fractions were stored at 4°C until use. 20 µl samples of each fraction were analysed by SDS-PAGE and Western blotting.

2.4.4 Generation of antiserum

Purified 6xHIS-MET3 or 6xHIS-MET2 protein was excised from an SDS-PAGE gel, placed in a pestle and mortar and pulverised in the presence of liquid nitrogen. 1 ml Freund's complete adjuvant was added and the sample homogenised by passage through a syringe needle. Immediately before use, the sample was sonicated for five 5-second blasts. Approximately 150 µl of this suspension was injected intra-peritoneally into each of three Balb/c mice, which had previously been pre-bled from a tail vein. This was done by Michael Miles at LSHTM. After 2 weeks, another protein sample was prepared in the same way, except that Freund's incomplete adjuvant was used, and administered to the mice. This was repeated twice more at two week intervals. Blood samples were left to clot at 4°C overnight, then centrifuged for 10 min at 13,000 rpm in a microfuge. The serum supernatant was stored in 50% glycerol, 0.02% sodium azide at -20°C.

2.4.5 Electrophoresis and blotting of proteins

SDS-PAGE. Parasite extracts were separated by discontinuous SDS-polyacrylamide gel electrophoresis (SDS-PAGE) in a Mini-PROTEAN 3 electrophoresis cell (Bio Rad) according to standard procedures. 10% or 12% polyacrylamide Ready Gels (Bio Rad) were used as required. Samples were mixed with sample buffer (0.4 M Glycine, 2% (w/v) SDS, 5% (w/v) 2-mercaptoethanol, 0.1% bromphenol blue) and boiled for 5 min prior to loading. Protein from 2 x 10^6 cells was loaded per lane. Precision protein standards™ (Bio Rad) were loaded in parallel. Electrophoresis was performed at 100 V in 1x running buffer (25 mM Tris/ 0.25 M glycine/ 0.1% (v/v) SDS). Gels were stained in 0.25% (w/v) Coomassie Brilliant Blue R 250/ 45% (v/v)

methanol/ 45% (v/v) ddH$_2$0/ 10% glacial acetic acid and destained in 5% (v/v) methanol/ 7.5% (v/v) glacial acetic acid or by boiling in ddH$_2$O.

Western Blotting. Proteins separated by SDS-PAGE were transferred to Protran[®]-nitrocellulose membranes (Schleicher and Schuell), in a Trans-Blot Semi-Dry Electrophoretic Transfer Cell (Bio Rad) according to the manufacturer's instructions, using Bjerrum and Schaefer-Nielsen transfer buffer (48 mM Tris, 39 mM glycine, 20% (v/v) methanol). Blotting was carried out at 15 V for 1 hour. Membranes were blocked in TBS/ 5% (w/v) milk powder for 30 min. Antibodies were diluted in TBS/ 5% milk powder and incubated at r/t for one hour. Unbound antibody was washed off with three changes of TBS. Primary antibodies were anti Xpress™ (Invitrogen), diluted 1:5,000; mouse anti c-myc (9E10) (Santa Cruz Biotechnology), diluted 1:1,000; mouse antiserum to MET3 (section 2.4.4), diluted 1:500; β-tubulin antibody KMX1 (a gift from K. Gull), diluted 1:500. Secondary antibodies were goat anti-mouse IgG horseradish peroxidase conjugate (Bio Rad), diluted 1:3,000. The blot was incubated with ECL Plus chemiluminescence substrate (Amersham Pharmacia Biotech) for 5 min prior to exposure to X-ray film for 1-5 min.

For detection of phosphoproteins, free binding sites on the membrane were blocked with Blotto B (TBS/ 1% (w/v) BSA/ 1% (w/v) milk powder) at r/t for at least 1 hour or at 4°C overnight. The following antibodies were used: Rabbit anti-Phosphoserine (PS) and Rabbit anti-Phosphothreonine (PT), polyclonal, 0.2 mg ml^{-1} (Zymed Laboratories Inc.), and Mouse Anti-Phosphotyrosine (4G10), monoclonal, 1 mg ml^{-1} (Upstate Biotechnology UK). The relevant antibody was diluted 1:500 in TBS/ 1% BSA and added to the membranes. Membranes were incubated either at r/t for one hour, or at 4°C overnight, with gentle agitation. Unbound antibody was washed off with three changes of TBS. The blots were then incubated as required with goat anti-rabbit or goat anti-mouse IgG horseradish peroxidase conjugate (Bio Rad) diluted 1:3,000 in Blotto B for 1 hour at room temperature, and washed as above prior to chemiluminescence detection.

2.4.6 Immuno-fluorescence and confocal microscopy

T. cruzi cells were fixed in suspension with 2% paraformaldehyde (PFA) in PBS for 20 min at r/t. Fixed cells were pelleted and washed twice with PBS and re-suspended in PBS. 20 µl cell suspensions were spotted onto multitest glass slides (ICN Bio-medicals Incorporated) and dried at r/t. To fix *T. cruzi*-infected macrophages grown on glass cover slips, fixative was added to the culture medium to a final concentra-tion of 2% PFA. After 20 min, cover slips were washed several times with PBS and left to dry. For immuno-staining, the slides or cover slips were placed in a humid chamber at r/t. Unreacted aldehyde groups were blocked with 50 mM NH$_4$Cl (in PBS), 5 min. After a rinse in PBS, non-specific protein binding sites were blocked with 50% horse serum in wash buffer (PBS/ 0.1% (w/v) saponin), 20 min. Anti-bodies were diluted in wash buffer/ 2% horse serum. Slides were incubated with the primary antibody for 45 min, and with the secondary antibody for 30 min, and rinsed in three changes of wash buffer. Primary antibodies were mouse anti c-myc (9E10) (Santa Cruz Biotechnology), diluted 1:200; mouse antiserum to MET3 (section 2.4.4), diluted 1:400; L1C6 monoclonal antibody (K. Gull, unpublished), diluted 1:200. Secondary antibodies were AlexaFluor 488- and 546-conjugated anti-mouse or anti-rabbit antibodies (as appropriate), all diluted 1:400. DNA was stained with 1 µM TOTO-3 (Molecular Probes) in PBS/ 0.1% saponin/ 10 mg ml^{-1} RNase A, for 20 min. Slides were washed twice in PBS and mounted in PBS/ 50% glycerol and viewed under a Zeiss LSM 510 axioplan confocal laser scanning microscope.

2.4.7 Protease assay

Detection of cruzipain activity in gelatin gels. For detection of protease activity, 10% polyacrylamide gels were prepared according to standard procedures, but 0.2% (w/v) gelatin was co-polymerised with the resolving gel (Greig and Ashall, 1990). Samples were diluted in sample buffer (0.4 M glycine, 2% (w/v) SDS, 0.1% bromphenol blue) without 2 -mercaptoethanol (2-Me) and were not boiled. Electrophoresis was at 100 V for 2 to 3 hours. After electrophoresis, gels were shaken in 200 ml of 2% (v/v) Triton X-100 for 1 hour for SDS removal. For protease detection, gels were incubated in digestion buffer (50 mM sodium citrate, 50 mM HEPES, 50 mM

79

sodium phosphate, 50 mM sodium Borate (Borax); pH 5.6) with 26 nM 2-Me for 18 hours at 37°C. Coomassie Blue staining was carried out as described in section 2.4.5 except that for destaining 10% (v/v) glacial acteic acid was used. Gels were photographed on a light box.

2.5 BIOINFORMATIC SEQUENCE ANALYSIS

Preliminary genomic data were accessed via http://www.genedb.org/. Most of the genomic data were provided by the Wellcome Trust Sanger Institute and The Institute for Genomic Research (TIGR), which are supported by the Wellcome Trust and the National Institutes of Health, USA (grant AI043062), respectively. Searches of GenBank were performed using the BLAST server at http://www.ncbi.nlm.nih.gov. Web-based analysis tools for protein sequence analysis were accessed via http://www.expasy.ch. In addition, for secondary structure predictions the programs PSIPRED (McGuffin et al., 2000), SSPRO2 (Pollastri et al., 2002), PROF (Ouali and King, 2000) were used. For 3D structure predictions, programs ESyPred3D (Lambert et al., 2002), 3D-PSSM (Kelley et al., 2000) and FUGUE (Shi et al., 2001) were used. The default parameters were used unless otherwise indicated.

3. *IN VITRO* METACYCLOGENESIS

The study of metacyclogenesis has been complicated by conflicting reports on the factors that trigger it (section 1.4.2), and a paucity of molecular data (section 1.4.3). Advances in genetic manipulation techniques in conjunction with the information available from the genome sequencing project provide new opportunities to address longstanding questions regarding the differentiation from replicating non-infective epimastigotes to non-replicating infective metacyclics. However, the design of informative experiments and the interpretation of results obtained with genetically altered parasites, rely on a model system for metacyclogenesis. This requires development of a robust *in vitro* metacyclogenesis assay, and a description of morphological and molecular correlates of differentiation.

The work reported in the first part of this chapter (section 3.1) therefore aimed to establish culturing conditions that promote metacyclogenesis in our laboratory stock of the genome sequencing strain *T. cruzi* CL Brener. The second part of this chapter describes morphological (section 3.2) and molecular changes (section 3.3) that are associated with metacyclogenesis in this strain. Specifically, the upregulation of protease activity was demonstrated using a qualitative method, and changes in phosphorylation patterns were investigated using antibodies against phosphorylated residues. The results of an in-depth characterisation of two novel genes that are upregulated during metacyclogenesis are presented in chapters 4 and 5.

3.1 *IN VITRO* METACYCLOGENESIS

To dissect the molecular mechanisms of differentiation *in vitro*, we require a system in which differentiation occurs at a high and reproducible rate. Early experiments with cultured *T. cruzi* epimastigotes established that the development of metacyclics is favoured in old cultures (Brener, 1973). Over the years, various culture conditions have been reported to enhance metacyclogenesis (section 1.4.2), but what works consistently well in one parasite strain in one laboratory has often been found not to work in another strain or laboratory. This could be a consequence of the genetic

diversity of *T. cruzi*. Moreover, the rate of metacyclogenesis not only differs between different *T. cruzi* strains, but it is negatively affected by the length of time that a strain has been maintained in culture (Brener, 1973; Zingales et al., 1997b). Because of this variability, it was first necessary to establish conditions under which epimastigotes of our CL Brener laboratory stock would differentiate to metacyclics, and to set up a system suitable for experimental manipulation.

3.1.1 Spontaneous differentiation in RPMI-1640 medium

Initially we observed the development of metacyclics in supplemented RPMI-1640 medium, which is the medium used for routine culture of epimastigotes. The development of different forms was quantified by taking samples at various time points, and counting Giemsa stained cells under the microscope as described in Materials and Methods, section 2.2.2. For a more detailed description of the morphological criteria by which cells were scored see also section 3.2 below.

Spontaneous differentiation in RPMI-1640 occurred as follows: During exponential growth the rapidly dividing epimastigotes are uniform in size and shape and less than 1% (typically none) of the cells are metacyclics or intermediate forms (i.e. forms with a kinetoplast in an intermediate position as described in section 3.2). In mid to late log phase, at a density of 4-5 x 10^7 cells ml^{-1}, epimastigotes become elongated and concomitantly metacyclics begin to appear. At the beginning of stationary phase (at 6-8 x 10^7 cells ml^{-1}), the average percentage of fully developed metacyclics is 5% (± 2%). Five independent cultures were left to incubate well into the stationary phase, and after 3 weeks yielded 4.5%, 19%, 27%, 33% and 41% metacyclics, respectively. After this prolonged incubation, however, accurate quantification becomes difficult, as cells start to show signs of degradation (such as missing kinetoplasts or nuclei, detached flagella, accumulating debris). The high variability in the proportion of metacyclics that are produced makes this an unsatisfactory system for experimental manipulation. Other conditions were therefore investigated.

82

3.1.2 Grace's insect medium supports rapid metacyclogenesis

Grace's Insect medium has previously been reported to enhance metacyclogenesis when added to CL Brener epimastigotes grown in liver infusion tryptone (LIT) medium (Zingales et al., 1997b). We found that transfer of epimastigotes at the beginning of exponential growth into a medium consisting of 80% RPMI-1640 and 20% Grace's Insect Medium (20% Grace's, section 2.2.2) led to rapid development of metacyclics (Figure 3.1A). On day six after transfer 24% (± 8.5) metacyclics were found in 20% Grace's, compared to 5% (± 2) metacyclics in cultures incubated in parallel in RPMI-1640 (mean value from 6 experiments ± standard deviation). No difference in growth rate was observed (Figure 3.1B).

The time-point when the maximal number of metacyclics in 20% Grace's was reached coincided with the end of the exponential growth phase. Longer incubation in 20% Grace's did not lead to a higher proportion of metacyclics (data not shown). Increasing the proportion of Grace's medium to 50% did not enhance the proportion of metacyclics, while culturing the cells in 100% Grace's without RPMI-1640 reduced the yield of metacyclics to 2.5%.
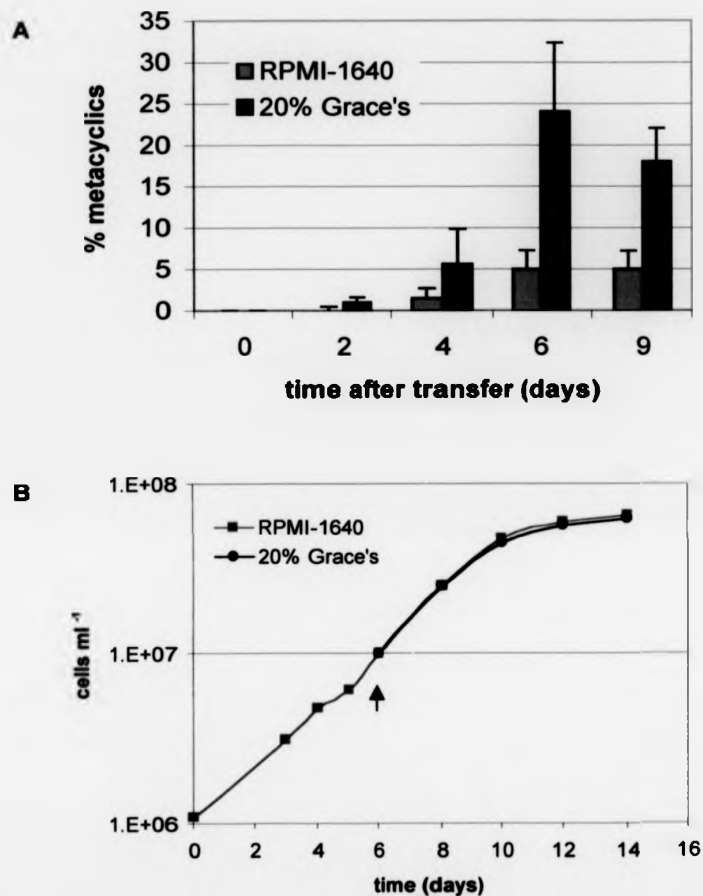
Figure 3.1: Timecourse of differentiation in 20% Grace's and RPMI-1640

T. cruzi Cl. Brener epimastigotes were grown in RPMI-1640 medium to a density of 1×10^7 cells ml^{-1} before transfer to fresh RPMI-1640 medium, with or without 20% Grace's. Samples were taken at the indicated time points. (A) The bars show the average proportion of metacyclics in each culture as determined from six experiments. Error bars indicate standard deviation. (B) Cell density vs. time in culture as measured in one representative experiment. The arrow indicates the timepoint of transfer to different media.

3.1.3 Triatomine Artificial Urine (TAU) does not induce
metacyclogenesis in CL Brener

The minimal medium TAU is routinely used to trigger metacyclogenesis in the
T. cruzi strain Dm28c, typically yielding 30-80% metacyclics (Contreras et al.,
1985a; Contreras et al., 1985b; Homsy et al., 1989; Krassner et al., 1990, see also
section 1.4.2). However, this method does not to work for all *T. cruzi* strains
(Krassner et al., 1990; Sanchez et al., 1990). To assess whether TAU would promote
metacyclogenesis in the CL Brener strain, cells were incubated according to the
procedure of Contreras, (Contreras et al., 1985b , section 2.2.2) and observed over a
period of up to three weeks.

Within the first 24 hours in TAU epimastigotes became long and thin compared to
control cells in RPMI-1640, and firmly attached to the culture flask. This is consis-
tent with the model proposed by Tyler and Engman (2000), that reduced levels of
glucose trigger attachment of epimastigotes to a hydrophobic surface (section 1.4.2).
However, very few unattached metacyclics developed in our experiments. The per-
centage of metacyclics in the cultures was determined after 10 days, and the follow-
ing values were obtained. 5% in TAU supplemented with 10 mM L-proline; 1% in
TAU supplemented with 2 mM D-glucose; < 1% in TAU supplemented with 0.035%
$NaHCO_3$. Over a period of three weeks the attached epimastigotes became even
longer and thinner. Cells remained motile, but no increase in cell number was
observed. When these epimastigotes were diluted back into RPMI-1640 after three
weeks in TAU, the epimastigotes resumed normal growth. Culturing CL Brener
epimastigotes in TAU thus resulted in the elongated morphology typical for starved
cells, and the cells ceased to proliferate, but did not differentiate.

3.1.4 cAMP analogues do not enhance the rate of metacyclogenesis

Several studies suggest that an increase in intracellular cAMP is involved in the
differentiation of trypanosomes (Fraidenraich et al., 1993; Gonzales-Perdomo et al.,
1988). Specifically, addition of cAMP analogues to *T. cruzi* epimastigote cultures
has been reported to enhance metacyclogenesis (Gonzales-Perdomo et al., 1988). We

attempted to replicate these findings. Epimastigotes were grown in supplemented RPMI-1640 medium as usual. Membrane permeable cAMP analogues (100 μM 8-Br-cAMP or 100 μM dibutyryl-cAMP) were added to mid log phase cultures. After continued incubation for six days, the proportion metacyclics was 5% in every culture. Exactly the same proportion of metacyclics developed in control cells grown in RPMI-1640 medium without the addition of cAMP analogues. Thus in our hands, the addition of exogenous cAMP to early log phase epimastigote cultures had no discernible effect on metacyclogenesis.

In summary, of the different conditions tested, only Grace's Insect Medium enhanced development of metacyclics in the CL Brener strain, consistent with previously published observations (Zingales et al., 1997b). Incubation in 20% Grace's medium was therefore used in all subsequent experiments to induce metacyclogenesis. Because neither TAU nor cAMP analogues enhanced differentiation above the level observed in RPMI-1640 in these preliminary experiments, these methods were not further investigated.

3.2 MORPHOLOGICAL MARKERS

Giemsa staining of trypanosomes (section 2.2.2) allows the visualisation of its over-all cell shape, and determination of the relative positions of kinetoplast, nucleus and flagellum. This easy method is traditionally used to distinguish different life-cycle stages, and was used routinely to determine the relative proportions of different cell types in cultures in this study. However, it is important to note here that the study of metacyclogenesis in the insect vector has always been complicated by the occurrence of various morphologically intermediate cell types of uncertain relationship, as detailed in the introduction (section 1.4.1). For example, Schaub (1989) described 18 morphologically distinct *T. cruzi* forms in the insect vector. Microscopic examination of CL Brener stationary phase *in vitro* cultures in 20% Grace's medium immediately revealed a similar situation: stationary phase cultures are mixed populations of various types of morphologically distinct cells.

epimastigote   metacyclic
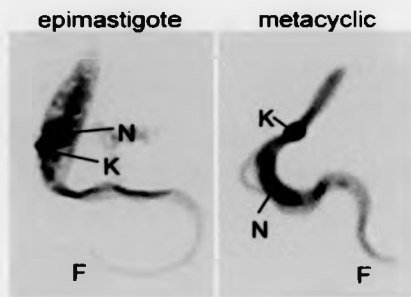
Figure 3.2: Epimastigotes and metacyclics in Giemsa stained preparations

T. cruzi epimastigotes and metacyclics are morphologically distinct cells. They can easily be distinguished in Giemsa stained preparations, as shown in the image. The epimastigote nucleus (N) is round and the kinetoplast (K) is located at the anterior end of the cell adjacent to the nucleus. The metacyclic nucleus is elongated, and the kinetoplast is located at the posterior end of the cell, away from the nucleus. The flagellum (F) is at the anterior end of the cell.

The key distinguishing criterion between epimastigotes and fully developed metacyclics is the position of the kinetoplast relative to the nucleus (Figure 3.2). In epimastigotes the kinetoplast is at the anterior end of the cell, in close proximity to the nucleus. In metacyclics the kinetoplast is at the posterior end of the cell with a gap between nucleus and the kinetoplast. Epimastigotes and metacyclics also differ in the shape of the nucleus (round in epimastigotes, elongated in metacyclics), the shape of the kinetoplast (disc shaped in epimastigotes and spherical in metacyclics) and the general shape of the cell body. Epimastigotes in stationary phase are more elongated than epimastigotes in log phase, but otherwise they are morphologically similar. A subset of cells in stationary cultures has an intermediate morphology. Based on two criteria, the position of the kinetoplast, and the shape of the nucleus, we established a simplified matrix that contains four intermediate types (Table 3.1) in addition to the classic epimastigote and metacyclic forms. We found this to be sufficient to describe all cells found in stationary cultures in 20% Grace's. Cells that did not fit into this matrix were encountered only very rarely. At present, the relationship between these forms, their replicative potential and commitment to differentiation is unclear.

Table 3.1: Morphologically distinct cell types in stationary T. cruzi *cultures*

T. cruzi CL Brener stationary cultures in 20% Grace's medium are mixed populations of morphologically distinct cell types. Based on the shape of the nucleus and the position of the kinetoplast, six types can be distinguished: epimastigotes, metacyclics (see also Figure 3.2) and four intermediate types. The intermediates are assumed to be in the process of differentiation, but the relationship between them is uncertain.

	round nucleus	*elongated nucleus*
anterior kinetoplast	**epimastigote**	intermediate 3
intermediate kinetoplast	intermediate 1	intermediate 4
posterior kinetoplast	intermediate 2	**metacyclic**

3.3 MOLECULAR MARKERS ASSOCIATED WITH METACYCLOGENESIS

3.3.1 Up-regulation of cruzipain activity in 20% Grace's medium

An increase in the activity of cruzipain, the major cysteine proteinase of *T. cruzi*, during the growth of epimastigotes, and a link between cruzipain activity and meta-cyclogenesis has previously been shown in the strain Silvio X10/6 (Tomas et al., 1997). We asked whether cruzipain activity also increased in the CL Brener strain during metacyclogenesis in Grace's medium. To test this, epimastigotes induced to differentiate in 20% Grace's medium were observed for nine days. The proteolytic activity associated with protein extracts taken at different time points during differentiation was detected in a qualitative assay, using the conditions optimised for detection of cruzipain activity (section 2.4.7). Detailed previous studies had assayed proteolytic activities in *T. cruzi* extracts at a range of pH values, and with different classes of protease inhibitors. These studies established that under the conditions employed here, the main proteolytic activity is derived from cruzipain. A weaker, higher molecular weight band corresponds to a developmentally regulated metallo-protease, which is optimally active at pH 9.0 (Tomas, 1995).
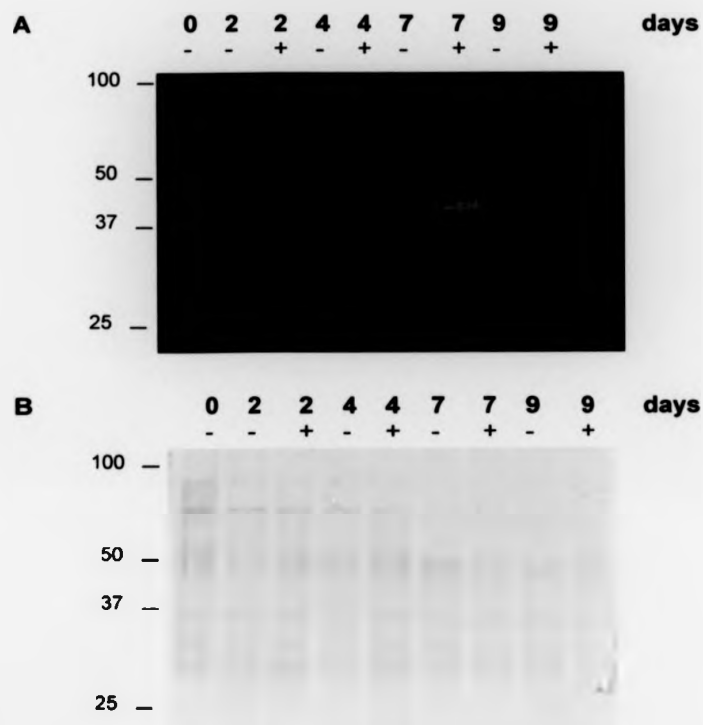
Figure 3.3: Detection of proteolytic activity in T. cruzi *extracts*

T. cruzi CL Brener epimastigotes were incubated in RPMI-1640 medium without (-) or with 20% Grace's (+). Cultures were incubated at 27°C/5% CO_2 for nine days. Samples were taken at the time points indicated above each lane. 4 μg of soluble protein was separated on a 10% SDS-PAGE gel, containing 0.1% gelatin. The gel in (A) was incubated in digestion buffer for 18 hours followed by staining of the gel with Coomassie. White bands appear where proteolytic activity has degraded the gelatin. (B) Equal loading in all lanes was confirmed by Coomassie staining of duplicate samples in a standard SDS-PAGE gel. Numbers on the left show molecular sizes in kilo Daltons.

Table 3.2: Cruzipain assay: cell density and proportion of metacyclics in samples

DAY	0	2	2	4	4	7	7	9	9
20% Grace's	-	-	+	-	+	-	+	-	+
cells ml^{-1} ($x\ 10^7$)	1.0	3.5	4.5	5.5	5.5	8.0	6.5	6.0	7.5
% metacyclics	< 1%	< 1%	2%	3%	5%	3.5%	18.5%	5%	15%

As the CL Brener cells reached stationary phase, increased proteolytic activity was clearly observed (Figure 3.3). Cruzipain activity increased over time in both cultures, with and without Grace's medium, with an apparent peak on day 7. However, cells in 20% Grace's medium, which produce more metacyclics, also showed a more pronounced up-regulation of cruzipain activity. The proportion of metacyclics rose from <1% to 5% over nine days in RPMI-1640. In 20% Grace's the proportion of metacyclics reached a peak of 18.5% on day 7, dropping slightly to 15% on day 9 (summarised in Table 3.2). This confirms that up-regulation of cruzipain activity correlates with an increased rate of metacyclogenesis. The greatest difference between the cells in RPMI-1640 and 20% Grace's medium is seen on day 4, when the difference in the number of metacyclics is still small. Up-regulation of cruzipain activity is thus an early marker for metacyclogenesis.

3.3.2 Changes in protein phosphorylation

Reversible protein phosphorylation is an important regulatory mechanism in many signal transduction pathways (section 1.5.1). The kinase activities and phosphorylation profiles associated with *T. cruzi* metacyclogenesis have not yet been studied. The initial objective of this project was to dissect the cAMP signalling pathway of *T. cruzi* and its role during metacyclogenesis. However, there are now considerable doubts regarding the results that suggested a role for cAMP in *T. cruzi* differentiation (section 1.5.2). Our own attempts to reproduce the finding that membrane permeable cAMP analogues induce metacyclogenesis yielded negative results (see above).

To begin a dissection of *T. cruzi* signalling pathways that might regulate meta-cyclogenesis, we sought to identify stage-specific protein phosphorylation. First, the pattern of protein phosphorylation in epimastigotes was analysed, and then it was compared with the pattern observed during metacyclogenesis.

In order to do this, commercially available anti-phosphoserine (PS) anti-phos-phothreonine (PT) and anti-phosphotyrosine (PY 4G10) antibodies were used to detect phosphorylated residues on 1D Western blots of *T. cruzi* protein extracts as described in sections 2.4.1 and 2.4.5. We optimised conditions for immunochemical detection of phosphorylated *T. cruzi* proteins, and established the pattern of phos-phorylation in wild type CL Brener epimastigotes and stationary cultures that contain metacyclics.

3.3.3 Phosphorylated proteins in *T. cruzi* CL Brener epimastigotes

Threonine phosphorylation. Numerous protein bands between 37 and 250 kDa were detected in both the insoluble and soluble of parasite lysates (Figure 3.4A). In the soluble fraction a band of about 37 kDa, several bands between 60 and 100 kDa and a band of about 150 kDa gave the strongest signals. Between experiments the overall pattern remained unchanged but the relative signal intensities of the different bands did vary. In particular the 37 kDa band gave a very strong signal in some but only a weak signal in other blots.

Serine phosphorylation. Two strong bands of about 75 and 30 kDa and a weak band of about 60 kDa were detected in the soluble fraction of epimastigote extracts. No bands are detected in the insoluble fraction (Figure 3.4B).

Tyrosine phosphorylation. A strong band of about 250 kDa was invariably detected in the insoluble fraction. In most blots, no bands were detected in the soluble fraction (Figure 3.4C). Occasionally, two weak bands (30 kDa and 40 kDa) were detected in the soluble fraction (not shown).
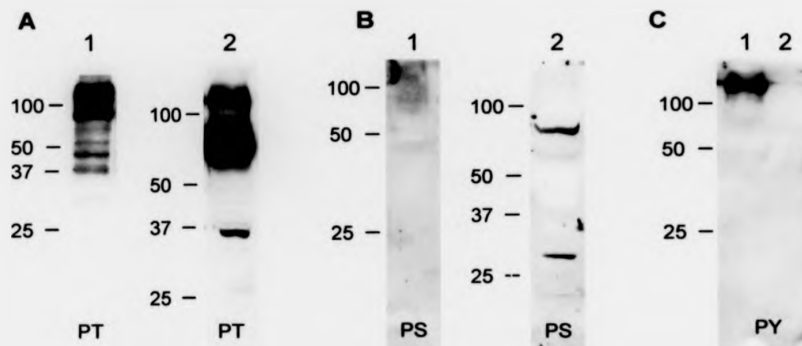
Figure 3.4: Western Blot analysis of phosphorylated epimastigote proteins
Western blots of protein from log phase epimastigotes were probed with antibodies against phospho-rylated residues. (A) anti phospho-threonine (PT); (B) anti phospho-serine (PS); (C) anti phospho-tyrosine (PY). Lane 1: insoluble protein; Lane 2: soluble protein (20 µg per lane). Numbers on the left indicate molecular sizes in kilo Daltons.

3.3.4 Phosphorylation of low molecular weight proteins
during metacyclogenesis

To obtain a first overview of changes in protein phosphorylation that occur during metacyclogenesis, protein extracts from cultures containing different proportions of metacyclics were compared to pure epimastigote extracts, using the same method as in the experiments described above.

No changes in the pattern of serine and tyrosine phosphorylation were detected in extracts from differentiating populations in Western blots of protein extracts sepa-rated on 1D SDS-PAGE gels (not shown). However, distinct changes in the pattern of threonine phosphorylation were observed. Several strongly phosphorylated low molecular weight bands (between 20 and 30 kDa) appeared in the soluble protein fraction from populations with >20% metacyclics. These bands were absent in epi-mastigote cultures, and appeared only weakly in an extract from a population con-taining 6% metacyclics (Figures 3.5A). In the insoluble fraction a signal of around 20 kDa was detected in an extract from a population containing 23% metacyclics but

not in extracts from populations with 14% or less metacyclics (Figure 3.5B). The identity of these proteins remains to be determined.

Unfortunately, it was not possible to analyse pure metacyclic preparations for technical reasons. Complement mediated lysis is a commonly used method to eliminate epimastigotes from mixed populations in other *T. cruzi* strains (Nogueira et al., 1975). This works because in those strains resistance to complement mediated lysis is developmentally regulated. This method could not be used here because CL Brener epimastigotes are resistant. Attempts to purify the CL Brener metacyclics from 20% Grace's medium by anion exchange chromatography (Alvarenga and Brener, 1979; Orr et al., 2000) yielded unsatisfactory results. Only a slight enrichment of metacyclics was achieved, and a large proportion of cells was lost in the procedure.
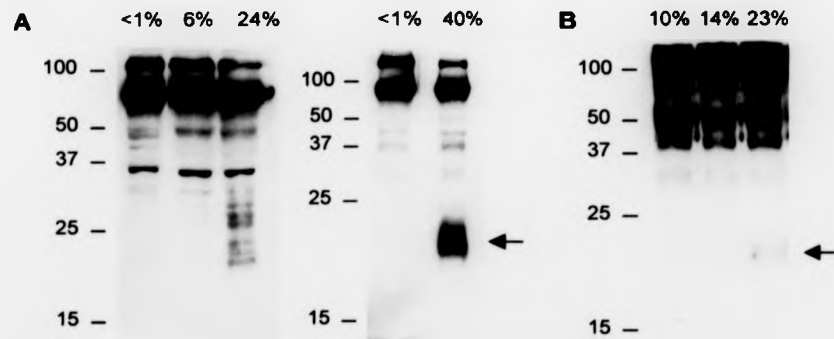
Figure 3.5: Western Blot analysis of phosphorylated proteins in differentiating populations
Protein extracts from *T. cruzi* CL Brener populations with different proportions of metacyclics were analysed on Western blots using the anti phospho-threonine antibody PT. (A) soluble protein (20 μg per lane) (B) insoluble protein fraction. The percentage of metacyclics in the sample is indicated above each lane. Phosphorylated proteins of approximately 20 kDa are detected only in extracts containing a high percentage of metacyclics (arrows). Numbers on the left indicate molecular sizes in kilo Daltons.

In summary, we have found that in epimastigote extracts numerous proteins are phosphorylated on threonine residues, but only few are phosphorylated on serine or tyrosine. Several, as yet unidentified low molecular weight proteins become phosphorylated on threonine during metacyclogenesis. These results indicate that a proteomics approach is likely to identify novel stage regulated proteins or protein modifications.

3.4 DISCUSSION

There are many conflicting reports in the literature as to the conditions that induce metacyclogenesis in *T. cruzi* (Bonaldo et al., 1988; Contreras et al., 1985b; de Isola et al., 1981; Homsy et al., 1989; Sullivan, 1982; Ucros et al., 1983). Several factors have been found to enhance the process, including starvation conditions (Contreras et al., 1985b; Krassner et al., 1990; Miles, 1993), transfer of epimastigotes to Grace's Insect medium (Heath et al., 1990; Sullivan, 1982; Zingales et al., 1997b) and the addition of cAMP analogues (Gonzales-Perdomo et al., 1988). Results obtained with one *T. cruzi* strain however, are often not generally applicable (Krassner et al., 1990; Sanchez et al., 1990), and the molecular mechanisms and signal transduction pathways that control the differentiation process are poorly understood (Parsons and Ruben, 2000).

Different methods were tested to study the conditions under which *T. cruzi* CL Brener, the genome reference strain, would differentiate reproducibly *in vitro*. Only two conditions supported the development of metacyclics. The best *in vitro* differentiation rates were obtained with Grace's Insect medium. Transfer of epimastigotes in the mid-log phase of growth from RPMI-1640 to a medium containing 20% Grace's Insect medium enhanced metacyclogenesis considerably. Up to 25% metacyclics were routinely obtained within six to eight days, while only 5% metacyclics developed in RPMI-1640 medium over the same time. The second method that typically resulted in a relatively high proportion of metacyclic cells is simply to let stationary phase cultures in RPMI-1640 medium incubate for several weeks. However, this method has two disadvantages. First, the significant number of dead cells, which accumulate in old stationary phase cultures, complicate analysis of any experiment.

Second, the time course of development of metacyclics cannot be predicted with the accuracy required to design experiments. The proportion of metacyclics produced varies enormously (between 4 and 40%), and it takes several weeks for development to occur.

It is important to note that Grace's medium did not increase the maximal number of metacyclics that could be obtained. The advantage of using 20% Grace's medium to induce metacyclogenesis in CL Brener is that development occurs much faster and that the cells follow a predictable time course of development. Six to eight days after addition of Grace's medium the maximum number of metacyclics will have accumulated. These cultures contain epimastigotes, metacyclics, and a number of intermediate forms of defined morphologies (Figure 3.2 and Table 3.1), without the decaying or dead cells that are observed in older stationary phase cultures. Using 20% Grace's medium is therefore the best method currently available for the purpose of comparing the effect of experimental manipulation on the rate of metacyclogenesis in CL Brener.

Still, this system is not ideal, as a yield of 25% metacyclics is too low to obtain clean results when doing comparative biochemistry, for example. The rationale behind using the CL Brener strain for this study was to work with the same model strain that was used for sequencing the genome. However, if higher rates of metacyclogenesis cannot be achieved using the CL Brener strain, then one must consider using a different strain for future work in this area. Ideally, a strain should be sought that differentiates well in the defined minimal medium TAU. That system is likely to be more reproducible and better suited for biochemical analysis than complex and undefined media. Strain Dm28c should be considered, as this strain consistently yields high proportions of metacyclics in TAU medium (Contreras et al., 1985b). The strain Silvio X10/6 should also be tested, because a cell line stably transfected with the genes for T7 polymerase and Tet repressor gene has already been generated in this laboratory . This cell line can be used for inducible gene expression (Martin Taylor, unpublished, see also section 1.7.1), which will be an invaluable tool for genetic research.

To dissect metacyclogenesis further, it will be important to untangle the relationship between the different morphological forms. This is still unclear, despite a number of detailed studies that have addressed this question (Brack, 1968; Schaub, 1989). Specifically, it is not clear in which order the observed morphological changes occur, whether each of the intermediate forms gives rise to a fully developed metacyclics, and which forms are infective to mammalian cells. Molecular markers should help to clarify the relationship between these morphologically distinct forms.

Cysteine protease activity has been implicated in the process of metacyclogenesis in the Dm28c strain (Bonaldo et al., 1991) and the Silvio X10/6 strain (Tomas et al., 1997). There are strain-specific differences in the proteolytic activities of epimastigotes (Greig and Ashall, 1990), but that study did not include the CL strain. We therefore investigated the proteolytic activity in CL Brener epimastigotes and stationary cultures. The results obtained confirm that cruzipain activity increases when CL Brener epimastigotes enter stationary phase, similar to the results obtained with the X10/6 strain (Tomas et al., 1997). Interestingly, purified metacyclics of the Dm28c strain showed very low cysteine protease activity (Bonaldo et al., 1991). Up-regulation of cruzipain activity thus appears to occur transiently during metacyclogenesis, but is not a marker for fully developed metacyclics.

The analysis of phosphoproteins in whole parasite extracts using antibodies against phosphorylated residues suggests that threonine-phosphorylation of several low molecular weight proteins is associated with metacyclogenesis. Using proteomics technology (section 1.7.2), there are two directions in which this potentially very fruitful area can be followed up: (i) a comprehensive analysis of the phosphorylation patterns associated with different *T. cruzi* life-cycle stages, and (ii) identification of the phosphorylated low molecular weight proteins associated with metacyclogenesis.

There are a number of techniques that allow a more detailed analysis of the phosphorylation profiles of log phase epimastigotes and metacyclics. A common approach used to study global protein phosphorylation is to grow the cells in medium containing inorganic ^{32}P. Lysates from these cells can then be separated on 2D gels and analysed by autoradiography to compare phosphorylation patterns from different

96

cell types. A protocol for the separation of trypanosome proteins by 2D gel electrophoresis has recently been published (van Deursen et al., 2003), which may be used as a basis to initiate such a large scale proteomics analysis in *T. cruzi*. To increase the sensitivity and detect proteins that are not abundant, it may be necessary to prepare samples that are enriched in phosphoproteins (Oda et al., 2001; Zhou et al., 2001).

Using mass spectrometry approaches it should be feasible to identify the nature of the low molecular weight phosphoproteins detected in the present experiments, and determine the precise site of phosphorylation (Ficarro et al., 2002; Soskic et al., 1999). The identification of these proteins would no doubt shed some light on the signal transduction pathways active in stationary *T. cruzi* cells. To pin down components that regulate metacyclogenesis, it will be particularly important to subsequently identify the corresponding protein kinase(s) by biochemical methods.

A global analysis of changes in phosphorylation patterns during metacyclogenesis may also help to clarify the role of cAMP signalling. The involvement of the cAMP signalling pathway in development of metacyclics is frequently cited as fact (Tyler and Engman, 2001). However, a closer scrutiny of the literature on *T. cruzi* metacyclogenesis reveals that the evidence for an involvement of cAMP is rather scarce (sections 1.4.2). Previous studies often based their conclusions on the effect of substances that activate or inhibit enzymes of the mammalian cAMP signalling pathway, such as the adenylyl cyclase activator forskolin, or the phosphodiesterase inhibitor IBMX. None of theses studies provide any supporting molecular data. Recently, cloning and analysis of the relevant genes from trypanosomes revealed important biochemical differences between the trypanosomatid enzymes and their mammalian counterparts (Bieger and Essen, 2001; Taylor et al., 1999; Zoraghi et al., 2001, see also section 1.5.3). The older data thus clearly needs to be reassessed.

Gonzales-Perdomo (1988) also reported that the simple addition of cAMP analogues enhanced the development of metacyclics directly. Our attempts to replicate these findings failed to confirm this. The addition of membrane permeable cAMP analogues to epimastigotes in RPMI-1640 or TAU did not enhance metacyclogenesis above the level observed in RPMI-1640 medium alone. These preliminary experi-

97

ments thus provided no support for the notion that cAMP is a trigger for metacyclo-genesis.

Metacyclogenesis occurs under conditions of stress and starvation. It is conceivable that cAMP signalling plays a role in these processes. How a starvation response is linked to differentiation is currently unknown. Clearly, there is scope for much interesting work in this area. The demonstration of developmentally regulated protein phosphorylation could be a first step towards dissecting the signal transduction pathways involved.

4. THE *MET3* GENE

A first step in dissecting the molecular mechanisms that control development of *T. cruzi* is to identify genes that are differentially expressed in different life-cycle stages. Goldenberg and co-workers have identified several genes that are upregulated during *in vitro* metacyclogenesis of *T. cruzi* strain Dm28c, using a method of selective amplification of RNA transcripts (Krieger et al., 1999; Krieger and Goldenberg, 1998). Initial characterisation of a small number of transcripts has confirmed that the corresponding proteins are expressed during metacyclogenesis in the Dm28c strain (Avila et al., 2001; Dallagiovanna et al., 2001; Fragoso et al., 2003). The specific function of these gene products has not been elucidated. Development of genetic tools for *T. cruzi* has progressed significantly in recent years and provides scope for experimental manipulation of *T. cruzi* genes (Taylor and Kelly, 2002). We have used transfection-based approaches to characterise in detail two of these stage-regulated genes, *MET2* and *MET3*. These genes show no apparent relatedness to previously characterised genes, and may therefore shed light on novel aspects of metacyclo-genesis in *T. cruzi*.

This chapter is divided into three sections. Section 4.1 describes the identification and detailed sequence analysis of the *MET3* gene from the genome sequencing strain CL Brener. Section 4.2 reports the results of confocal microscopy studies to determine expression and localisation of the MET3 protein. Stage-regulated nuclear localisation of native MET3 was observed. MET3 associates with the nucleolus, and we identified MET3 sequence elements that can direct the protein to the nucleolus. Section 4.3 describes the generation and analysis of *MET3* null mutants. The results from their analysis show that *MET3* is not essential for *T. cruzi* metacyclogenesis.

4.1 CHARACTERISATION OF *MET3* FROM *T. CRUZI* CL BRENER

4.1.1 Identification of two *MET3* alleles from CL Brener

We identified the CL Brener homologues of *MET3* by using the cDNA sequence from the Dm28c strain (metaciclina III, GenBank AC:AF312553) to screen DNA sequence databases. A search of GenBank initially identified a single open reading frame (ORF) on a 21 kb genomic contig from CL Brener (GenBank AC:090078). This ORF was named *MET3-CL1*. Subsequent searches of the *T. cruzi* CL Brener genome shotgun sequences (section 2.5) with the *MET3* sequence identified a second allele that we called *MET3-CL2*. The names for the CL Brener genes were chosen according to the recommended nomenclature for trypanosome genes (Clayton et al., 1998). The *MET3* gene from strain Dm28c is referred to in the literature alternatively as *Met-I* (Avila et al., 2003), *metacyclin-III / Met-III* (Yamada-Ogatta et al., 2004) or *metaciclina III* (GenBank AF312553). For consistency, and to clearly distinguish it from the CL Brener genes, we refer to it in this text as *MET3-Dm28c*.

A comparison of the three sequences reveals considerable sequence variation (Figure 4.1). Sequence identity between the two CL Brener alleles is only 97%, and between the CL Brener alleles and *MET3-Dm28c*, the sequence identity is 97% and 96%, respectively. A similar degree of sequence divergence between CL Brener alleles has been described for other genes (Bromley et al., 2004; Tran et al., 2003) and is likely to be a consequence of the hybrid nature of this strain (Machado and Ayala, 2001; Pedroso et al., 2003; Sturm et al., 2003). Surprisingly, the length of the *MET3* coding sequence is different in the two *T. cruzi* strains. This difference is caused by a single nucleotide insertion at position 570 in the coding sequence of both CL Brener alleles (Figure 4.1), which causes a frameshift relative to the *MET3-Dm28c* coding sequence. *MET3-CL1* and *MET3-CL2* extend over 576 bp, terminating at the stop codon TAA at position 574-6. *MET3-Dm28c* extends over 600 bp, terminating at stop codon TAG at position 598-600. To exclude the possibility that this frameshift was due to a sequencing error in the CL Brener contig, we compared all available shotgun sequences (from different sequencing centres) and found that the frameshift relative to *MET3-Dm28c* is present in all the relevant CL Brener sequences.

The CL Brener sequences were also confirmed experimentally, by sequencing PCR fragments derived from genomic DNA, and RT-PCR fragments, derived from cDNA templates (section 2.3.2 and 2.3.3). Sequencing cloned cDNA also identified the 5' splice-acceptor site in *MET3-CL1*: the spliced leader sequence (SL) is added to the RNA at an AG dinucleotide, 59 bases upstream of the start codon. The cDNA sequence of *MET3-Dm28c* deposited in GenBank also contains the SL, and it is spliced at the same position as *MET3-CL1*. As yet we have not identified the splice acceptor site for *MET3-CL2*.

Figure 4.1 (overleaf): Alignment of MET3 *DNA sequences*

The sequences designated *MET3-CL1* and *MET3-CL2* correspond to the genomic sequence of the two *MET3* alleles from *T. cruzi* CL Brener. The sequences were identified in searches of GenBank and GeneDB, using *MET3-Dm28c* as query. *MET3-CL1* is contained on cosmid AC090078. *MET3-Dm28c* corresponds to the *metaciclina III* cDNA sequence (GenBank AC:AF312553). The spliced leader sequence is not shown. The splice acceptor site (underlined) is at the same position in *MET3-CL1* and *MET3-Dm28c*. Start and stop codons are marked with bold letters. One nucleotide insertion at position 570 in the CL Brener sequences (highlighted) causes the reading frame to shift relative to *MET3-Dm28c*. This could be due to a sequencing error in *MET3-Dm28c*. Note how the poly-A stretch tail at the 5' end of the *MET3-Dm28c* sequence (putative poly A-tail) aligns with an A rich stretch in the genomic sequences of CL Brener.

```
CL1    -119 CATTCGGTGCGTTTATTTATTTTTTCTTCTTCTTGTTTCTCCGCTGCGTCACACAATAAT
CL2    -116 ...............G.........CT...................---
Dm28c

CL1    -59 AGACATTGAGGAAAGAAGCCACGGGGGGTGGCTCCAAAGGCATAAGAGGTGGCGGAAGTA 1
CL2    -59 ...........................................................A 1
Dm28c  -57 ...........G...............................................A 1

CL1        TGTCGCGACGCAAAGATAACAAATCCCACCTGAAAGAATTCCTCCGTCGACTTCGCGATC 61
CL2        TG.....................................G................... 61
Dm28c      TG........................................................ 61

CL1        AGGGCCGTCAGGGAACGCCTCAGGCATCGGCCGGGCGGGGGAGAGCACGCACGGGTGAAG 121
CL2        ............................................G............. 121
Dm28c      ............................................G............. 121

CL1        GCAGCGGACGGTACGCTCTCTCCTTTTCAGGGCGTTCCGTGGTGGCAACCGCCCCCACAC 181
CL2        ....A...................................................... 181
Dm28c      .......................T.........................G......... 181

CL1        GTGGCCGCTCCTATCTCCCACGCTCGCGCAAGGATAACGAAGATGACAGCATGCCTCTTG 241
CL2        ..................................C......................A 241
Dm28c      ..................................C......................A 241

CL1        CTGTTGCCAATGCCGCCTTGTCGTCAACGGATGACGCGCGGGATCAAGGACGGGAGCATC 301
CL2        .......G......T.........................G................. 301
Dm28c      ...C.A....G........................GA................... 301

CL1        AGCGTCGCCACCGTGACGAGGAGACAAGGACGTCAGAGGAGAGTTCACGGAGGCGCCCTT 361
CL2        .........T......................G........C............... 361
Dm28c      ........................................C............... 361

CL1        CTCGCTCTACGGAAAGCGACCCACATGAGGTCGTGTTGTCAGAGAGCCGCGCGTCTGTCC 421
CL2        ..........................C..T........................... 421
Dm28c      ................A...................C................... 421

CL1        TCCCGTCCACATCGACAGGGAGCCGAGCAAGCGGCAGCACTGATCCCGTTTCCGAACACG 481
CL2        ....T...........................A....................... 481
Dm28c      ..............................C.....T................... 481

CL1        CTGCAAGGCGGGCGGACGCGGGACGTGTAATAGGCAGAAGGAGAGGCGGGGACAGTGAGA 541
CL2        ...........T...A..A......G...G........................... 541
Dm28c      ...........T.........G................................... 541

CL1        AGAGTCTAAAGCGCCCGAGGACAGGTGGGCCCTAACGTGAATTGTCGTTCCCATCATTAG 601
CL2        .............................G...TAA....................... 601
Dm28c      .......................-.............C...............TAG 600

CL1        CAGTTTCTGCGCCAATTGAGGGCAAATAATGAGAGGGAGTAAACGGCAGCAAAGAAACAA 661
CL2        ...................................A..........            647
Dm28c      ......................................................... 660

CL1        GAAAAAGATTCAAATAGCAACGAGTGTGGAAAACGCAAATAAATTATTTCGCGTGTCGA 721
CL2
Dm28c      ......................................................... 720

CL1        GGGCTCCAAAACCTCATTGCTGCATCTGATTCTTTTTTTTGTGTCTG--TGCATAATTGG 779
CL2
Dm28c      ........................................G..TG........... 780

CL1        CAAAGAGGGGGGAACCAGTAAAAAAGAAACAAAA                          813
CL2
Dm28c      ......A.......A...........A...A.....                       814
```

Figure 4.1

102

4.1.2 Genomic organisation of *MET3*

A contiguous 54 kb fragment of genomic DNA from *T. cruzi* CL Brener, spanning the MET3 locus was assembled using two contigs from the genome sequencing project (TSKTSC_8550 and AC090078). Analysis of the sequence, using the Artemis software (Rutherford et al., 2000), showed that *MET3* is a single copy gene, in an area of high gene density. A BLAST search of GenBank and GeneDB, identified 22 putative coding sequences (CDS) within this contig – one for every 2.5 kb. All genes are encoded on the same strand of DNA. The average GC content is 47.6%. *MET3* is flanked by two CDS with predicted functions (Figure 4.2). A CDS that shows high similarity to delta-1-pyroline-5-carboxylate dehydrogenases (P5CDH; enzyme that catalyses the second step in conversion of proline to glutamate) ends 522 nucleotides upstream of *MET3*. A putative protein kinase gene (similar to MAP kinases) is located 2.8 kb downstream of *MET3*. The intergenic sequences up- and downstream of *MET3* contain long runs of single nucleotides (mainly T) and dinucleotide repeats.

Gene order in kinetoplastids is highly conserved (Bringaud et al., 1998; Ghedin et al., 2004). A comparison of the *T. cruzi* genomic contig with sequences from other kinetoplastids identified regions on *T. brucei* chromosome 10 and *L. major* chromosome 3 that show a high level of synteny in the genes flanking *MET3*. They do not encode any gene similar to *MET3* itself (Figure 4.2). No similarity between the nucleotide sequences of *T. cruzi*, *T. brucei* and *L. major* were found in pair-wise comparisons using the sequence in between *P5CDH* and *MAPK* of *T. cruzi* (3.8 kb, including *MET3*), *T. brucei* (4.5 kb) and *L. major* (1.9 kb). More significantly, BLASTX searches that translate the *T. brucei* and *L. major* nucleotide sequences in all six reading frames found no amino acid sequences similar to any part of MET3.
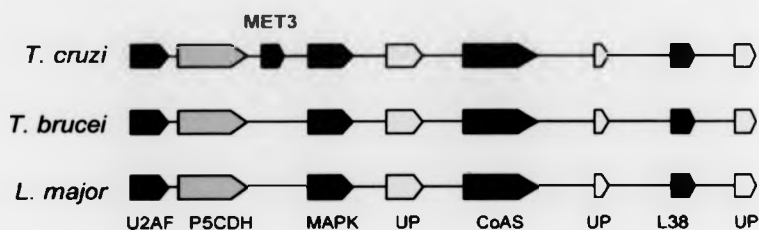
Figure 4.2: Genomic organisation of MET3 *and comparison of syntenic loci*

The diagram shows a schematic alignment (not drawn to scale) of ~17 kb surrounding the *MET3* locus with syntenic loci from *L. major* chromosome 3, and *T. brucei* chromosome 10. *MET3* is unique to *T. cruzi*, while the genes flanking *MET3* are preserved in al three species. The predicted genes (arrows) code for the following. *U2AF*: U2 snRNP auxiliary factor small subunit; *P5CDH*: delta-1-pyroline-5-carboxylate dehydrogenase; *MAPK*: protein kinase; *UP*: hypothetical protein; *CoAS*: fatty acyl CoA synthetase; *L38*: ribosomal protein.

In silico analysis suggests that *MET3* is a 'single copy' gene, with two allelic variants present in the genome of *T. cruzi* CL Brener. This was confirmed experimentally by Southern blot analysis of genomic DNA (Figure 4.3A). DNA was digested with ten different restriction enzymes and probed with the *MET3* CDS. Each digest results in either a single band or a doublet, where both bands are of equal intensity. This suggests that restriction site polymorphisms are present that cause small variations in the size of restriction fragments derived from each locus. Furthermore it suggests that only two allelic copies of *MET3* are present, despite the partially trisomic nature of the CL Brener genome (unpublished observations). The hybridisation pattern obtained with a fragment from the upstream gene *P5CDH* is consistent with this interpretation. *P5CDH* is also a single copy gene with two allelic variants, and the restriction site polymorphisms are consistent with the pattern obtained when hybridising with *MET3*. The information derived from Southern blot analysis, and from the genomic sequences, was used to draw a detailed restriction map of an 18 kb region surrounding the *MET3* gene (Figure 4.3B). This facilitated the gene deletion experiments outlined in section 2.3. The *MET3* locus was mapped to a 2 Mb chromosome by pulse field gel electrophoresis (Figure 4.4).
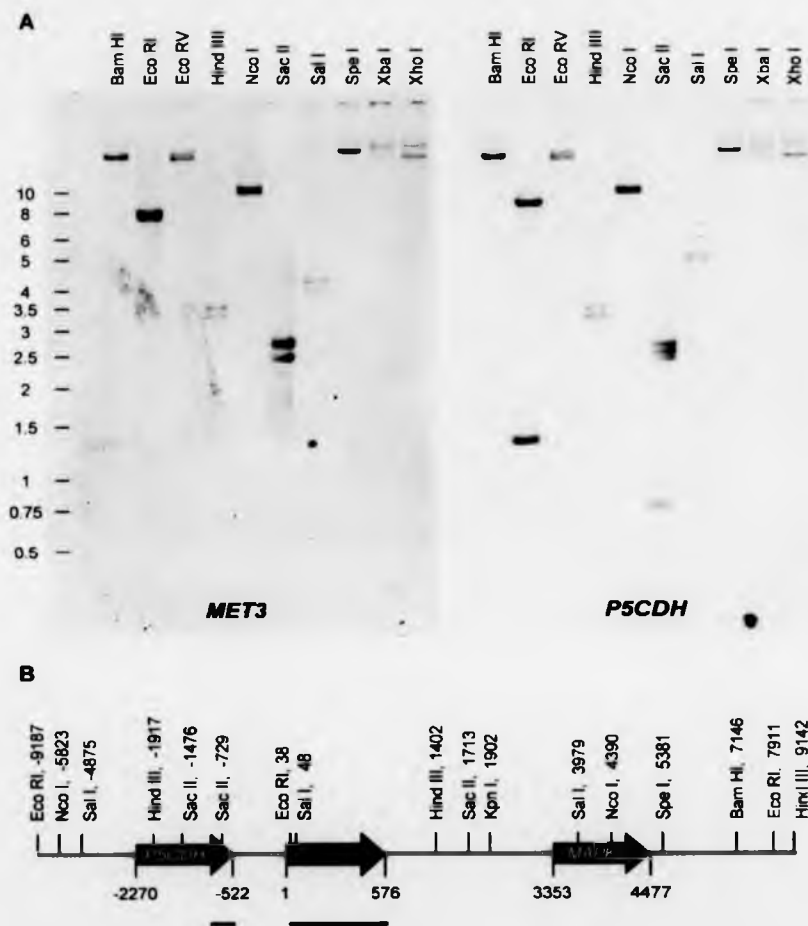
Figure 4.3: Restriction map of the MET3 *locus*

(A) Genomic DNA from *T. cruzi* CL Brener was digested with restriction enzymes as indicated above each lane. The banding pattern on a Southern blot probed with *MET3* or a fragment of *P5CDH* (the gene upstream of *MET3*) is consistent with single copy genes. Two bands are seen in most lanes because of restriction site polymorphisms between allelic loci. Numbers on the left indicate size in kb. (B) A restriction map of the *MET3-CL1* locus, drawn based on information from the genome sequence and experimentally confirmed by the Southern blot. Numbers indicate nucleotide positions. The first nucleotide in the *MET3* start codon was designated position '1'. Black lines show location of probes used for hybridisation.
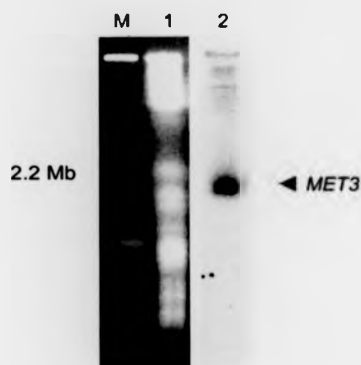
Figure 4.4: MET3 is localised on a 2 Mb chromosome
Chromosomes of *T. cruzi* CL Brener were separated by CHEFE (lane 1). A blot of the gel probed with the *MET3* coding sequence shows that it is localised on a chromosome of approximately 2 Mb (lane 2). *Saccharomyces cerevisiae* chromosomes were run as size markers (lane M).

4.1.3 *MET3* is found only in *T. cruzi*

Extensive searches of all available kinetoplastid databases and of GenBank (section 2.5) failed to identify any *MET3* homologues. Nor were any partially similar sequences identified that could be used to assign a biochemical function to the MET3 protein. Only sequences with spurious similarities were identified in a BLAST search of kinetoplastid sequences in GeneDB (version 2.1;section 2.5). Equally, none of the 'hits' in BLAST searches of other sequence databases is likely to be related to *MET3*, based on sequence alignment scores. The most recent search of the EMBL data yielded no hits better than an expect value $E = 3.0$. Interestingly, however, these weak hits contain a linker histone protein (HMG-I/Y) from *Arabidopsis thaliana* (40% identity over 40 amino acids). This may be significant in light of our experimental results, which show that the MET3 protein localises to the nucleus (section 4.2). These results and the comparative analysis of the *MET3* locus (Figure 4.2) show that *MET3* is clearly absent from *T. brucei* and *L. major*, and appears to have no homologues in any other organism sequenced to date.

4.1.4 MET3 protein sequence analysis predicts nuclear localisation
and possible nucleic-acid binding capacity

In order to extract as much information as possible from the predicted protein sequence of MET3, we utilised a wide range of bioinformatic analysis tools (section 2.5). *MET3* codes for a protein with a calculated molecular weight of 21 kDa and a pI of 11.4. The protein contains a high number of serine residues, and charged amino acids (Table 4.1). An alignment of the derived amino acid sequences for MET3-CL1, MET3-CL2, and MET3-Dm28c is shown in Figure 4.5. MET3-CL1 and MET3–CL2 are eight amino acids shorter than MET3-Dm28c, due to the frameshift in the *MET3-Dm28c* gene. Non-synonymous nucleotide substitutions cause an additional eight to twelve (depending on the allele) amino acid substitutions throughout the protein (Figure 4.5). All three sequences were subjected to bioinformatic analysis as described in materials and methods. Despite small differences in primary sequence, the three sequences are the same with respect to the features described below.

Table 4.1: Properties of MET3 amino acid sequence

Properties of the deduced amino acid sequences of MET3-CL1, MET3-CL2 and MET3-Dm28c. These proteins contain high numbers of serine residues and charged amino acids (mainly arginine). The difference in molecular weight is due to a frameshift in the sequence of the *MET3-Dm28c* gene.

	MET3-CL1	MET3-CL2	MET3-Dm28c
number of amino acids	191	191	199
molecular weight	21 kDa	21 kDa	22 kDa
theoretical pI	11.35	11.46	11.41
negatively charged amino acids (D+E)	26 (13.6%)	26 (13.6%)	25 (12.6%)
positively charged amino acids (R+K)	41 (21.5%)	42 (22%)	43 (21.6%)
arginines (R)	35 (18.3%)	36 (18.8%)	36 (18.1%)
serines (S)	29 (15.2%)	29 (15.2%)	30 (15.1%)

```
MET3-CL1     MSRRKDNKSHLKEFLRRLRDQGRQGTPQASAGRGRARTGEGSGRYALSFSGRSVVATAPT 60
MET3-CL2     ...............V........................................... 60
MET3-Dm28c   ........................................................... 60


MET3-CL1     RGRSYLPRSRKDNEDDSMPLAVANAALSSTDDARDQGREHQRRHRDEETRTSEESSRRRP 120
MET3-CL2     ............T.......T..D..........G......................... 120
MET3-Dm28c   ..................TATK...........G......................... 120


MET3-CL1     SRSTESDPHEVVLSESRASVLPSTSTGSRASGSTDPVSEHAARRADAGRVIGRRRGGDSE 180
MET3-CL2     ........D....................T..............V.TR.GM........ 180
MET3-Dm28c   .......T....................................V....G......... 180


MET3-CL1     KSLKRPRTGGP*          191
MET3-CL2     ...........*          191
MET3-Dm28c   ...........NVNCRSHH*  199
```

Figure 4.5: Alignment of MET3 protein sequences

The alignment compares the deduced MET3 amino acid sequences of the two alleles from *T. cruzi* CL Brener with the *T. cruzi* Dm28c protein (GenBank AAK37449). A dot (.) denotes amino acids identical to the MET3-CL1 sequence. A star (*) denotes a stop codon. Note that the predicted MET3-Dm28c protein is longer, due to a frameshift in the DNA sequence.

```
MET3  MSRRKDNKSHLKEFLRRLRDQGRQGTPQASAGRGRARTGEGSGRYALSFSGRSVVATAPTRGRSYLPRSRKDNEDDSMPL
PKC   <->                                            <->      <->                 <->
CK2                                                                                <-->
ASN         <-->
MYR                         <---->

MET3  AVANAALSSTDDARDQGREHQRRHRDEETRTSEESSRRRPSRSTESDPHEVVLSESRASVLPSTSTGSRASGSTDPVSEH
PKC                                      <  >
CK2       <-->                        <-->       <->
CAMP                                     <->
MYR                                                                            <---->

MET3  AARRADAGRVIGRRRGGDSEKSLKRPRTGGP
PKC                   <-><->
AMI         <-->
```

Figure 4.6: Distribution of potential modification sites in the MET3 protein

Arrows indicate the positions of potential modification sites in the MET3 protein sequence. A Prosite scan identified potential phosphorylation sites for protein kinase C (PKC), casein kinase II (CK2) and cAMP- and cGMP-dependent protein kinase (CAMP). Putative glycosylation (ASN), N-myristoylation (MYR) and amidation sites (AMI) were also identified.

Figure 4.7: Summary of MET3 sequence features

(A) A schematic representation of the MET3 protein sequence shows potential nuclear localisation signals (NLS) dispersed throughout the sequence. The filled red box indicates a bipartite NLS, The open boxes represent monopartite NLS. (B) Vertical black lines indicate SR/RS dipeptides that are found scattered throughout the sequence. (C) The grey box indicates a region which corresponds to the consensus core sequence of the RNA-binding motif RNP-1 (see also Figure 4.8). (D) MET3 is predicted to fold into a mix of α helices (orange boxes) and β sheets (yellow boxes). Numbers on top indicate amino acid positions.

```
TcMET3      SGRYA SFSGRSVVATAPTRGRSYLPRS
Tbp34/37    RRYAI FFENAAAVRKAIDLFNEKEVLG
RNP-1 CSC   xUxVxFxxxxxZxxA
```

Figure 4.8: MET3 contains a putative RNA recognition motif

Sequence alignment of amino acid positions 42-69 of MET3 and the RNP-1 motif of *T. brucei* RNA-binding protein p34/37 (Zhang and Williams, 1997). These sequences conform to the consensus structural core sequence of the RNP-1 motif (RNP-1CSC; Birney et al., 1993). Conserved amino acids are shown in red, and conservative substitutions in orange. X = any residue; U = uncharged residue; Z = U + S, T. The sequence similarity between MET3 and p34/37 does not extend beyond this motif.

MET3 contains a number of potential post-translational modification sites (Figure 4.6). Several putative phosphorylation sites for casein kinase II (CK2), and cAMP- and cGMP-dependent protein kinase are found in the central part of the sequence. Putative PKC phosphorylation sites are dispersed throughout MET3. Putative ASN glycosylation site, and two N-myristoylation sites and one amidation site were also detected in a motif scan of Prosite patterns and profiles (Hulo et al., 2004). In addition, we found at least ten potential arginine methylation sites (not shown on Figure): arginine residues in the contexts RGG, RXR, GRG or RG may be targets of arginine methylation (McBride and Silver, 2001).

MET3 is predicted to localise to the cell nucleus [PSORTII (Nakai and Kanehisa, 1992), 74% probability]. A scan of Prosite profiles, and PSORT II predictions identified a bipartite nuclear localisation signal (NLS) at the N-terminus, and four mono-partite NLS distributed throughout the sequence (Figure 4.7A). The program DBS-Pred (Ahmad et al., 2004) predicts that MET3 binds to DNA (91.7% probability), based on amino acid composition. MET3 has several other features characteristic of nucleic-acid binding proteins: the protein is classified as arginine-rich by Prosite profile scan, and nine RS or SR dipeptides are found dispersed throughout the sequence (Figure 4.7B). Furthermore, a sequence reminiscent of RNP-1 is present in the N-terminal half of MET3 (amino acid positions 44-58) (Figures 4.7C and 4.8). RNP-1 is one of two submotifs of the RNA recognition motif RRM (Burd and Dreyfuss, 1994), a region of around 80 amino acids found in many RNA binding proteins (of prokaryotes and eukaryotes). MET3 corresponds to the conserved structural core sequence of RNP-1 (Figure 4.8) (Birney et al., 1993), but the second submotif, RNP-2, is absent from MET3. MET3 is clearly distinct from the recently discovered family of *T. cruzi* RNA-binding proteins which contain the canonical RRM domain (De Gaudenzi et al., 2003). There is however similarity between MET3 and the RNP-1 domain of the developmentally regulated nuclear RNA binding proteins p34/37 of *T. brucei* (Zhang and Williams, 1997) (Figure 4.8). None of these features alone is a definitive indicator of a particular function. Taken together however, they suggest that MET3 may bind to nucleic acids.

We extended sequence analysis of MET3 to structure, to gain maximal insight into its potential function. Since accurate *ab initio* protein structure prediction has remained beyond the limits of current bioinformatic analysis, and since there are no homologous sequences available on which MET3 structure could be modelled, our analysis was limited to the prediction of likely secondary structures, and the search for conserved protein folds. The latter can identify more distant evolutionary relationships that are not detected by standard sequence alignment, and give important clues to a putative function. MET3 is predicted to fold into a mixture of α helices and β strands (Figure 4.7D), as determined from the consensus output from four different prediction programs (section 2.5). We also used 3D structure prediction programs (section 2.5) to calculate if MET3 could adopt any protein fold represented in the Protein Data Bank PDB, but found no known conserved folds. This suggests that MET3 represents a novel structure.

In summary,

- *MET3* is a *T. cruzi* specific gene with no apparent homologues in other organisms.
- *MET3* is a single copy gene, in a gene-rich region on a 2 Mb chromosome.
- The MET3 protein sequence contains a high proportion of basic residues and several putative nuclear localisation signals.
- Sequence analysis suggest that MET3 is a nuclear protein that may interact with DNA or RNA.

4.2 MET3 IS A STAGE-SPECIFIC NUCLEAR PROTEIN

4.2.1 *MET3* RNA is upregulated during metacyclogenesis

MET3 was initially identified in a screen for RNA transcripts that are enriched in metacyclics of *T. cruzi* strain Dm28c (Yamada-Ogatta et al., 2004). We established by Northern blot analysis that *MET3* expression is also upregulated in CL Brener during metacyclogenesis (Figure 4.9A). RNA was extracted from cultures containing 0%, 6% and 24% metacyclics, respectively. A double band migrating at about 3.5/3.9 kb hybridises with a probe derived from the *MET3* CDS in all samples. The

signal is much stronger in the sample from the culture with 24% metacyclics com-
pared to the ones with fewer metacyclics. The abundance of *MET3* transcripts thus
correlates with development of metacyclics in CL Brener. This pattern of upregu-
lation is similar to that found for *MET3-Dm28c*. As an internal control, the same
RNA was hybridised with a probe for the metacyclic-specific gene *gp82*. A probe
corresponding to the 5' end of the coding sequence of *gp82* (Fragment J18b coding
for amino acids 224-516, (Araya et al., 1994)) hybridised to a single 2.2 kb band in
all RNAs tested (Figure 4.9B). The signal was much stronger in RNA from a popu-
lation containing 24% metacyclics than in a pure epimastigote culture, consistent
with the reported up-regulation of steady-state levels of *gp82* RNA in metacyclics.
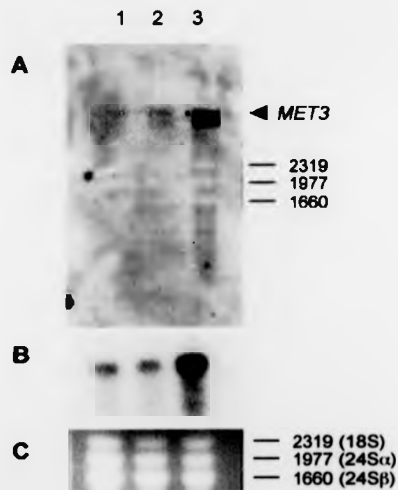
Figure 4.9: Northern blot shows upregulation of MET3 *transcripts in stationary phase cultures*
Total RNA was extracted from cultures containing different proportions of metacyclics and analysed
on Northern blots. Lane 1: log-phase epimastigotes; lane 2: 6% metacyclics; lane 3: 24% metacyclics.
(A) Two transcripts are detected with a *MET3* probe (calculated size: 3.9 and 3.5 kb, respectively).
MET3 transcripts are strongly upregulated in cultures containing 24% metacyclics. (B) *gp82*, a gene
known to be upregulated in metacyclics. (C) RNA loading was controlled by ethidium bromide
staining of the ribosomal RNA bands (18S, 24Sα and 24Sβ) prior to blotting. 10μg RNA was loaded
per lane. Numbers on the right indicate size in bases.

4.2.2 MET3 protein expression is stage-specific

To determine whether MET3 is a marker specifically for metacyclics, or whether MET3 is expressed in all cells of a stationary phase population, we raised antisera against recombinant MET3 protein as described in sections 2.4.2 to 2.4.4. The full-length *MET3* gene was cloned in expression vector pTrcHis C (section 2.3.4). The resulting vector pTrcHis-met3 was transformed into *E. coli* BL21, and expression of the recombinant 6xHis-MET3 protein was induced with IPTG. Western blots of *E. coli* lysates were screened, using the anti-Xpress antibody. 6xHis-MET3 is expressed at low levels in *E. coli*, and is predominantly found in the insoluble fraction. Expression of the protein at 18°C or 4°C instead of 37°C did not yield higher proportions of soluble protein. Addition of 1% Triton X-100 to the lysis buffer, increased the amount of soluble protein and allowed purification of soluble 6xHis-MET3 under native conditions on a Ni-NTA agarose column (Figure 4.10A).

The purified recombinant protein was used to raise antisera in mice. Antisera that detect recombinant MET3 (Figure 4.10B), as well as native MET3 from *T. cruzi* lysates (Figure 4.10C) were obtained from two mice. No protein was detected with the pre-immune sera (Figure 4.10C), confirming that the antisera specifically recognise MET3. Western blot analysis using these antisera shows that a protein of approximately 21 kDa (calculated molecular weight for MET3) is expressed at high levels in cells from stationary phase cultures, but not detectable in lysates from log-phase epimastigotes (Figure 4.11).

We noticed that MET3 is solubilised at high salt concentrations. At NaCl concentrations lower than 0.3 M, MET3 is present in the soluble as well as insoluble fraction of *T. cruzi* protein extracts (Figure 4.11). At higher concentrations (0.5 M to 1.2 M) MET3 was detected exclusively in the soluble fraction (data not shown).
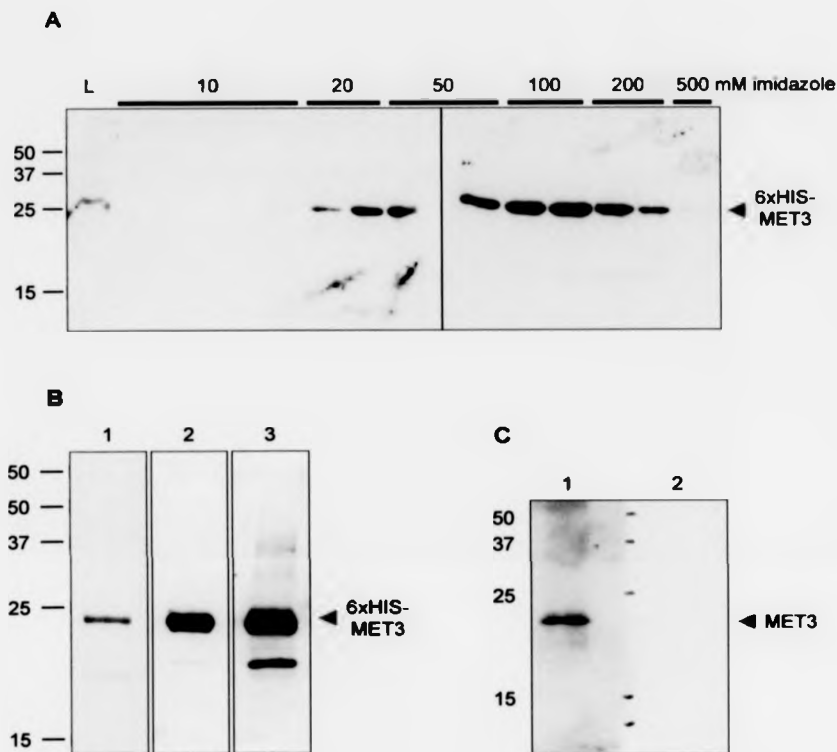
Figure 4.10: Generation of MET3 antiserum

(A) Purification of recombinant MET3 from *E. coli*. Full-length MET3, with an N-terminal 6xHis tag and an Xpress epitope tag (6xHis-MET3) was expressed in *E. coli* BL21, and purified under native conditions using a Ni-NTA agarose column (section 2.4.3). The image shows detection of 6xHis-MET3 with the anti-Xpress antibody in a Western blot of the *E. coli* lysate (L), and protein fractions eluted from the column using increasing concentrations of imidazole (indicated above each lane). (B) Purified 6xHis-MET3 was used to raise antiserum in mice (section 2.4.4). Lane 1: Purified 6xHis-MET3 (24 kDa) in a Coomassie stained SDS-PAGE gel. Lanes 2 and 3: Antisera from two mice detect the purified 6xHis-MET3 on a Western blot. (C) Lane 1: Detection of native MET3 (21 kDa) on a Western blot of *T. cruzi* stationary phase cell lysates, using the antiserum from mouse 1. Lane 2: Pre-immune serum from mouse 1 does not detect any protein in the *T. cruzi* lysate. Numbers on the left indicate molecular sizes in kilo Daltons.

114
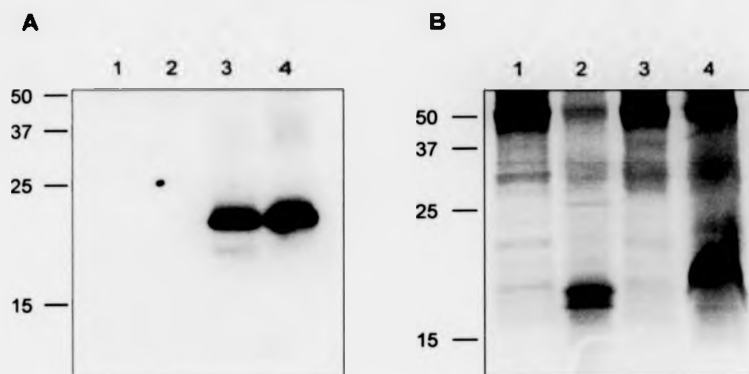
Figure 4.11: Western blot shows stage-specificity of MET3 expression

Protein lysates from *T. cruzi* log phase epimastigotes (lanes 1 and 2) and stationary phase cells (lanes 3 and 4) were analysed on a Western blot, using the MET3 antiserum. Prior to fractionation on SDS-PAGE, the lysates were separated in a soluble fraction (lanes 1 and 3) and an insoluble fraction (lanes 2 and 4) by centrifugation (section 2.4.1). (A) High levels of MET3 are detected in the stationary phase cell lysates, both in the soluble (lane 3) and insoluble fraction (lane 4). In epimastigote lysates no MET3 protein is detected (lanes 1 and 2). (B) Loading control: Coomassie stained protein gel. Numbers on the left indicate molecular sizes in kilo Daltons.

4.2.3 MET3 is a nuclear protein that associates with the nucleolus

To determine at which stage during the *T. cruzi* life-cycle the MET3 protein is expressed, we used the MET3 antiserum to visualise the protein in individual cells by immuno-fluorescence confocal microscopy (Figures 4.12 - 4.17). We found that MET3 expression is specific to a subset of cells in stationary phase culture: MET3 is strongly expressed in fully developed metacyclics where it localises to the nucleus (Figures 4.12 and 4.13). It is also detected in the nucleus of cells that are morphological intermediates between epimastigotes and metacyclics, as defined by the position of the kinetoplast (Figures 4.12 and 4.14). Cells with epimastigote morphology are negative for MET3 staining, both in stationary phase cultures (Figures 4.12 and 4.14), and in log phase cultures (Figure 4.15), consistent with Western blot

results (Figure 4.11). As negative controls, pre-immune mouse sera, and secondary antibodies alone were tested and found not to stain *T. cruzi* (not shown). Metacyclics frequently exhibit a striking pattern of MET3 distribution. Four to six spherical clusters of MET3 staining are spread in regular order along the anterior-posterior axis of the elongated nucleus (Figure 4.13A-C). Some cells show punctate MET3 staining throughout the nucleus which appears less regular (Figure 4.13D-E), while in others just one or two clusters of stronger staining persist with most of the signal dispersed diffusely throughout the nucleus, (Figure 4.13F). Areas of strong MET3 staining tend to stain weakly with the DNA stain TOTO-3. This pattern of staining suggested a possible nucleolar localisation.

Distribution of MET3 in the nucleus of intermediate forms also appears punctate, with one area staining strongly for MET3 and patches of weaker staining arranged along the edge of the nucleus (Figure 4.14). MET3 was also detected in the nucleus of a subset of amastigotes at early stages of macrophage infections where it localises to the nucleolus (Figure 4.16A-B). Amastigotes at later stages of macrophage infections (Figure 4.17A), or trypomastigotes released from infected macrophages six days after infection (Figure 4.17B-C) were negative for MET3 staining. Thus, MET3 expression is a specific marker for cells that go through metacyclogenesis: MET3 expression is first seen in intermediate forms in stationary phase, and is maintained in fully developed metacyclics. MET3 expression ceases at an early time point after development of the amastigote form.
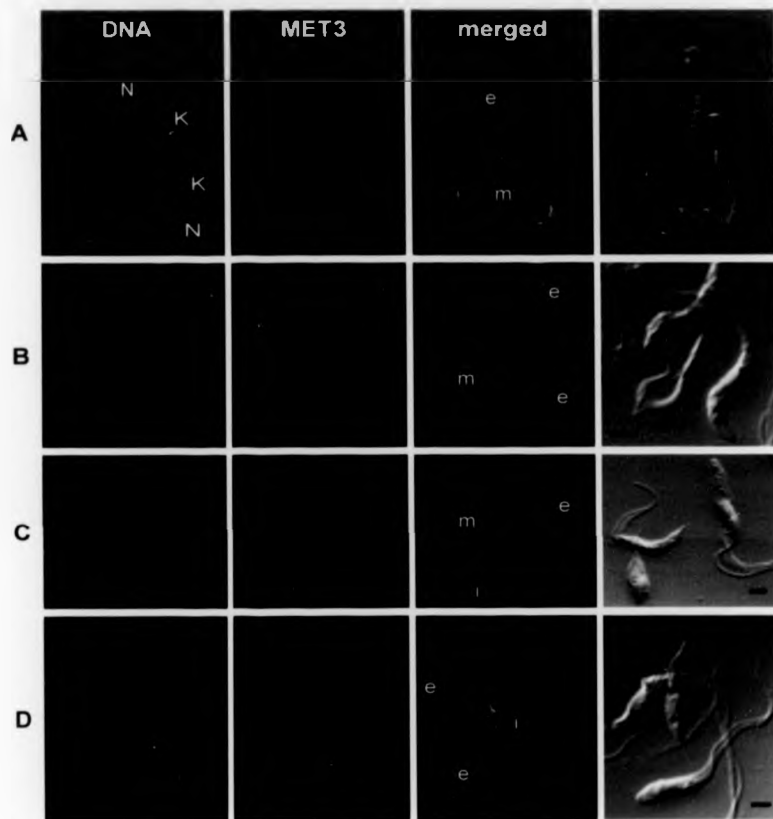
Figure 4.12: MET3 is expressed in metacyclics and intermediate forms

Cells from a stationary phase culture were fixed and immuno-labelled with MET3 antiserum (MET3). Staining is seen in the nucleus of metacyclics (m, panels **A-C**) and intermediate forms (i, panels **C** and **D**), but not in the nucleus of epimastigotes (e, panels **A-D**). Kinetoplast (K) and nuclear DNA (N) is stained with TOTO-3. Scale bar = 2 μm.

Figure 4.13 (overleaf): Localisation of MET3 in the metacyclic nucleus

A detailed examination of the metacyclics in stationary cultures (see Figure 4.12) revealed that MET3 is found dispersed in the elongated nucleus of these cells. Most metacyclics show a punctate pattern, which appears structured into segments in some cells (**A-C**), and dispersed in other cells (**D-E**). A few cells show diffuse MET3 staining throughout most of the nucleus (**F**). Scale bar = 2 μm.

117
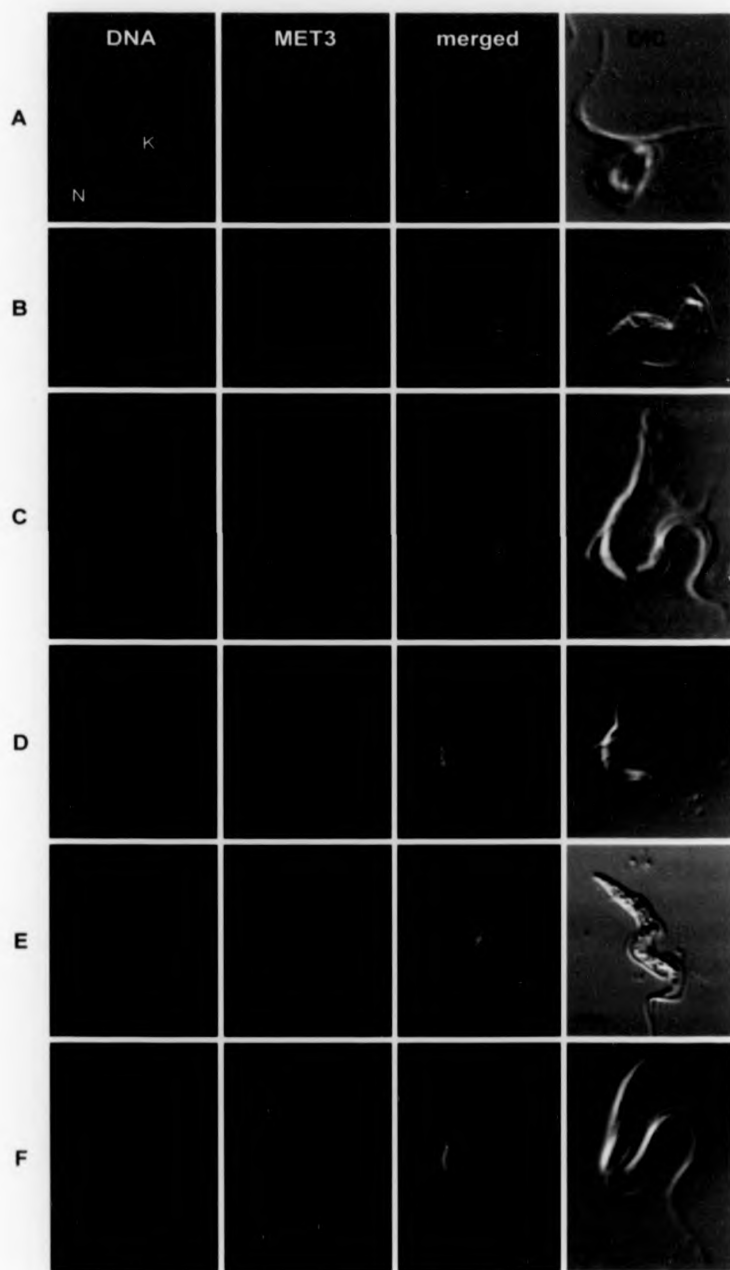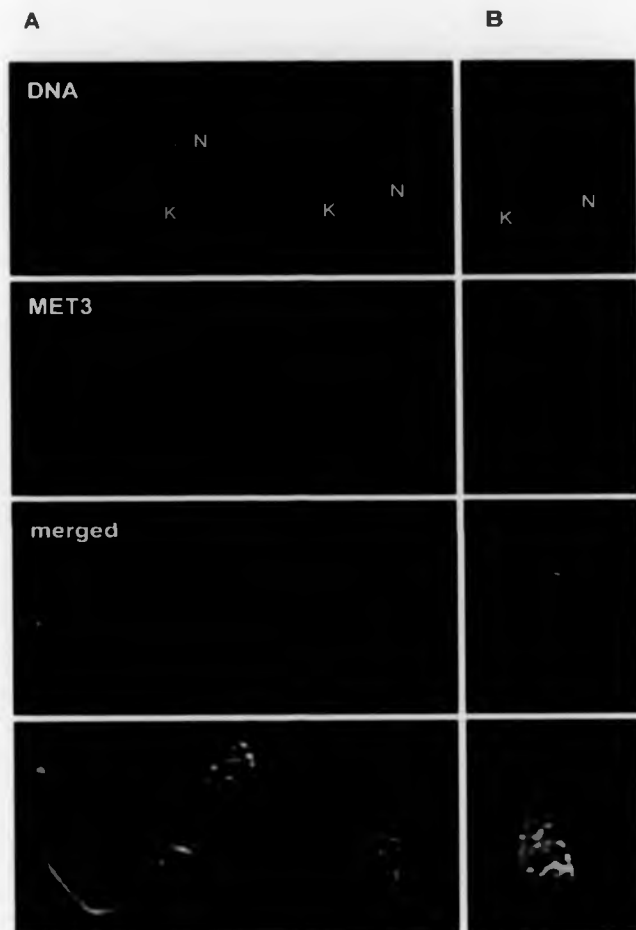
Figure 4.13

Figure 4.14: Subnuclear distribution of MET3 in intermediate forms
Cells from a stationary phase culture were fixed and immuno-labelled with MET3 antiserum (MET3).
Kinetoplast (K) and nuclear DNA (N) is stained with TOTO-3. In epimastigotes the kinetoplast is
positioned at the anterior end of the nucleus. In intermediate forms the kinetoplast has moved to the
side of the nucleus. MET3 is detected in the nucleus of intermediate forms (A and B). MET3 appears
to be concentrated in patches. No MET3 is detected in the nucleus of epimastigotes (panel A, cell on
the left). Scale bar = 2 μm.
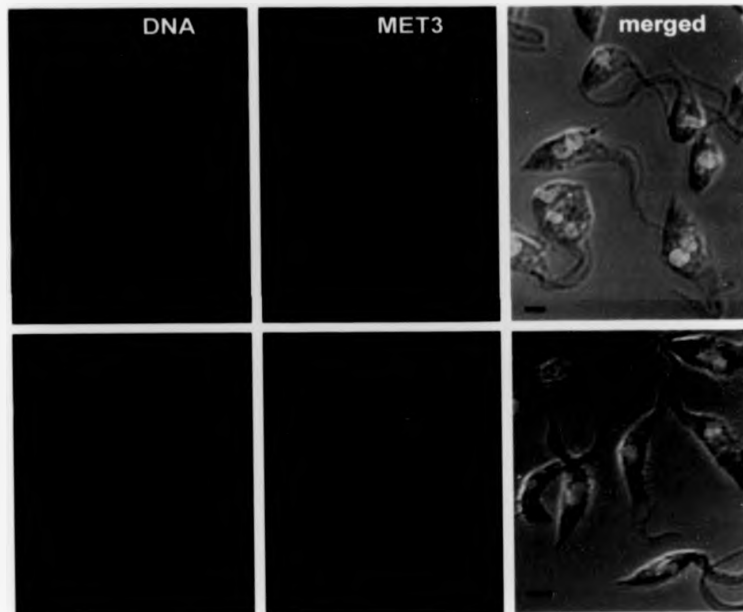
Figure 4.15: MET3 is not expressed in proliferating epimastigotes

Epimastigotes in log phase of growth were fixed and immuno-labelled with MET3 antiserum (MET3). No MET3 is detected in these cells. Kinetoplast and nuclear DNA is stained with TOTO-3. Scale bar = 2 μm.
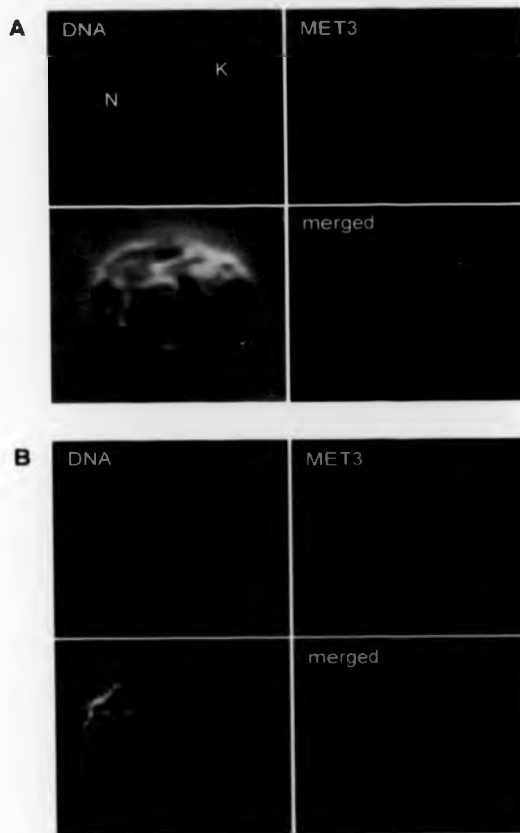
Figure 4.16: MET3 in the nucleolus of early amastigotes

T. cruzi-infected macrophages were fixed 24 h post infection. Cells were immuno-labelled with MET3 antiserum, and kinetoplast (K) and nuclear DNA (N) is stained with TOTO-3. **(A)** An enlarged image of one amastigote shows that MET3 staining is concentrated in the nucleolus, which stains weakly with TOTO-3. **(B)** Amastigotes inside a macrophage. MET3 is detected in the nucleus of three amastigotes. Two amastigotes are negative for MET3 staining. Scale bar = 2 μm.
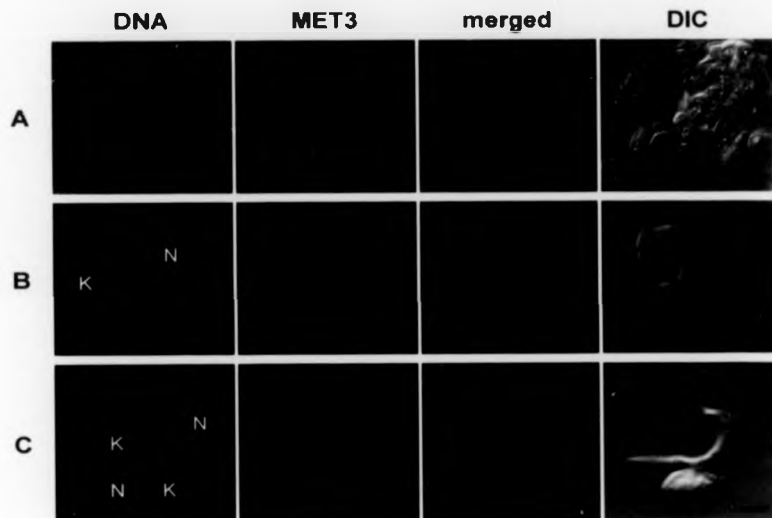
Figure 4.17: Replicating amastigotes and trypomastigotes express no MET3

T. cruzi-infected macrophages and cells from the supernatant of infected cultures were fixed and immuno-labelled with MET3 antiserum. Kinetoplast (K) and nuclear DNA (N) were stained with TOTO-3. (A) Six days post infection amastigotes have replicated inside macrophages. The antiserum fails to detect any MET3 in these amastigotes. (B) A trypomastigote that has emerged from the infected macrophages on day six is also negative for MET3 staining. (C) Two weeks post infection, trypomastigotes and amastigotes are abundant in the supernatant of infected macrophage cultures. The MET3 antiserum fails to detect any MET3 in the nucleus of these cells. Scale bar = 2 μm.

4.2.4 The nucleolus

Earlier studies have documented extensive nuclear reorganisation that takes place during *T. cruzi* differentiation. Heterochromatin becomes widespread and nucleolar antigens disperse in the nucleus of trypomastigote forms (Elias et al., 2001b). The nucleolus is a morphologically defined, yet highly dynamic structure. Its primary function is ribosome biogenesis, which involves transcription of rDNA, processing of pre-rRNA and assembly of rRNA with ribosomal proteins into ribonucleoprotein complexes (Melese and Xue, 1995). How the functional architecture of the nucleolus is regulated is only partially understood and remains a focus of ongoing investigation (Scheer and Hock, 1999). Interestingly, in recent years it has emerged that in addition to its role in ribosome biogenesis, the nucleolus also plays regulatory roles in the control of cell proliferation, through sequestration of regulatory molecules (Carmo-Fonseca et al., 2000; Scheer and Hock, 1999).

In trypanosome nuclei stained with TOTO-3, the nucleolus is visible s a single acentric spherical region from which the DNA dye is excluded. *T. brucei* bloodstream forms are shown in Figure 4.18, and replicating *T. cruzi* cells in Figures 4.18 and 4.19. In the nucleus of non-replicating *T. cruzi* cells, a nucleolus is not readily discernible. In these cells, DNA-staining is non-homogenous with a number of patches that stain weakly for DNA (Figure 4.20). Since the distribution of MET3 resembled closely the reported distribution of nucleolar antigens in metacyclics (Elias et al., 2001b), and since MET3 clearly localises to the nucleolus in early amastigotes (Figure 4.16), we examined in more detail the dynamics of nucleolar antigens. To compare the localisation of MET3 and nucleolar antigens directly we used antibodies against nucleolar antigens, and expressed tagged versions of MET3.

4.2.5 Distribution of nucleolar antigen L1C6 throughout the life-cycle of *T. cruzi*

To follow the distribution of nucleolar antigens in the *T. cruzi* nucleus, we tested two antibodies: The rabbit NOG1 antiserum, raised against a well-characterised highly conserved *T. brucei* nucleolar G-protein with a role in ribosome biogenesis (Jensen et al., 2003; Park et al., 2001) and the L1C6 antibody, raised against an uncharacter-

123

ised nucleolar *T. brucei* antigen (K. Gull, unpublished). To determine with certainty that L1C6 detects a nucleolar antigen we stained *T. brucei* cells with NOG1 and L1C6. The results clearly demonstrate co-localisation of NOG1 and L1C6 in the area of the nucleus that stains weakly with TOTO-3 (Figure 4.18A-B). When *T. cruzi* epimastigotes were stained with these antibodies, identical results were obtained with L1C6 (Figure 4.18C), but the NOG1 antiserum gives extensive background staining in most *T. cruzi* cells (Figure 4.18C).

In replicating cells, i.e. epimastigotes taken from log phase of growth, and in amastigotes, the nucleolus is visible as a distinct dot in a round nucleus as determined by weak TOTO-3- and strong L1C6 staining (Figure 4.19). In dividing cells, each of the two daughter nuclei has a well-defined nucleolus. Occasionally the L1C6 signal appears slightly dispersed, but remains restricted to a small area of the nucleus (Figure 4.19B). This pattern is maintained in epimastigotes until the end of log phase growth.

In non-replicating cells taken from stationary phase cultures, the nucleolar antigen disperses (Figure 4.20). L1C6 staining is visible as a punctate pattern distributed throughout the nucleus of these non-dividing epimastigotes. In some cells a small number of discrete dots are observed (Figure 4.20A-B), which coincide with areas of weak DNA staining, whereas others show large patches of diffuse staining throughout the nucleus (Figure 4.20C). Cells that are intermediate forms between epimastigotes and metacyclics show a dispersed L1C6 staining pattern (Figure 4.21A), similar to that in stationary phase epimastigotes (Figure 4.20). In fully developed metacyclics, L1C6 detects a small number of discreet dots along the anterior-posterior axis of the elongated nucleus (Figure 4.21B-D), which again coincide, with areas of weak DNA staining.

The distribution of the L1C6 nucleolar antigen in intermediate forms, metacyclics and early amastigotes follows essentially the same pattern as the distribution of MET3. Of particular note is the segmented pattern in some metacyclics that is seen with both antibodies (compare Figures 4.13 and 4.21).

124
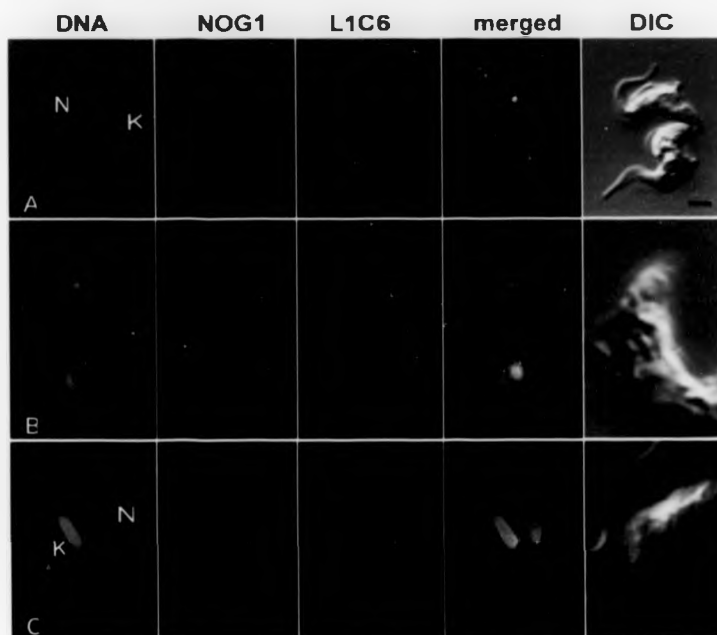
Figure 4.18: Markers for the trypanosome nucleolus

(A) and (B) *T. brucei* bloodstream forms were fixed and double labelled with rabbit antiserum against nucleolar G-protein NOG1 and mouse monoclonal antibody L1C6. Kinetoplast (K) and nuclear DNA (N) is stained with TOTO-3. The nucleolus stains weakly with TOTO-3 (panel **B**, arrow). The merged images show that the L1C6 antigen co-localises with NOG1. (**C**) *T. cruzi* epimastigotes were fixed and double labelled with the anti-NOG1 and L1C6 antibodies. The L1C6 antigen is localised to the nucleolus, which stains weakly with TOTO-3 (arrow). Antiserum against NOG1 (upper panel) does not specifically stain the *T. cruzi* nucleolus in this cell. Scale bar = 2 μm.

Technical difficulties have prevented us from proving co-localisation directly by double staining. Because both the MET3 antiserum and the L1C6 antibody were raised in mice, they would have to be directly labelled with a fluorescent substance in order to distinguish the signals. The rabbit NOG1 serum which could be used in double labelling experiments gives extensive background staining in most *T. cruzi* cells and was therefore considered unsuitable to analyse the dispersed pattern in metacyclics.
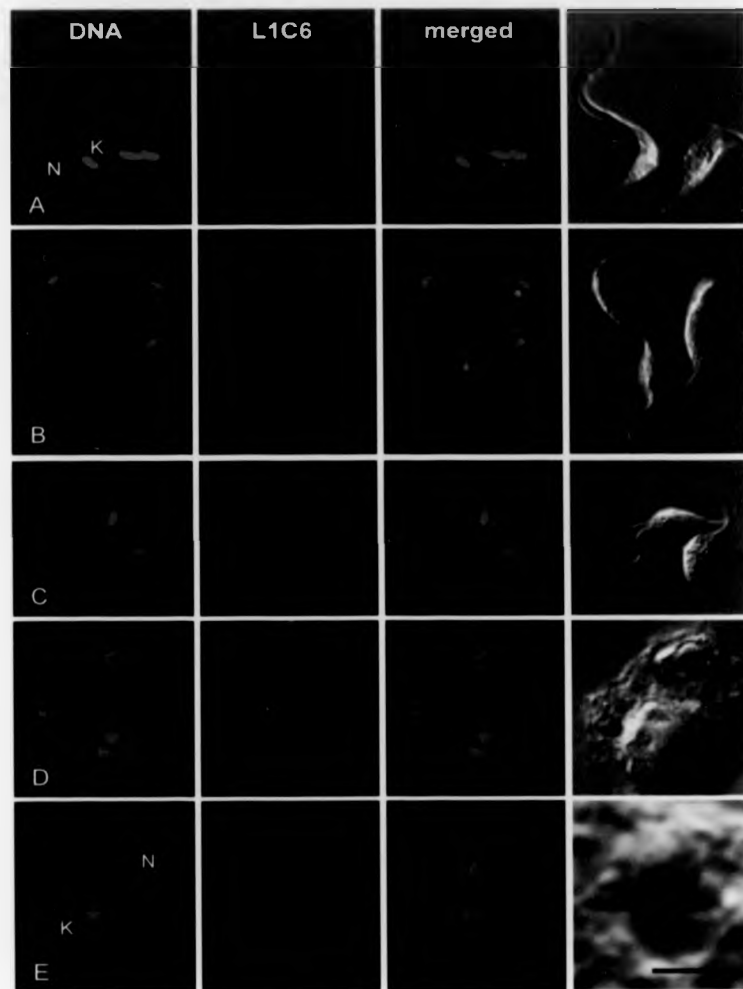
Figure 4.19: The nucleolar antigen L1C6 in replicating T. cruzi *cells*

Cells were fixed and labelled with monoclonal antibody L1C6. Kinetoplast (K) and nuclear DNA (N) is stained with TOTO-3. The nucleolus is visible as area of weak TOTO-3 staining. A single dot of L1C6 staining is visible in the nucleolus in log phase epimastigotes (A-C), and in amastigotes (D-E). Scale bar = 2 µm.

	DNA	L1C6	merged	DIC
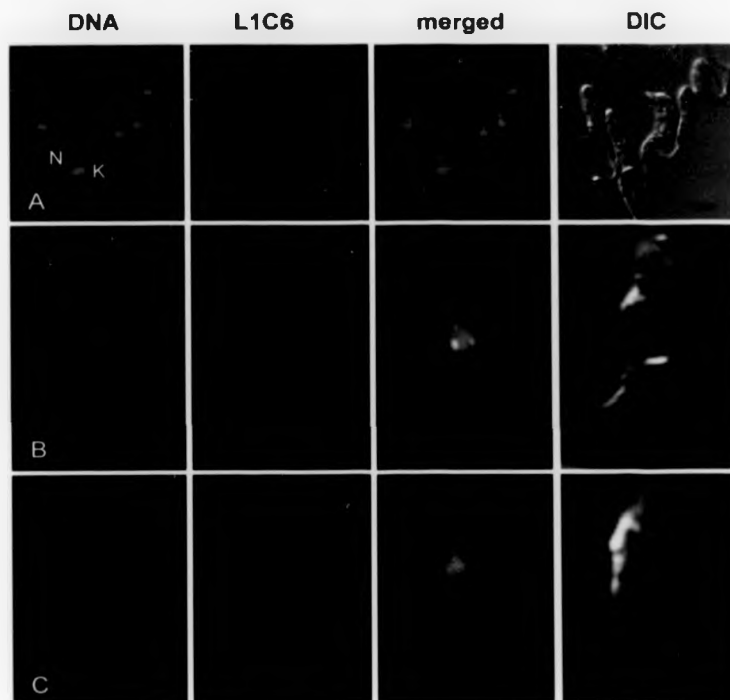
Figure 4.20: Dispersal of the nucleolar antigen L1C6 in non-replicating T. cruzi *epimastigotes*
T. cruzi stationary phase cells were fixed and labelled with monoclonal antibody L1C6. Kinetoplast
(K) and nuclear DNA (N) is stained with TOTO-3. In stationary phase epimastigotes, L1C6 staining
appears as a dispersed punctate pattern. The L1C6 nucleolar antigen clusters in distinct foci on a
background of diffuse staining. Scale bar = 2 μm.
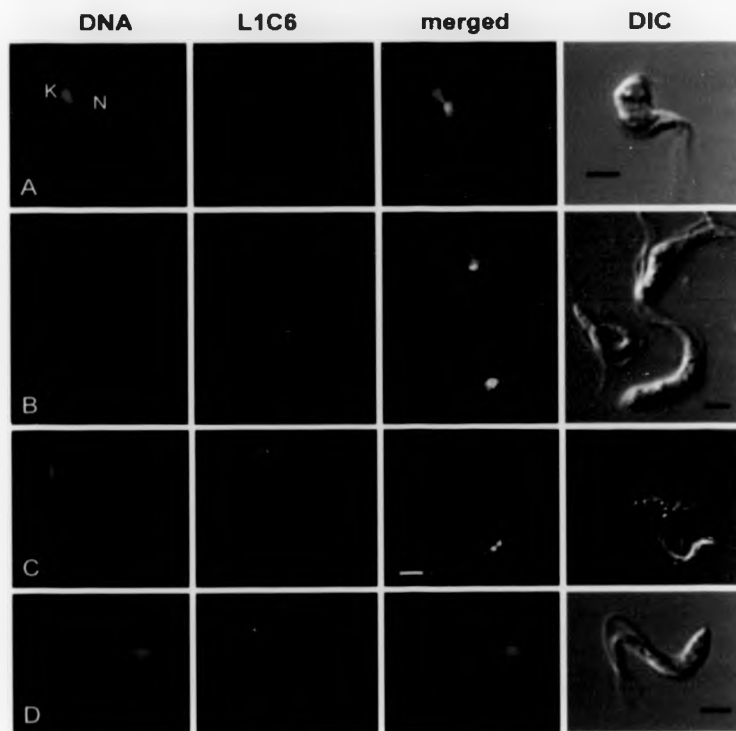
Figure 4.21: Distinct clusters of nucleolar antigen L1C6 in the nucleus of metacyclics

T. cruzi stationary phase cells were fixed and labelled with monoclonal antibody L1C6. Kinetoplast (K) and nuclear DNA (N) is stained with TOTO-3. In intermediate forms (**A**), and in stationary epimastigotes (**B**, **C**) a dispersed L1C6 staining pattern is observed. In metacyclics (**B-D**) the L1C6 antigen clusters in distinct foci along the axis of the elongated nucleus. Scale bar = 2 μm.

4.2.6 Episome-encoded MET3 localises to the nucleolus
of replicating epimastigotes

In wild-type *T. cruzi*, MET3 is not expressed in replicating epimastigotes. Episomal expression vectors were therefore constructed (section 2.3.4) to constitutively express tagged versions of MET3 in epimastigotes. These were used to test whether the recombinant protein localises to the nucleolus and to determine if inappropriate expression of MET3 enhances metacyclic development. Vector pTEX-MET3-9E10 was designed for expression of MET3 with a C-terminal c-myc tag (MET3-9E10). Vector pTEX-MET3-GFP was designed for expression of MET3 fused at the C-terminus to green fluorescent protein (GFP). Transfectants for both constructs were readily obtained, and used for localisation of the recombinant proteins by immuno-staining, or direct visualisation of GFP fluorescence.

In epimastigotes, MET3-GFP localises to the nucleolus. Double-labelling shows co-localisation of MET3-GFP with the L1C6 nucleolar antigen (Figure 4.22). MET3-9E10 also localises to the nucleolus in most cells (Figure 4.23A-B). In a few cells, MET3-9E10 is found along the nuclear periphery with only weak staining of the nucleolus (Figure 4.23C). Identical staining patterns were observed with the antibody against the c-myc tag (not shown).

Importantly, we did not observe dispersal of the nucleolar antigens as a consequence of constitutive MET3-GFP expression. Epimastigotes that express tagged versions of MET3 appear morphologically no different from wild-type epimastigotes, which do not express any form of MET3.
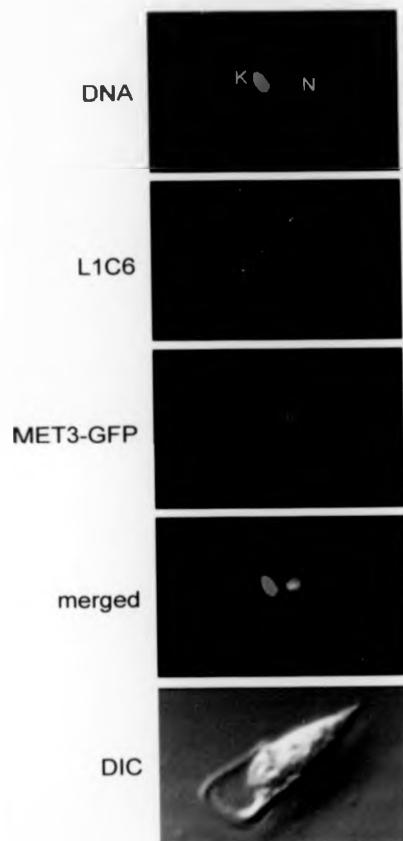
Figure 4.22: Co-localisation of MET3-GFP and L1C6 in the epimastigote nucleolus

T. cruzi epimastigotes transfected with pTEX-MET3-GFP were fixed and stained with monoclonal antibody L1C6. MET3-GFP was localised by directly visualising GFP fluorescence. Kinetoplast (K) and nuclear DNA (N) were stained with TOTO-3. The merged image shows that MET3-GFP and L1C6 co-localise in the nucleolus.
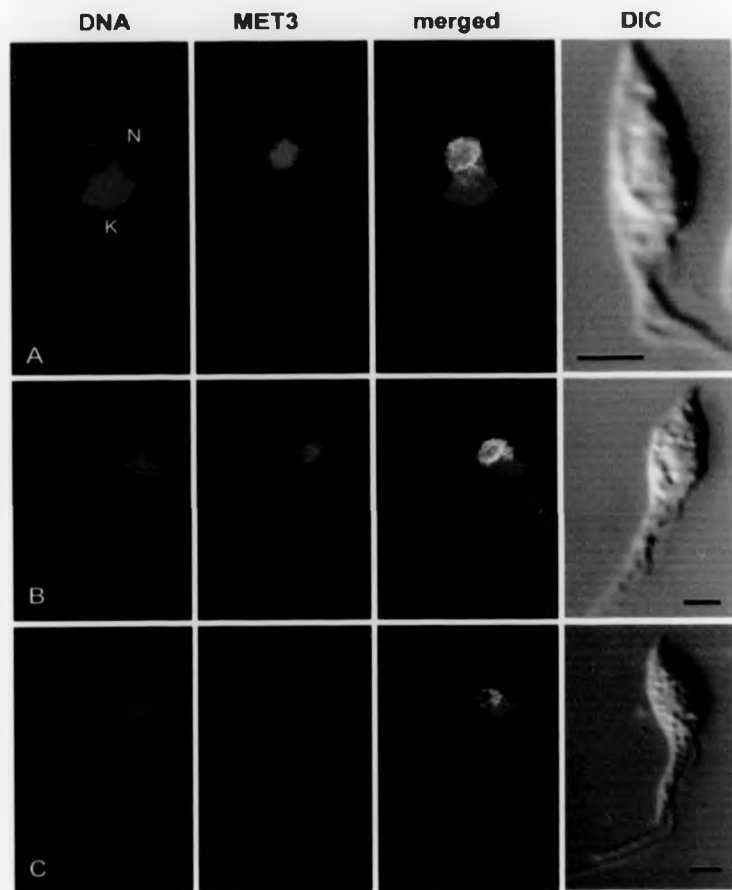
Figure 4.23: Subnuclear localisation of MET3-9E10 in log phase epimastigotes
T. cruzi epimastigotes transfected with pTEX-MET3-9E10 were fixed and labelled with MET3 antiserum. Kinetoplast (K) and nuclear DNA (N) is stained with TOTO-3. (**A-B**) In most cells, strong MET3 staining is observed in the nucleolus, where DNA staining is weak. (**C**) In some cells MET3 staining is observed at the nuclear periphery, with no signal in the nucleolus. Scale bar ▬ 2 μm.

Western blot analysis shows that transfected epimastigote populations express high levels of MET3-9E10 (Figure 4.24). Additional bands recognised by the antiserum, above the 21 kDa MET3, are possibly modified forms of the MET3 protein. The proportion of cells that express MET3-9E10 is very high, and high expression levels are maintained after continuous passage (>1 year) with no apparent deleterious effects. We observed a rapid decrease in the proportion of cells that express MET3-GFP over time. This is consistent with a toxic effect of GFP expression in *T. cruzi* that was noted in other contexts (unpublished observations).
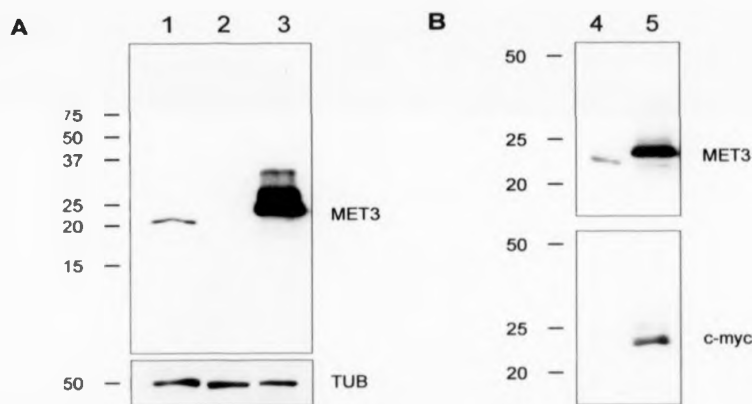
Figure 4.24: Detection of MET3-9E10 in Western blots
(A) Western blot of whole cell extracts from non-transfected *T. cruzi* stationary phase cells (lane 1), non-transfected epimastigotes (lane 2) and pTEX-MET3-9E10 transfected epimastigotes (lane 3). Antiserum against MET3 detects native MET3 in non-transfected stationary phase cells (lane 1), and recombinant MET3-9E10 in transfected epimastigotes (lane 3). Non-transfected epimastigotes express no MET3 (lane 2). The same blot was probed with anti-β-tubulin antibody (TUB) as a loading control. (B) Western blot of the soluble protein fraction from non-transfected *T. cruzi* stationary phase cells (lane 4) and from pTEX-MET3-9E10 transfected epimastigotes (lane 5). MET3 antiserum detects the native MET3 (lane 4) in non-transfected cells, and recombinant MET3-9E10 in transfected cells (lane 5). An antibody against the c-myc tag (9E10) detects the recombinant MET3-9E10 (lane 5). Numbers on the left indicate molecular weight in kilo Daltons.

In non-replicating cells, the tagged MET3 protein becomes dispersed. MET3-9E10 is found in a dispersed punctate pattern in stationary phase epimastigotes (Figure 4.25), differentiating cells (Figure 4.26A) and metacyclics (Figure 4.26B-D). Regions of strong MET3-9E10 signal tend to stain weakly for DNA. This pattern resembles closely the dispersal of the L1C6 nucleolar antigen we found in wild-type cells.
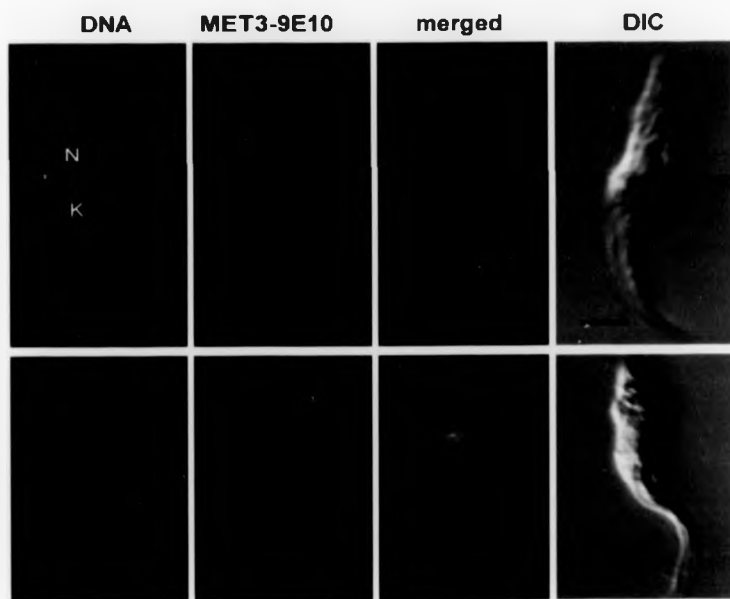
Figure 4.25: Dispersal of MET3-9E10 in the nucleus of stationary phase epimastigotes
Stationary phase epimastigotes expressing MET3-9E10 were fixed and labelled with MET3 anti-serum. Kinetoplast (K) and nuclear DNA (N) is stained with TOTO-3. MET3 staining is visible as dispersed pattern. Strong MET3 staining coincides with areas where DNA staining is weak. Scale bar = 2 µm.
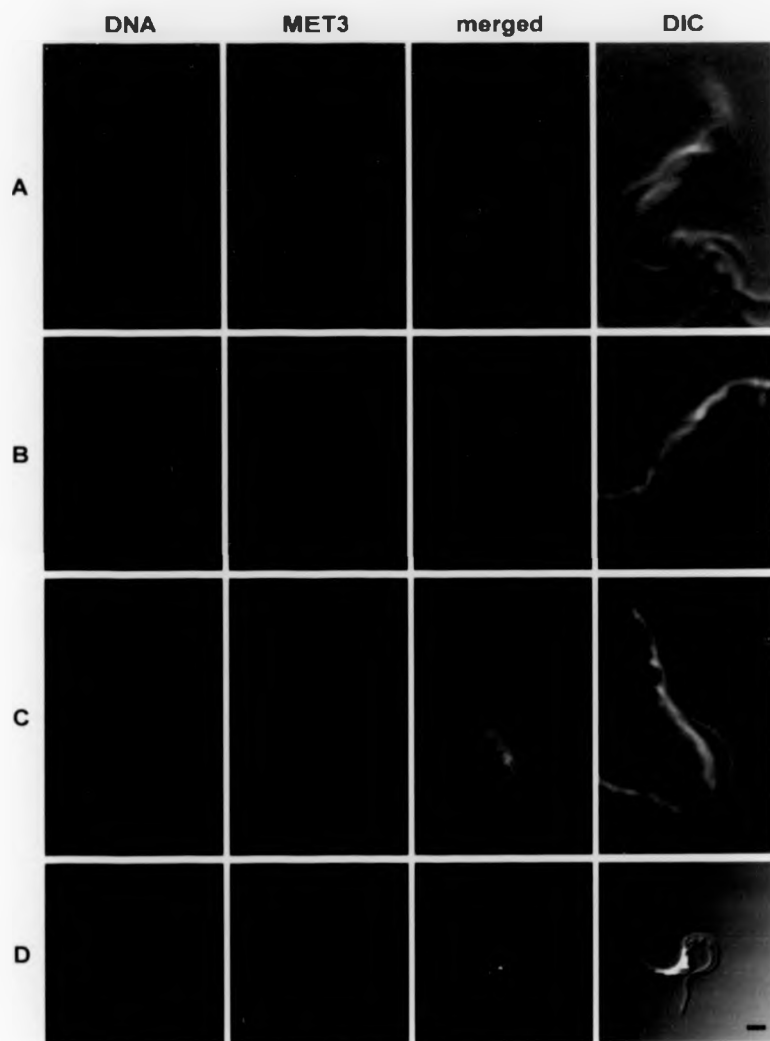
Figure 4.26: Clusters of MET3-9E10 in the metacyclic nucleus
Stationary cells expressing MET3-9E10 were fixed and labelled with MET3 antiserum. Kinetoplast and nuclear DNA is stained with TOTO-3. In metacyclics (A-C) and intermediate forms (D), MET3-9E10 clusters in distinct foci, particularly in areas where DNA staining is weak. Scale bar = 2 μm.

In summary,

- *MET3* RNA levels are upregulated during metacyclogenesis
- MET3-specific antiserum was generated using recombinant MET3, expressed in *E. coli.*
- MET3 protein expression is stage-specific. The protein was only detected in intermediate forms, metacyclics and early amastigotes.
- MET3 is a nuclear protein.
- The subnuclear localisation of native MET3, and tagged versions of MET3 suggests association with nucleolar components.
- Dispersal of nucleolar antigens was observed in all non-replicating *T. cruzi* cells.
- Overexpression of MET3 in epimastigotes had no apparent effect on the cells.

4.3 IDENTIFICATION OF TWO SEQUENCE ELEMENTS THAT CAN MEDIATE NUCLEOLAR LOCALISATION

A commonly used method to locate targeting sequences in a protein is to fuse fragments of the protein to green fluorescent protein (GFP) (Chalfie et al., 1994), or to an epitope tag (such as c-myc), and determine the subcellular localisation by fluorescence microscopy. Both approaches have been used successfully in trypanosomes. With regard to nuclear targeting, GFP-fusion experiments demonstrated that a monopartite NLS in the La protein homologue, and a bipartite NLS in histone H2B mediate nuclear import in *T. brucei* (Marchetti et al., 2000). Nuclear import in trypanosomes has been found to rely on signals and machinery similar to that in higher eukaryotes (Marchetti et al., 2000).

Whereas NLS can be predicted with some confidence (Cokol et al., 2000), nucleolar targeting signals are much more diverse. Very few sequence motifs have been characterised in detail that can mediate nucleolar targeting in trypanosomes (Das et al., 1998; Hoek et al., 2000; Marchetti et al., 2000).

Bioinformatic analysis of the MET3 amino acid sequence detected one bipartite and four monopartite NLS in MET3 (Figure 4.7A). The experiments presented here

aimed to determine which elements in the MET3 sequence are critical for the specific subnuclear localisation of MET3. Ten expression constructs were designed for this purpose (overview in Figure 4.27). Each construct contains a fragment of the MET3 coding sequence, fused in-frame with the full length sequence of GFP, or with the c-myc (9E10) epitope (section 2.3.4). Every construct was expressed in epimastigotes. Localisation of the fusion proteins was analysed in epimastigotes and in metacyclics by direct observation of GFP fluorescence, or where applicable, by observation of immuno-stained cells using the anti-c-myc antibody.

Figure 4.27 (overleaf): Summary of MET3 fusion constructs

(A) Schematic diagram of expression vector pTEXeGFP (section 2.3.4). Fragments of the *MET3* gene were cloned into the multiple cloning site (MCS) of this vector to generate in-frame fusions with the green fluorescent protein (*GFP*) coding sequence. Fusion proteins were expressed in *T. cruzi* epimastigotes and their localisation determined by microscopic examination. (B) Schematic representation of the MET3 protein. (C) Fragments of MET3 that localise to the *T. cruzi* nucleolus when fused to c-myc (construct 7/8) or GFP (all other constructs). (D) Fragments of MET3 that do not specifically localise to the nucleolus when fused to GFP. These fusion-proteins are found throughout the cytoplasm and nucleus. Numbers on top identify two sequence elements that are critical for targeting of the proteins to the *T. cruzi* nucleolus. Element 1 consists of a bipartite NLS (filled red box) and a sequence with similarity to the RNA-binding motif RNP-1 (grey box). Element 2 consists of two NLS at the C-terminus (filled red boxes). Elements 1 and 2 can mediate nucleolar localisation independently. The open boxes represent predicted NLS sequences that were found not to target GFP to a specific location.

136
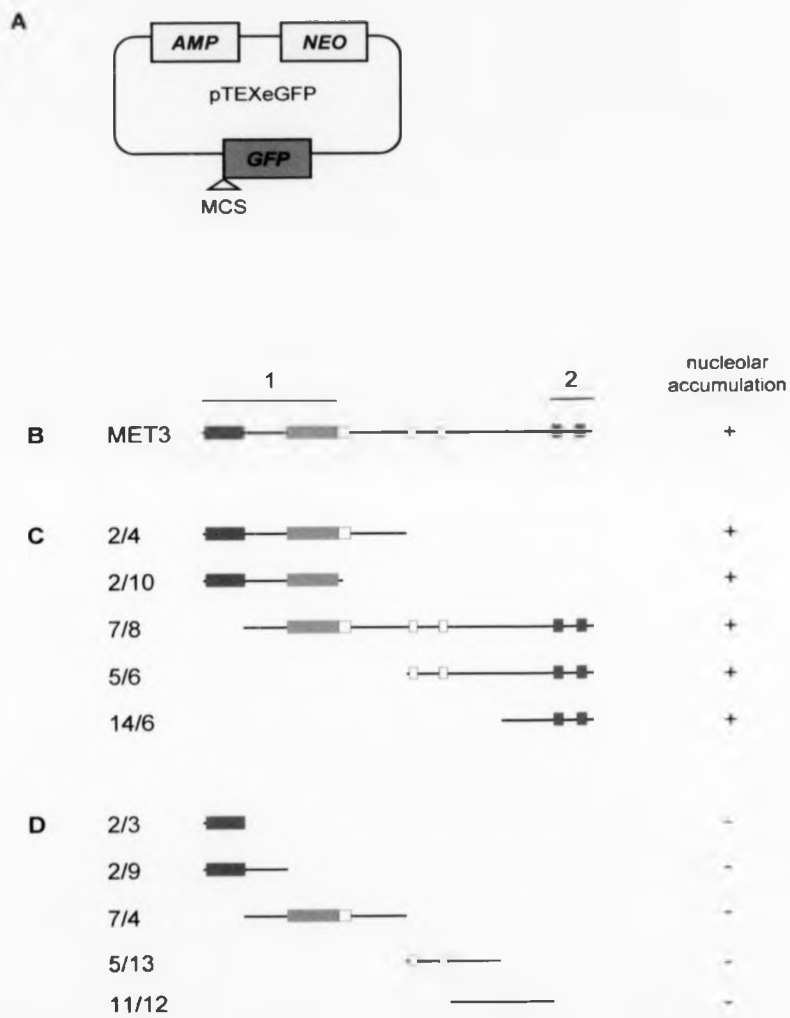
Figure 4.27

The fusion proteins fall into two classes. The first class localises exclusively to the nucleolus in epimastigotes (Figure 4.28 A-D, I-J, M-O), and disperses in the nucleus of metacyclics (Figure 4.28 E-H, K-L, P). These fusion proteins show the same localisation as full-length MET3 (Figure 4.22). The second class is dispersed throughout the whole of the cell cytoplasm, including the flagellum, and the nucleus, but excluding the kinetoplast (Figure 4.29). These fusion proteins show the same localisation as GFP on its own (Figure 4.31). We noted that expression levels of GFP and GFP-fusion proteins differ markedly between individual cells in a population (Figures 4.29F and L and data not shown).

Detailed analysis shows that at least two independent elements within the MET3 sequences can mediate nucleolar targeting. Non-overlapping fragments of the N-terminal half (2/4; Figure 4.28 A-B, E-F) and the C-terminal half (5/6; Figure 4.28 C-D, G-H) of the MET3 fused to GFP both localise to the nucleolus. Further truncation of these fragments show that elements responsible for nucleolar localisation are near the N- and C-terminus, respectively.

An N-terminal fragment containing the predicted bipartite NLS and the adjacent putative RNA binding domain is sufficient to restrict localisation of the fusion protein to the nucleolus (2/10; Figure 4.28 I-L). This exclusive nucleolar localisation is completely abolished when the two domains are separated: constructs that contain only the predicted bipartite NLS (2/3, 2/9; Figure 4.29 A-I) or only the putative RNA binding domain (7/4; Figure 4.29 J-O) are all dispersed throughout the cytoplasm and nucleus. Those MET3 fusion proteins that did not specifically accumulate in the nucleolus were not excluded from the nucleus, presumably because they are small enough to enter the nucleus freely by diffusion.

A second element at the extreme C-terminus can target the fusion protein to the nucleolus independently of the N-terminal sequence. Fusion proteins 5/6 (Figure 4.28 C-D, G-H), 14/6 (Figure 4.28 M-P) and 7/8 (Figure 4.30A) all show exclusive nucleolar localisation. Fusion proteins 5/13 (Figure 4.29 P-S) and 11/12 (Figure 4.29 T-Y), which lack the extreme C-terminus as well as the entire N-terminal half of MET3, are completely dispersed. This indicates that the last 19 amino acids (the only

sequence element present in 14/6 but absent from 11/12) are important in determining localisation.

The smallest fragment that conferred nucleolar localisation is 46 amino acids long. This fragment contains two closely spaced NLS. At the N-terminus, a sequence that conforms to the consensus for a bipartite NLS is critical for nucleolar accumulation, but it only confers specificity when it is present jointly with the sequence that resembles a RNA binding motif. We found no evidence of a role in protein localisation for the clusters of positively charged amino acids in the central part of MET3. Fusion proteins that contain these elements, but neither of the N- or C-terminal sequences described above (i.e. constructs 7/4 and 5/13), are dispersed throughout the cell. Fusing the c-myc tag instead of GFP to MET3 (7/8; Figure 4.30A) demonstrates that nucleolar localisation is not an artefact of fusing GFP to fragments of MET3.
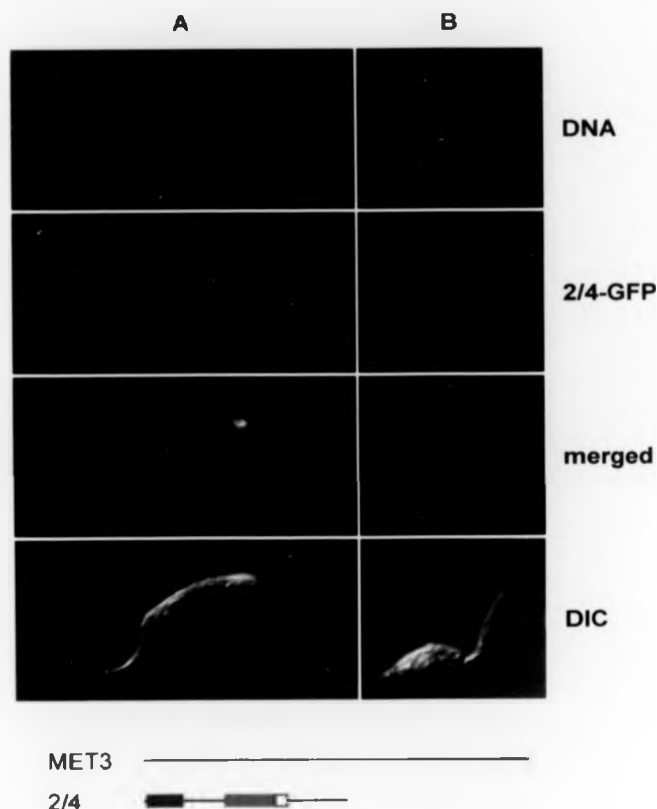
Figure 4.28 MET3-GFP fusion constructs that localise to the nucleolus

(A-P) Images of cells that show nucleolar localisation of the fusion protein in epimastigotes. In meta-cyclics (E-G, H-K, L, P), these fusion proteins disperse in the nucleus, similar to wild-type MET3. The diagram below each set of images shows schematically the full length MET3 sequence (MET3) and underneath it the fragment of MET3 that was fused to GFP. Critical for targeting of GFP to the nucleolus are the nuclear localisation signals (NLS) at the N- and C- terminus (filled red boxes) and a sequence with similarity to the RNA-binding motif RNP-1 (grey box). Predicted NLS in the central part of MET3 (open boxes) were found not to target GFP to a specific location.

Figure 4.28C-D

141
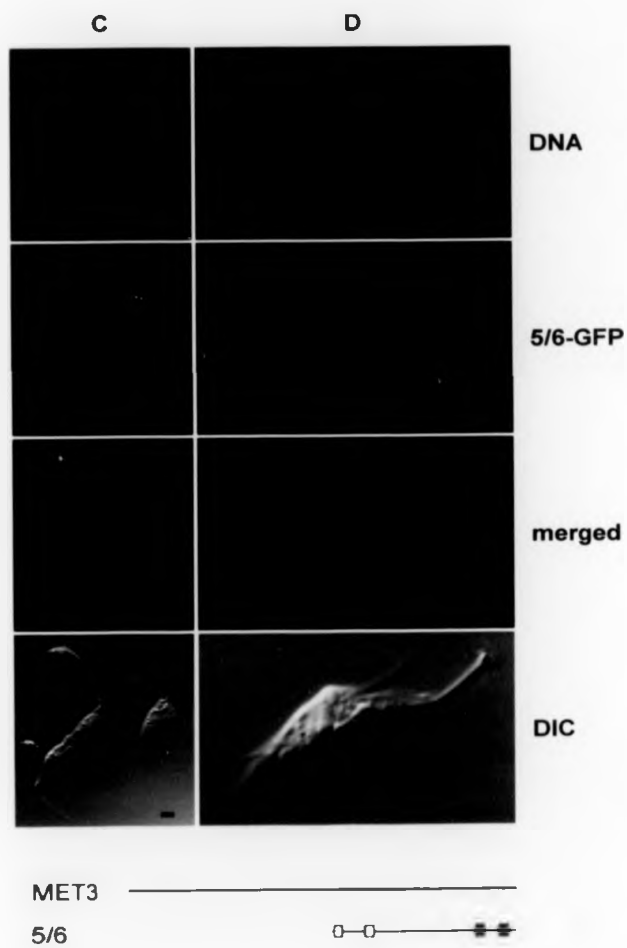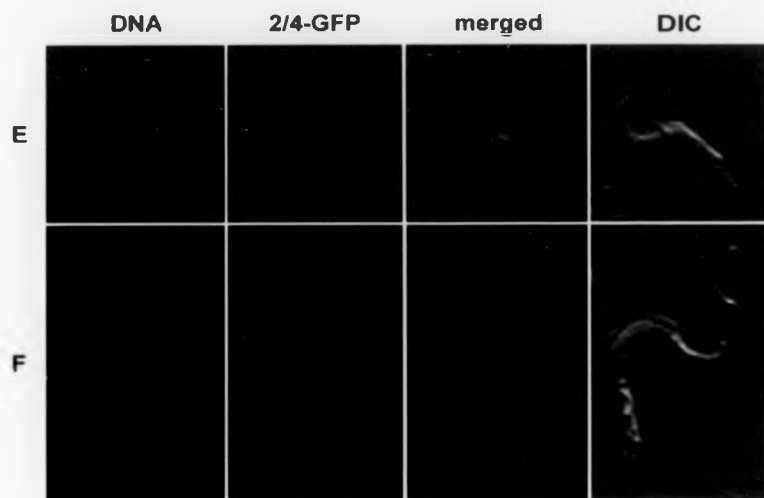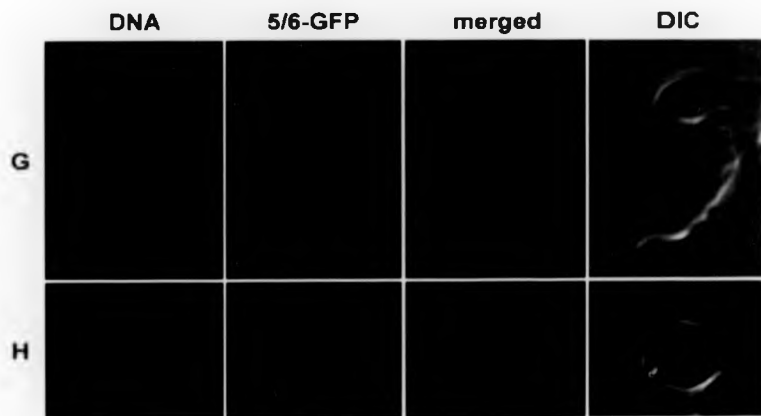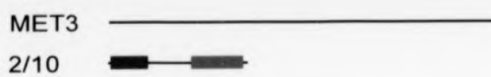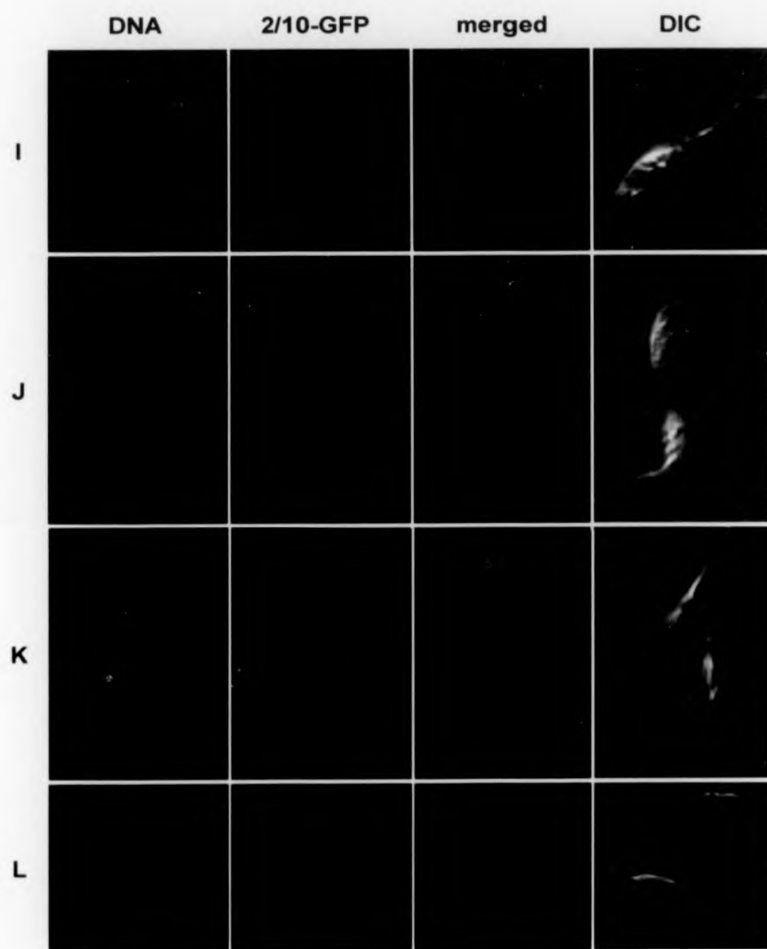
Figure 4.28E-H

Figure 4.28I-L

Figure 4.28M-P
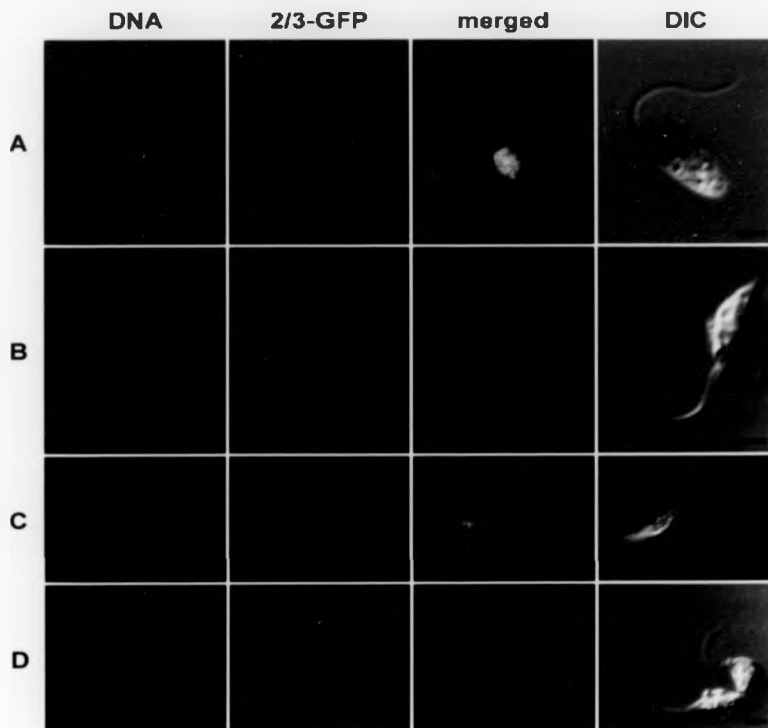
Figure 4.29: MET3 fusion constructs that do not specifically localise to the nucleolus

(A-O) Images of cells that show no specific localisation of the fusion proteins. In epimastigotes and metacyclic trypomastigotes, these non-localised proteins are found throughout the cell, including the nucleus and along the flagellum. They are excluded from the kinetolpast. A diagram below each set of images shows schematically the full length the MET3 sequence (MET3) and underneath it the fragment of MET3 that was fused to GFP. The N-terminal NLS (filled red box) on its own (A-I)or the putative RNA-binding motif (grey box) on its own (J-O) are both insufficient for targeting GFP to the nucleolus. The non-localised fusion proteins also lack the C-terminal NLS (red box). Open boxes indicate predicted NLS sequences that were found not to target GFP to a specific location.

Figure 4.29E-G

Figure 4.29H-I

Figure 4.29J-L

DNA **7/4-GFP** **merged** **DIC**

M

N

O

MET3 ——————————————————

7/4 ——■□—

Figure 4.29M-O

| DNA | c-myc | merged | DIC |

MET3 ——————————————————

7/8 ————▭—————▭—▭—————■■— c-myc
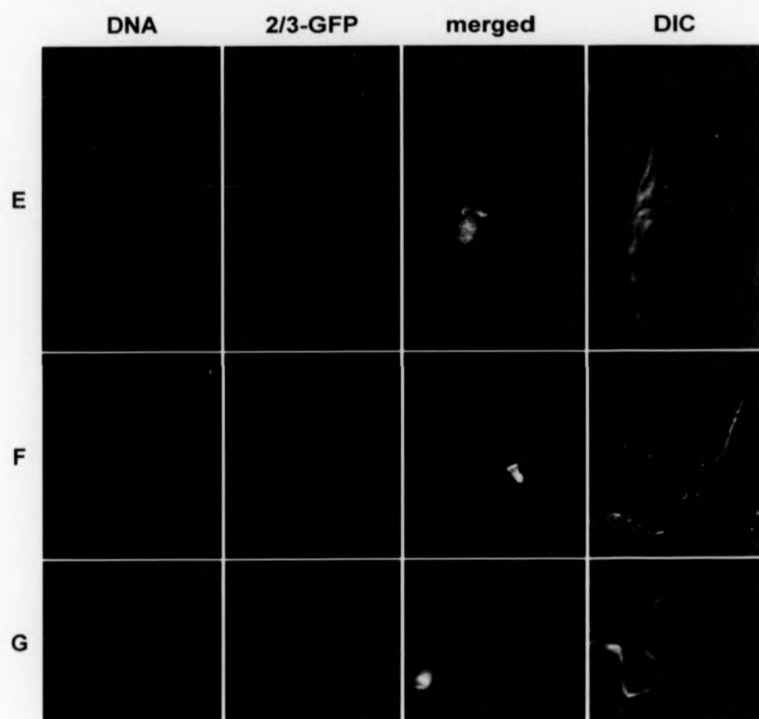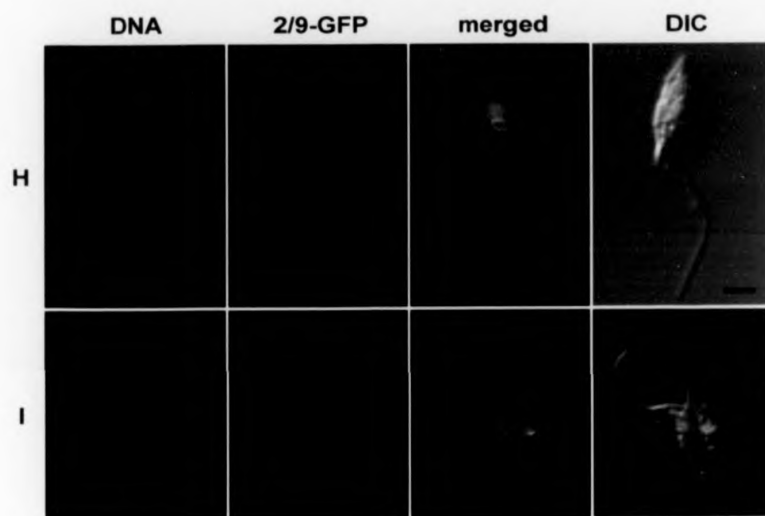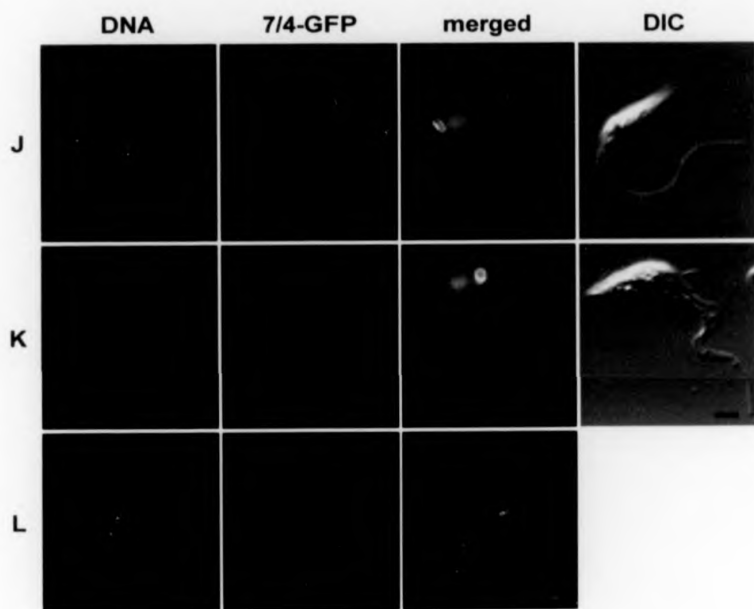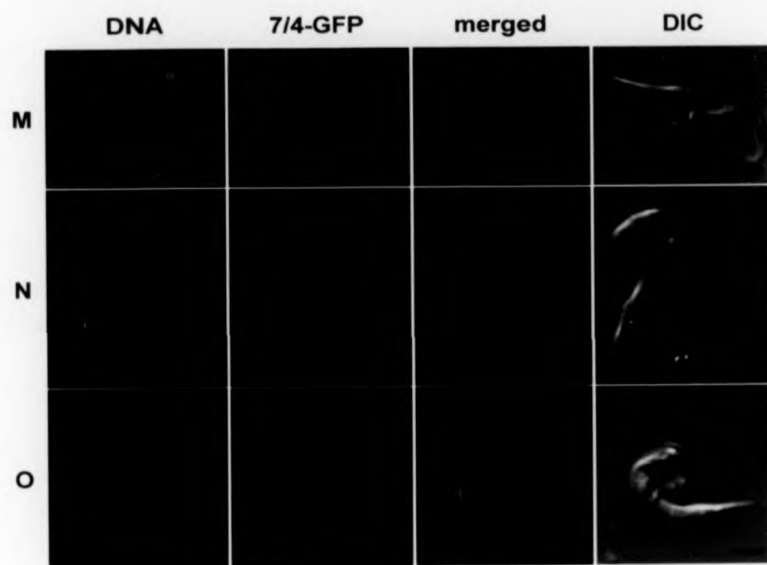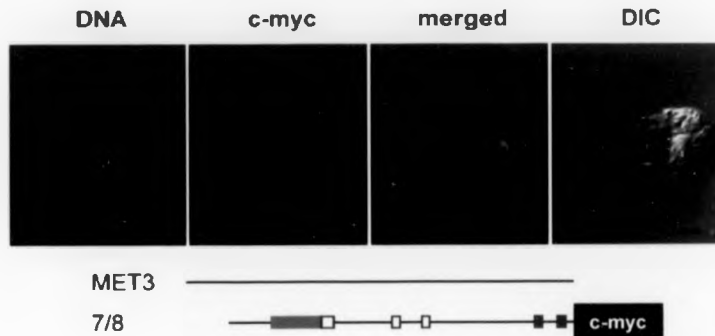
Figure 4.30: Nucleolar localisation of a control construct: truncated MET3 with a c-myc tag
The fusion protein 7/8 consists of an N-terminally truncated fragment of MET3, with a C-terminal c-myc epitope tag. The fusion protein was detected in fixed cells using an anti-c-myc antibody (c-myc). This fusion protein also localises to the nucleolus in epimastigotes. Nucleolar localisation thus occurs independently of a GFP-fusion.
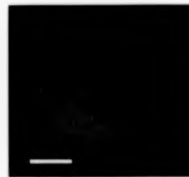
Figure 4.31: Epimastigote expressing GFP
When GFP is expressed on its own, it is found dispersed throughout the cell, including the nucleus and along the flagellum (image courtesy of Liz Bromley, (Bromley et al., 2004)). Scale bar = 5 µm.

In summary,

- Expression of MET3-GFP fusion proteins in epimastigotes identified two sequence elements that can direct the protein to the nucleolus.

- These elements can direct nucleolar localisation independently.

- Nuclear and nucleolar localisation could not be uncoupled.

4.4 GENERATION OF KNOCKOUT MUTANTS DEMONSTRATES THAT *MET3* IS NOT AN ESSENTIAL GENE

4.4.1 Deletion of the *MET3* gene by homologous recombination

We hypothesised that the metacyclic stage-specific protein MET3 would not be essential for parasite survival in the epimastigote stage, and therefore proceeded to delete both alleles by homologous recombination. The *MET3* gene was deleted from the genome of *T. cruzi* CL Brener by successive replacement of both alleles with drug selectable marker genes. Targeting constructs contained the hygromycin phosphotransferase gene (pkoMET3-HYG) or the puromycin acetyl transferase gene (pkoMET3-PAC) flanked by *GAPDH* processing signals, and with 5' and 3' untranslated sequences from the *MET3* locus to target homologous recombination (Figure 4.32A-B). For transfection, linear targeting fragments were delivered to epimastigote cells by electroporation (section 2.3.8).

In the first round of transfection, a number of drug resistant clones were obtained with both constructs. Deletion of one copy of *MET3* occurred in three clones transfected with the *HYG* construct, and in one clone transfected with the *PAC* construct. Based on restriction analysis, it was established that the same allele was deleted in all four single knockout clones (data not shown). The size of the restriction fragments suggest that the deleted allele corresponds to the one designated *CL-1* (Figure 4.1). The three *MET3/HYG* single knockout clones (sko) were used for the second round of transfection, using construct pkoMET3-PAC. One double knockout clone (dko) was obtained from each single knockout clone, and Southern blot analysis demonstrated that the *MET3* gene was no longer present in these clones (Figure 4.32C; data for the third clone not shown). Southern blot analysis also confirmed single insertions of the *HYG* and *PAC* gene, respectively, as well as linkage of the insertions to the *P5CDH* gene upstream of *MET3* (Figure 4.32C). Southern blot analysis therefore suggests correct integration of the targeting constructs into the *MET3* locus, resulting in *MET3* null mutant cell lines.

Figure 4.32 (overleaf): Generation of T. cruzi MET3 *null mutants*

(A) Targeting fragments pkoMET3-HYG and pkoMET3-PAC, and the *MET3* genomic locus. The targeting fragments contain the drug selectable marker genes *HYG* or *PAC*, flanked by *T. cruzi GAPDH* intergenic sequences (black lines) and *MET3* intergenic sequences (grey lines). H, *Hin* dIII restriction site. (B) *T. cruzi HYG* and *PAC* genomic loci generated by integration of the targeting fragments. (C) Genomic Southern blot of *Hin* dIII digested genomic DNA from wild-type (wt), two single knockout clones (sko 1 and 2), and two double knockout clones (dko 1 and 2). A *MET3* probe detects two bands in wt DNA, corresponding to the two alleles. In the sko clones, only one *MET3* allele remains. The *MET3* probe fails to hybridise to DNA from dko clones indicating that both alleles have been deleted. Hybridising fragments detected with *HYG*, *PAC* and *DPCDH* probes are of the expected size. Numbers on the left indicate size in kb.
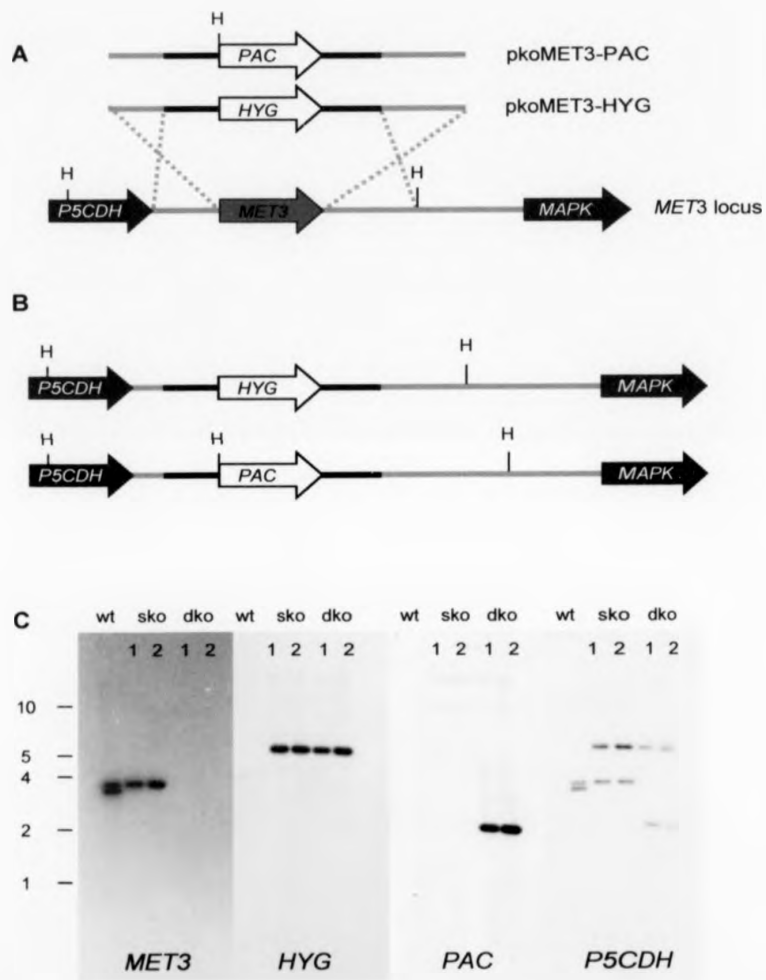
Figure 4.32

4.4.2 One of the two *MET3* alleles from CL Brener is a pseudogene

The single knockout cells were used to determine whether the two *MET3* transcripts detected in wild-type RNA (Figures 4.9 and 4.33) are derived from the same or separate alleles. Total RNA was extracted from early stationary phase cultures of single knockout and wild-type cells, and compared in Northern blots probed with the *MET3* coding sequence. Only one of the two transcripts is present in RNA from the single knockout, whereas two transcripts (with equal signal intensities) are detected in wild-type RNA (Figure 4.33A). This is consistent with the hypothesis that the two transcripts in the wild-type are derived from different alleles. The size difference between the two transcripts (estimated to be around 380 bp) could be due to alternative splicing of the two transcripts.

Western blot analysis was performed to compare the levels of MET3 protein in the wild-type and single knockout clones, and to confirm the null mutants no longer express any MET3. As expected, a strong band of 21 kDa was detected by the MET3 antiserum in wild-type lysates of stationary phase cells, but not in lysates from null mutants (Figure 4.33B). Surprisingly, no MET3 protein was detected in the single knockout lysate (Figure 4.33B). Northern blot analysis shows that these cells lack the longer *MET3* transcript. The levels of the shorter *MET3* RNA transcript are the same in single knockout and wild-type cells (Figure 4.33A). These results suggest that the smaller transcript is not translated into protein, and that the MET3 protein in the wild-type is derived from only the larger transcript. One of the MET3 alleles thus appears to be a pseudogene, and the single knockout cells are functional MET3 null mutants.
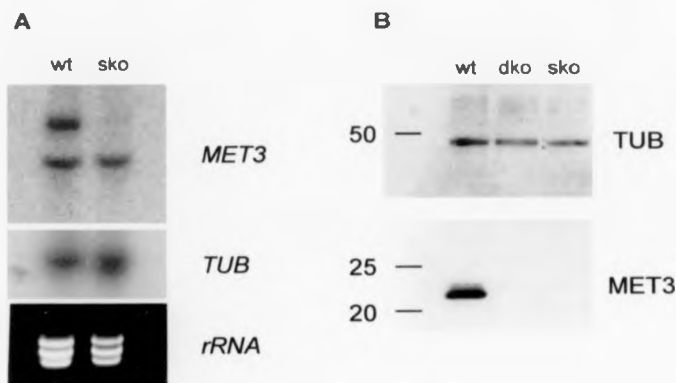
Figure 4.33: The MET3-CL2 *allele appears to be a pseudogene*

(A) Total RNA from wild-type (wt) and *MET3* single knockouts (sko) was analysed on Northern blots. A *MET3* probe detects two transcripts in the wild-type, but only one in the single knockout. This suggests that the two transcripts are derived from different alleles of *MET3*. (B) Western blot analysis of whole cell lysates from stationary cultures of wild-type (wt), double knockouts (dko) and single knockouts (sko) were analysed with the MET3 antiserum. MET3 protein is detected in wild-type, but is absent from the single- and the double knockouts. All lysates contained 4% metacyclic trypomastigotes. The blot was subsequently probed with an anti-β-tubulin antibody (TUB) to control for loading. Numbers on the left indicate molecular weight in kilo Daltons.

4.4.3 *MET3* is not essential for metacyclogenesis

Since the MET3 protein is expressed only during metacyclogenesis, deletion of the *MET3* gene was expected to have no effect on epimastigotes. Examination of epimastigote morphology by light microscopy revealed no differences between wild-type and MET3 knockout cells. *MET3* dko1 epimastigotes grew at the same rate and to the same density as wild-type epimastigotes (Figure 4.35B). *MET3* dko2 cells grew marginally slower than the *MET3* dko1 and wild-type cells (not shown).

The *MET3* double knockout cells were used to determine whether the *MET3* gene, which is upregulated during metacyclogenesis in wild-type cells, is essential for

155

development of metacyclics. To examine the effect of *MET3* gene deletion on the ability of cells to produce metacyclics, wild-type and *MET3* dko1 and dko2 cells were induced to differentiate *in vitro*, using Grace's insect medium (section 2.2.2). Initial analysis of the cell lines revealed that metacyclics were present in wild-type controls and in both the *MET3* knockout cultures (Figure 4.34A). *MET3* is therefore not essential for development of metacyclics. However, a quantitative comparison showed that the *MET3* knockout cell lines produced significantly fewer metacyclics than the wild-type (Figure 4.34B). Wild-type cultures contained four times more metacyclics on average than *MET3* dko1 cells (11.1% in wild-type, compared to 2.8% in dko1, t-test: $p<0.001$). Conversely, the number of intermediate forms in *MET3* dko1 cultures was more than twice that in wild-type cultures (Figure 29B; 23.5% in dko1 compared to 10.5% in wild-type, t-test: $p<0.001$). Interestingly, one in four dko1 intermediates had an elongated nucleus and an anterior kinetoplast (type 3 in Figure 4.34A; see also section 3.2). In the wild-type, less that one in ten intermediates is of this morphological type. The number of epimastigotes was not significantly different (78.3% in wild-type compared to 73.6% in dko1, t-test: $p>0.2$). This indicated that the same proportion of cells initiated differentiation, but the dko1 cells progressed more slowly to the fully developed metacyclic form. The *MET3* dko2 cells also produced significantly fewer metacyclics than wild-type controls (Figure 4.34B); 0.3% in dko2 compared to 11.1% in wild-type, t-test: $p<0.001$). Contrary to the results obtained with *MET3* dko1 cells, however, the *MET3* dko2 cells did not show a significant accumulation of intermediate forms. These initial experiments thus showed that *MET3* is not essential for the development of metacyclics, but revealed a difference in the rate of metacyclogenesis between wild-type and *MET3* knockouts. They also revealed slight differences between the two knockout cell lines.

A   **B**
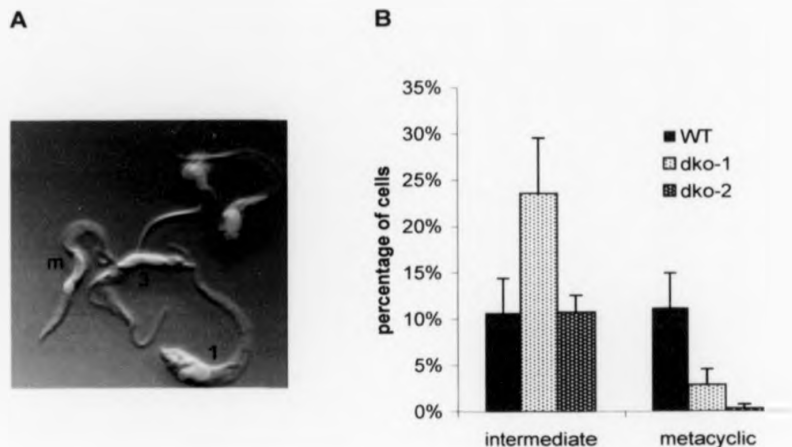
Figure 4.34: Differentiation of T. cruzi *wild-type and* MET3 *null mutants*

(A) Image of cells from stationary culture of *MET3*-dko1 shows that these cells can develop to inter-
mediate forms and metacyclics (m). The image shows three types of intermediates which differ in the
position of the kinetoplast relative to the nucleus, and shape of the nucleus (see also Table 3.1):
(1) kinetoplast on the side of a round nucleus; (3) elongated nucleus and anterior kinetoplast;
(4) kinetoplast on the side of elongated nucleus. A stationary epimastigote is also shown (e). Nuclear
and kinetoplast DNA are stained with TOTO-3 (pseudo-coloured red). (B) Differentiation of cells to
intermediate forms and metacyclics was determined in stationary phase cultures, eight days after
addition of Grace's medium. Giemsa stained cells were scored based on morphology. The two *MET3*
knockout cell lines (dko1 and dko2) produce significantly fewer metacyclics than the wild-type (WT).
The dko1 cells produce significantly more intermediate forms than WT. Each graph indicates the
average percentage and the standard deviation in eight cultures. 200 cells were counted for each
culture.

To determine if impaired metacyclogenesis was a consequence of the loss of *MET3*, a copy of the gene was re-introduced into both knockout cell lines. Plasmid pTEX-MET3-9E10, which encodes full-length MET3, with the c-myc epitope 9E10 at the C-terminus, was used for transfection of knockout epimastigotes. Expression of genes from the pTEX vector is constitutive, and high levels of the MET3-9E10 protein were detected in Western blots from epimastigote lysates of transfected cells (Figure 4.35A). Overexpression of the MET3-9E10 protein did not affect epimastigote growth rates (Figure 4.35B).
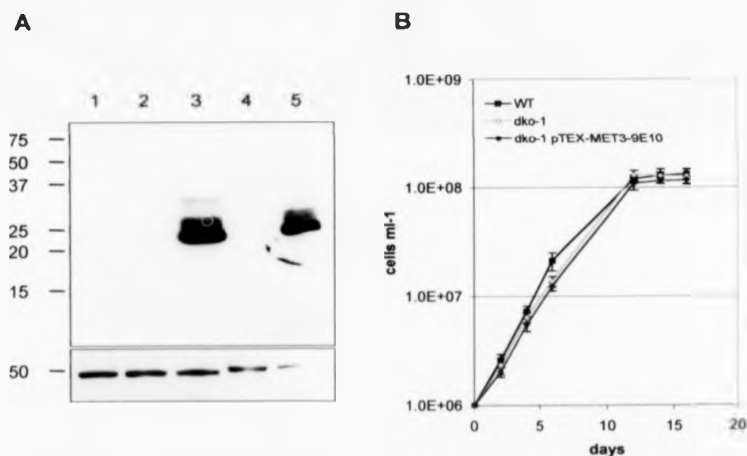
Figure 4.35: Expression of MET3-9E10 protein in MET3 *null mutants*

(A) Epimastigote protein lysates from different cell lines were analysed on Western blots. Lane 1: wild-type; lane2: *MET3*-dko1; lane 3: *MET3*-dko1 pTEX-MET3-9E10; lane 4: *MET3*-dko2; lane 5: *MET3* dko2 pTEX-MET3-9E10. Upper panel: MET3 antiserum detects MET3-9E10 in complemented knockout cells (lanes 3 and 5). Non-transfected wild-type and knockout cells express no MET3 (lanes 1, 2 and 4). Lower panel: The same blot was probed with anti-β-tubulin antibody as a loading control. (B) Overexpression of MET3-9E10 does not alter the growth rate of epimastigotes. Each point represents the average cell count from two experiments performed in triplicate. Error bars indicate standard deviations.

The *MET3* dko1 cells complemented with MET3-9E10 were analysed systematically for their ability to differentiate to metacyclics *in vitro*. Overexpression of MET3-9E10 clearly does not restore the rate of metacyclogenesis to wild-type levels. The proportion of metacyclics in the MET3-9E10 complemented cell line is similar to that observed in *MET3* dko1 cultures and significantly lower than in the wild-type controls (Table 4.2 and Figure 4.36). The MET3-9E10 complemented cell line produced slightly more intermediate forms than the wild-type (Figure 4.36), but the accumulation was not as marked as in MET3 dko1 cultures (Figures 4.34B and 4.36). An accumulation of the intermediate cells with elongated nuclei was not apparent in the MET3-9E10 complemented cell line. Similarly, complementation of *MET3* dko2 cells did not result in a noticeable increase in the number of metacyclics. This was not systematically quantified, however. In summary, these complementation experiments do not support the hypothesis that it is the loss of *MET3* that caused a reduction in the number of fully developed metacyclics.

Table 4.2: Differentiation of T. cruzi WT, MET3-dko1 and complemented cell line

T. cruzi wild-type (WT), *MET3* null mutant (dko1), and dko1 complemented with pTEX-MET3-9E10 were assayed for their ability to differentiate in 20% Grace's medium. Samples were taken six, eight and ten days after addition of Grace's medium, fixed and stained with Giemsa. 200 cells from each sample were counted to determine the relative proportion of epimastigotes, intermediates and metacyclics. Each value represents the average percentage of the respective cell type as determined in two experiments performed in triplicate. Numbers in brackets indicate standard deviation.

cell type	epimastigote			intermediate			metacyclic		
DAY	6	8	10	6	8	10	6	8	10
WT	89.0 (2.3)	79.8 (2.2)	77.1 (3.0)	5.6 (1.6)	9.2 (1.3)	11.0 (1.4)	5.4 (1.4)	11.0 (1.5)	11.9 (2.4)
dko1	86.8 (3.1)	83.3 (1.8)	83.3 (3.4)	11.7 (3.2)	14.3 (1.4)	14.4 (3.9)	1.6 (0.7)	2.3 (0.8)	2.3 (0.7)
dko1 pTEX-MET3-9E10	87.6 (1.1)	88.3 (2.5)	88.1 (1.5)	11.3 (1.3)	10.8 (2.5)	10.3 (1.7)	1.1 (0.7)	0.8 (0.6)	1.6 (0.4)
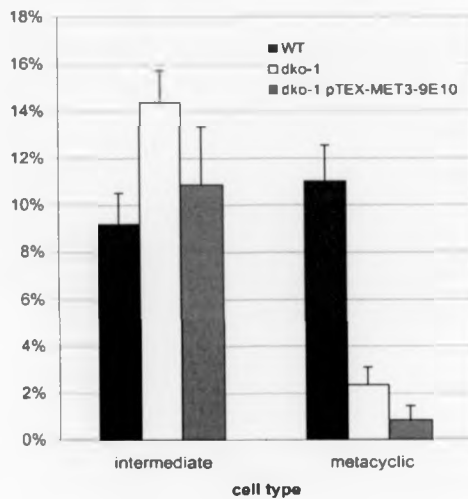
Figure 4.36 Expression of MET3-9E10 *in knockouts does not restore wild-type levels of metacyclics*
Differentiation of cells to intermediate forms and metacyclic trypomastigotes in cultures of wild-type
(WT), *MET3* dko1 and *MET3* dko1 pTEX-MET3-9E10. The *MET3* knockout cells produce signifi-
cantly fewer metacyclics than the wild-type. The *MET3* dko1 cells complemented with *MET3-9E10*
produce equally low numbers of metacyclic trypomastigotes. Expression of *MET3-9E10* thus does not
restore the wild-type rate of metacyclogenesis. Each graph indicates the average percentage and the
standard deviations from six cultures. 200 cells were counted for each culture.

4.4.4 *MET3* is not essential for completion of the *T. cruzi* life-cycle

To verify whether the metacyclics that develop from *MET3* knockout cells are genuine metacyclics, their ability to infect macrophages was tested *in vitro*. Qualitatively, no difference was observed between the wild-type and *MET3* dko1 cells in the time course and severity of infection when equal numbers of stationary phase cells were used to infect macrophages. Like the wild-type controls, *MET3* dko1 cells can invade macrophages, differentiate into amastigotes (Figure 4.37A-B), replicate intracellularly (Figure 4.37C-D), and emerge as trypomastigotes from the ruptured macrophages (Figures 4.37E-F and 4.38C-D).

These results show that *MET3* is not essential for development of metacyclics and for completion of the life-cycle *in vitro*. The *MET3* knockout mutants are able to produce metacyclics that are morphologically indistinguishable from wild-type metacyclics, and they are able to infect macrophages *in vitro*, replicate as amastigotes and develop into trypomastigotes. The two *MET3* knockouts do show a significant reduction in the number of metacyclics, compared to wild-type, but expression of MET3-9E10 does not restore the rate of metacyclogenesis to wild-type levels. Furthermore, the two *MET3* knockout cell lines dko1 and dko2 are dissimilar with respect to the production of intermediate forms and the growth rate of epimastigotes. Therefore, it is very likely that factors other than the loss of *MET3* caused the differences between clones.
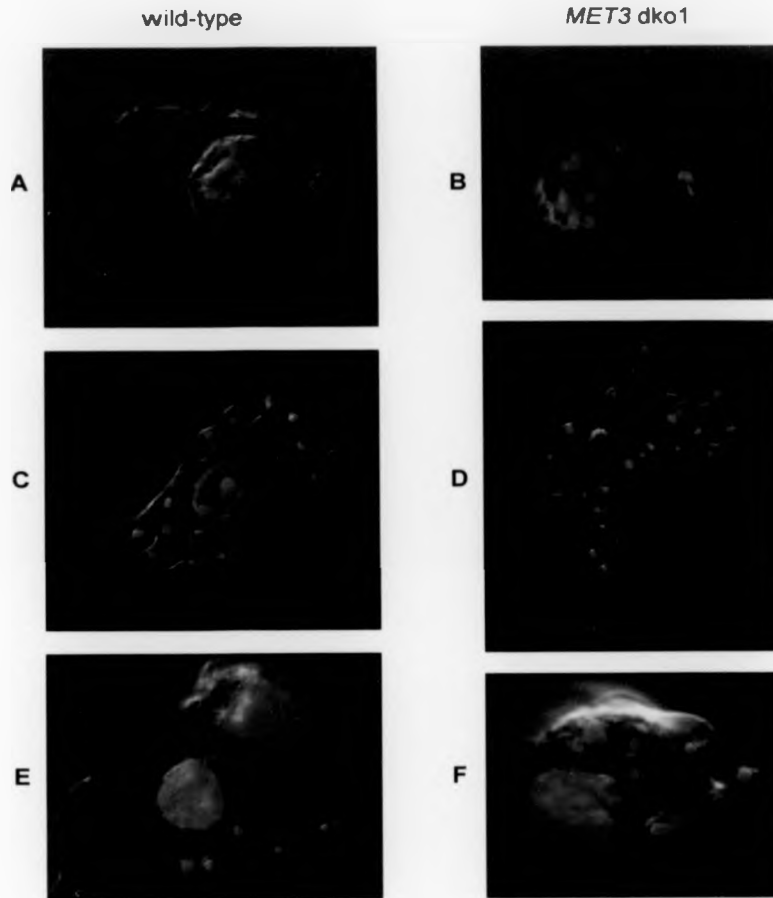
wild-type   MET3 dko1

Figure 4.37: MET3 *knockouts can complete the life-cycle* in vitro

Mouse macrophages were infected *in vitro* with stationary phase *T. cruzi* wild-type or *T. cruzi MET3*-dko1 cells as indicated. Representative DIC images of three time points are shown, merged with the corresponding images showing the DNA (stained with TOTO-3, pseudo-coloured red) to visualise the *T. cruzi* kinetoplast and nucleus, and the macrophage nucleus. Panels **A** and **B**: amastigotes inside a mouse macrophage 3 hours after infection. Panels **C** and **D**: amastigotes six days after infection. The large number of amastigotes inside the macrophage indicates that intracellular replication has occurred. Panels **E** and **F**: trypomastigotes emerging from ruptured macrophages six days after infection. Scale bar = 2 μm.
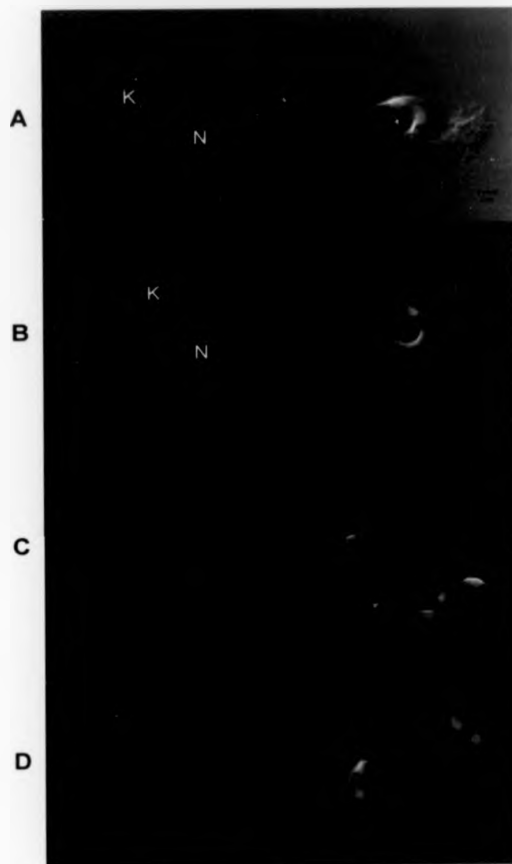
Figure 4.38. Wild-type and MET3-dko1 *cell-derived trypomastigotes*

Six days post infection, trypomastigotes are abundant in the supernatant of *T. cruzi*-infected macro-phages. Wild-type trypomastigote populations are a mix of broad cells with a round nucleus (**A**) and slender cells with an elongated nucleus (**B**). Most *MET3*-dko1 trypomastigotes are broad (**C** and **D**). The mean length of nuclei in wild-type is 2.6 µm compared to 2 µm dko1 (t-test: p 0.001, n 30). Panels on the left show kinetoplast (K) and nuclear DNA (N) stained with TOTO-3. Panels on the right are merged fluorescence and DIC images. Scale bar = 2 µm.

4.4.5 Nuclear morphology in *MET3* null mutants

Trypomastigotes from the supernatant of macrophage cultures infected with *MET3* knockout cells appeared smaller, and more stumpy than trypomastigotes from wild-type controls. Trypomastigotes in the mammalian host are pleomorphic, consisting of broad trypomastigotes, and slender trypomastigotes (Brener, 1973; Tyler and Engman, 2001). The significance of these different forms has remained controversial since their first description by Chagas in 1909 (Brener, 1973). Recent research has shown that the slender form precedes the broad form (Tyler and Engman, 2001). Microscopic examination of fixed *MET3* dko1 trypomastigotes cells revealed that these cells are predominantly broad, with round nuclei (Figure 4.38C-D). In wild-type trypomastigote populations on the other hand, a mix of slender cells with elongated nuclei (Figure 4.38B) and broad cells (Figure 4.38A) are observed, similar to that described in the literature. To obtain a quantitative measure of the difference between the wild-type and the *MET3* dko1 trypomastigotes, the nuclei of fixed cells were measured. The average length of nuclei is significantly greater in wild-type than in *MET3* dko1 cells (mean in wild-type 2.6 μm compared to 2 μm in dko1, t-test: $p < 0.001$, n = 30). It remains to be determined whether this accumulation of broad trypomastigotes in the *MET3* dko1 cells is due to the deletion of *MET3* or a consequence of other differences between cloned cell lines.

In the nuclei of wild-type cells, *MET3* expression coincides with elongation of the nucleus and dispersal of nucleolar antigens when epimastigotes differentiate to meta-cyclics. We therefore investigated whether *MET3* is required for remodelling of the nuclear and nucleolar structure during metacyclogenesis (section 1.2.2). During exponential growth, the nuclei of *MET3* knockout epimastigotes are round, with a single nucleolus, indistinguishable from wild-type nuclei (Figure 4.39A). As *MET3* knockout cells enter stationary phase, the nucleolar antigen L1C6 disperses (Figure 4.39B), like in wild-type cells. Upon differentiation to amastigotes, the nucleus becomes round again, and a single L1C6 signal is clearly visible in the nucleolus of *MET3* knockout cells (Figure 4.39C-D). These data show that MET3 is not necessary to cause dispersal of nucleolar antigens, and it is not essential for re-assembly of the

nucleolus in amastigotes. Deletion of MET3 thus does not result in a readily discernible alteration to nuclear structure in epimastigotes, metacyclics and amastigotes.

In summary,

- We have generated *MET3* knockout mutants.
- *MET3* is not essential for development. Knockouts can differentiate to metacyclics and complete the life-cycle *in vitro*.
- Knockouts do show a significant reduction in the number of metacyclics but expression of MET3-9E10 does not restore the rate of metacyclogenesis to wild-type levels.
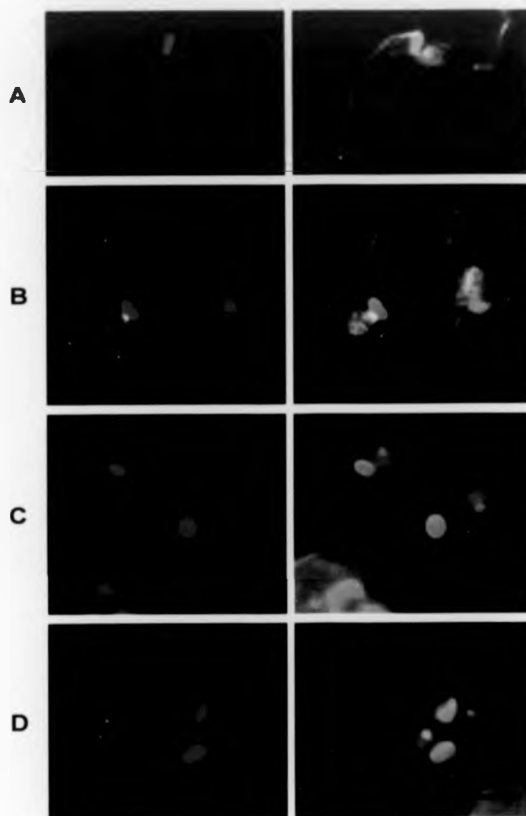
Figure 4.39: Deletion of MET3 *has no apparent effect on the dispersal and re-assembly of the nucleolar antigen LIC6*

Images on the left show the merged fluorescence image of LIC6 antibody staining (pseudo-coloured red) and TOTO-3 DNA staining (pseudo coloured blue). Images on the right show merged images of fluorescence and DIC. (**A**) *MET3*-dko1 epimastigotes in the log phase of growth have a clearly defined nucleolus, which is stained with the LIC6 antibody. (**B**) In the nucleus of *MET3*-dko1 epimastigotes in stationary phase, the LIC6 antigen is dispersed. *MET3*-dko1 and dko2 stationary phase cells (including infectious forms) were used to infect mouse macrophages. Panels **C** and **D** show amastigotes that subsequently developed inside the macrophage. The LIC6 antigen is concentrated again in a single nucleolus. (**C**) *MET3*-dko1 amastigotes, (**D**) *MET3*-dko2 amastigotes. Scale bar = 2 μm.

4.5 DISCUSSION

4.5.1 Identification and characterisation of *MET3* from *T. cruzi* CL Brener: stage-specific expression and localisation

As a result of this study, we now have detailed information about the expression and subcellular localisation of the MET3 protein. MET3 was originally identified in a screen for RNAs that are upregulated during metacyclogenesis in *T. cruzi* strain Dm28c (Yamada-Ogatta et al., 2004). Here, we have identified and characterised the *MET3* homologue of the genome reference strain *T. cruzi* CL Brener. Surprisingly, the *MET3-CL* coding sequence is slightly shorter than *MET3-DM28c*. Whether this is another example of the genetic diversity of *T. cruzi*, or due to a sequencing error in the *MET3-Dm28c* sequence remains to be investigated. Northern and Western blot analysis was used to show that up-regulation of MET3 expression during metacyclogenesis in the CL Brener strain follows a similar pattern to that described for the Dm28c strain (Yamada-Ogatta et al., 2004), where it was shown that MET3 protein is expressed in metacyclics, but not in log phase epimastigotes. In addition, we demonstrate specifically that in CL Brener cells, MET3 protein expression is induced before the repositioning of the kinetoplast has been completed. The expression of MET3 protein thus clearly distinguishes intermediates and metacyclics from stationary epimastigotes, and represents a marker for metacyclogenesis in *T. cruzi* strains from both lineages.

We extended protein expression analysis to include the mammalian stages of the parasite, which have not been studied in the Dm28c strain. Our immunofluorescence results show that in *in vitro* infection experiments, MET3 disappeared from *T. cruzi* cells soon after entry into the macrophage, but only after the morphological transformation from metacyclic to amastigote had been completed. All macrophage-derived trypomastigotes examined in three separate experiments were negative for MET3 immuno-staining. A more comprehensive analysis may be undertaken in the future to confirm the observed pattern of MET3 expression in the mammalian stages by an independent method, for example by comparing the levels of RNA transcripts in these stages.

167

Cell-derived trypomastigotes have a cell and nuclear morphology very similar to metacyclics, and both cell types are infective and non-replicating. However, they differ with respect to their commitment to the production of amastigote forms (Tyler and Engman, 2001) and expression of surface proteins (Araya et al., 1994). They represent distinct life-cycle stages, but markers to distinguish between these cell types are scarce. MET3 represents a new metacyclic-specific marker and will be valuable in the study of the infectious forms of *T. cruzi*.

At present we can only speculate about the mechanisms that regulate MET3 expression. The results of this study show that stage-regulated MET3 protein expression in CL Brener correlates with differences in levels of *MET3* RNA, similar to the observed pattern in Dm28c (Yamada-Ogatta et al., 2004). A number of stage-regulated *T. cruzi* genes have been shown to contain regulatory elements within their 3' UTR (D'Orso and Frasch, 2001a; Nozaki and Cross, 1995; Teixeira et al., 1995). Based on the size of *MET3* transcripts in Northern blots, the estimated size of the 3' UTR is 2.8 kb (Figure 4.9), and it is possible that the steady-state *MET3* transcript levels are controlled by *cis*-acting elements within this long 3' UTR. Such elements may stabilise or destabilise *MET3* mRNA in a stage-specific manner. However, epimastigotes contain detectable levels of *MET3* transcripts, but no detectable protein (Figure 4.11). This indicates that MET3 expression is regulated at multiple levels. There is evidence to suggest that a number of mRNAs, including *MET3-Dm28c* are selectively recruited to polysomes in differentiating cells (Avila et al., 2001; Dallagiovanna et al., 2001; Fragoso et al., 2003; Yamada-Ogatta et al., 2004). The *MET3* gene may provide a useful model to further dissect the mechanisms of metacyclic-specific gene expression.

To investigate this further, reporter gene assays could be used to test the hypothesis that the MET3 3'UTR contains regulatory elements. This would involve transfection of *T. cruzi* with reporter constructs containing a reporter gene (such as luciferase) fused to fragments of the *MET3* 3' UTR. Expression levels of different constructs in transfected *T. cruzi* could then be quantified and compared, to localise and identify regulatory sequence elements. A comparison between CL Brener and the MET3-Dm28c sequences could help to pinpoint conserved sequence elements (or possibly

168

RNA secondary structures) that might be important in the regulation of MET3. However, the extent of the MET3-Dm28c 3'UTR is presently unclear. The MET3-Dm28c cDNA deposited in GenBank (metaciclina III, GenBank AF312553) contains both the spliced leader sequence at the 5' end and a poly A tail at the 3' end, implying that this corresponds to the entire mRNA sequence. However, this sequence is only 891 bp in length, and thus only a third of the size of the CL Brener transcripts. The published Dm28c Northern blots unfortunately lack size markers (Yamada-Ogatta et al., 2004). It will be important to verify the size of the *MET3-Dm28c* transcripts experimentally, before speculating about the significance of such a large difference in transcript length between the strains.

Interestingly, the *MET3* knockout experiments suggest that *MET3-CL2*, is a pseudogene. Wild-type cells contain two *MET3* RNA transcripts that differ in size by an estimated 0.4 kb. Deletion of the *MET3-CL1* allele resulted in a single-knockout cell line (sko1), which produced only the shorter transcript, and expressed no detectable MET3 protein (Figure 4.33). We therefore hypothesise that the *MET3 CL-2* transcript is *trans*-spliced at a different position to *MET3-CL1*, resulting in a shorter transcript that is not translated. Amplification of the short *MET3* transcript from sko1 by RT-PCR and sequencing of the resulting cDNA will determine whether this hypothesis is correct.

4.5.2 Evidence that MET3 associates with nucleolar components

The results from immuno-fluorescence studies presented here and transmission electron micrographs of *T. cruzi* Dm28c (Yamada-Ogatta et al., 2004) clearly established that MET3 is a nuclear protein. This is consistent with the presence of multiple predicted NLS in the MET3 sequence (Figure 4.7). The authors of the Dm28c study report that MET3 specifically associated with condensed chromatin (Yamada-Ogatta et al., 2004) but given the low resolution of the presented EM images, and the lack of any biochemical data, it is not clear how this conclusion was reached. The detailed analysis of the subnuclear localisation of MET3 in CL Brener presented here provides several lines of evidence to suggest that MET3 is associated with nucleolar components.

Confocal microscopy was used to gain information on the subcellular localisation of MET3. Successful purification of recombinant MET3 from *E. coli* and subsequent generation of a MET-specific antiserum provided a tool to visualise native MET3 in wild-type *T. cruzi*. In amastigotes early after macrophage infection, the nucleolus is visible in TOTO-3 stained nuclei as area of low DNA staining. In these cells, the MET3 signal was clearly confined to the nucleolus. In metacyclics and intermediate forms, nucleolar components become dispersed. In these cells, a dispersed MET3 signal was observed, clustered in areas that stained weakly for DNA. The same pattern was observed when the cells were stained for the nucleolar antigen L1C6. The extent to which the dispersed antigens co-localise in a single cell remains to be determined. This could be achieved by using directly labelled primary antibodies for further studies (both MET3 and L1C6 antibodies were raised in mice), or by using a different antibody against a nucleolar protein.

In a second set of experiments, a number of transgenic *T. cruzi* cell lines were created which expressed different tagged versions of MET3. Full-length MET3, tagged with either a c-myc epitope or GFP, localised exclusively to the nucleolus in epimastigotes, and the association of tagged MET3 with the nucleolar antigens L1C6 and NOG1 was shown directly in double labelling experiments. At the onset of stationary phase, these recombinant proteins dispersed in a pattern that was again extremely similar to that of L1C6 and native MET3. Because MET3 is not normally expressed in epimastigotes, localisation of recombinant MET3 to the epimastigote nucleolus does not prove that MET3 is functionally associated with the nucleolus in wild-type cells. However, the results from wild-type cells and from cells expressing recombinant MET3 taken together provide strong evidence that MET3 associates with nucleolar components under a variety of conditions.

4.5.3 Identification of elements that mediate nucleolar targeting

To gain some insight into the domain structure of the MET3 protein, a number of fusion proteins were expressed in *T. cruzi* to identify elements in the MET3 sequence

that direct nucleolar localisation. To our knowledge, this is the first detailed analysis of nucleolar targeting of a *T. cruzi* protein.

Expression of a number of fusion proteins identified two sequence elements, located at the N- and C-terminus respectively, which can direct GFP-fusion proteins to the nucleolus independently of each other (Figure 4.27). A consensus nucleolar localisation signal does not exist, and the signals that can direct proteins to the nucleolus are diverse (Melese and Xue, 1995). However, a number of cases have been documented where the domains critical for nucleolar localisation contain multiple clusters of basic amino acids (Hoek et al., 2000). The results obtained with the MET3 fusion proteins are consistent with this: The N-terminal sequence conforms to the consensus of a bipartite NLS, which is two clusters of basic residues separated by 9-12 spacer residues (Cokol et al., 2000). The C-terminal sequence also contains two clusters of basic residues, separated by 8 spacer residues. There are some similarities between the smallest MET3 fragment that conferred specific nucleolar localisation (construct 14/6, which contains the last 46 amino acids of MET3), and a 32 amino acid domain that is sufficient for nucleolar targeting of *T. brucei* ESAG8 (Hoek et al., 2000). Like the C-terminal domain of MET3, this ESAG8 domain contains two clusters of basic residues, as well as hydrophobic and a small number of acidic residues. A second similarity between the targeting of MET3 and ESAG8 is the fact that all fragments tested show either exclusive nucleolar localisation or completely non-specific localisation.

A bipartite NLS is also contained within the first 40 amino acids of *T. cruzi* histone H2B. When this fragment was fused to GFP, the fusion protein (H2B40-GFP) localised to the nucleolus in *T. brucei* (Marchetti et al., 2000) and *T. cruzi* (Elias et al., 2001b), while full-length H2B localised to the nucleus (Marchetti et al., 2000). The nucleolar localisation of H2B40-GFP is considered to be an experimental artefact due to a serendipitous interaction of the fusion protein with unidentified nucleolar components. The miss-targeting of H2B40-GFP to the nucleolus highlights one caveat of relying on a single reporter system to determine protein localisation (Marchetti et al., 2000). However, it appears highly unlikely that the localisation of the MET3 fusion proteins in the present study is due to a similar serendipitous inter-

171

action for two reasons. First, two constructs with the c-myc epitope tag instead of GFP (constructs MET3-9E10 and 7/8), were also targeted to the nucleolus. By contrast, the association of histone H2B40-GFP with the nucleolus was completely abolished in an altered context, by fusing H2B40-GFP to β-Gal (Marchetti et al., 2000). Second, wild-type MET3 in early amastigotes clearly localised to the nucleolus, confirming that the MET3 sequence contains the information for nucleolar targeting.

Targeting of proteins to membrane bounded organelles such as the mitochondrion, nucleus or glycosome typically relies on a specific targeting signal. By contrast, no membrane separates the nucleolus from the nucleoplasm. The current consensus is that nucleolar localisation is dependent on a protein's interaction with other nucleolar components, and the sequences that mediate nucleolar localisation are more likely to act as retention signals rather than classical targeting signals (Carmo-Fonseca et al., 2000; Melese and Xue, 1995). Based on sequence analysis that identified a region reminiscent to a RNA binding motif in the N-terminal half of MET3, we would speculate that interaction with a specific nucleolar RNA may cause MET3 to localise to the nucleolus. Since a C-terminal element can direct localisation independently, MET3 is likely to interact with a number of nucleolar components. Identification of factors that interact with MET3 will be a crucial next step in elucidating its function.

4.5.4 Functional analysis of MET3

As a result of genome sequencing projects, the complete catalogue of genes is now available for many organisms including *T. cruzi*. However, a large proportion of putative coding sequences have no known function (Hall et al., 2003). In the absence of any similarity to previously characterised genes, it is a major challenge for 'post genomic' research to assign a function to these novel genes. The description of the subcellular localisation of the MET3 protein is a first step towards this goal.

It is important to note that the association of MET3 with nucleolar components does not automatically imply a role in ribosome biogenesis. Several exciting studies in recent years have put the nucleolus in the spotlight, as it emerged that the nucleolus

172

is more than a 'ribosome factory'. The nucleolus is also a site for modification and maturation of various small RNAs and ribonucleoprotein complexes (Gerbi et al., 2003), and it plays a role in the control of cell proliferation and gene silencing (Carmo-Fonseca et al., 2000; Garcia and Pillus, 1999). Comprehensive studies of the human nucleolar proteome found that the function of nearly one third of the 271 detected proteins is unknown, and many proteins in the nucleolus have no obvious function in ribosome biogenesis (Andersen et al., 2002; Scherl et al., 2002).

We have attempted to further elucidate the biological function of *MET3* by transfection-based approaches. The results of our studies show that neither the deletion of the gene nor the over-expression of the protein results in an easily discernible phenotype. The most important question to ask was whether MET3 is essential for differentiation. The answer is clearly negative. The results from the MET3 knockout experiments show that metacyclics can develop in the absence of MET3 expression. These metacyclics are infective to macrophages, and can progress through the intracellular stages of the *T. cruzi* life-cycle. MET3 is therefore neither essential for production of metacyclics nor for the development of subsequent developmental stages *in vitro*.

The second question was whether MET3 plays an important role during metacyclogenesis, and whether manipulation of MET3 would affect the rate of differentiation. The two null mutant cell lines that were examined in detail both showed a significant and reproducible reduction in the number of metacyclics that developed compared to the parental wild-type population. Initially this suggested that MET3 might play a regulatory role during differentiation. If all the results are taken together, however two observations suggest that alternative explanations are more likely to account for this result.

First, the two *MET3* knockout cell lines behave differently with respect to the production of intermediate forms. The *MET3* dko1 cells accumulate high numbers of intermediate forms in stationary cultures. Such an accumulation of intermediate forms is not apparent in dko2 cells. Southern blots are consistent with a single integration of the targeting fragments into the *MET3* locus and revealed no differences between dko1 and dko2 cells. The genetic manipulation itself is thus unlikely to

173

account for the differences between dko1 and dko2. These differences could be due to variations between individual cells in the parental wild-type population. Clonal propagation of transfected cells would have brought such variations to light. During the process of transfection and selection, those cells that are able to proliferate under stress conditions have a strong selective advantage over those cells that readily and irreversibly differentiate to non-dividing metacyclics. The reduced rate of meta-cyclogenesis in the dko1 and dko2 cell lines may therefore be a consequence of this selection, and not caused by the loss of *MET3*. More importantly, expression of an ectopic copy of *MET3* (*MET3-9E10*) failed to complement the phenotype of the knockouts. This also suggests that the reduced rate of metacyclic development in the null mutants was not caused by the deletion of *MET3*.

However, there are a number of potential technical problems that could account for the failure of pTEX-MET3-9E10 to complement the loss of *MET3*. The simplest explanation would be that the c-myc tag at the C-terminus of the protein interferes with its function. This can be tested by replacing MET3-9E10 with an untagged version of MET3. Two other important differences between the expression of native MET3 and expression of MET3-9E10 from the pTEX vector are the levels and the timing of expression. As Figure 4.35 shows, MET3-9E10 is expressed at very high levels in epimastigotes, and the presence of higher molecular weight bands indicates that there may be modified forms of the protein. There could be mechanisms for modifying the MET3 protein that prevent it from performing its function at the wrong time in the life-cycle. In short, expression of MET3-9E10 from the pTEX vector does not closely resemble wild-type MET3-expression. A better strategy for complementation of the null mutants may be the re-insertion of the *MET3* coding sequence, flanked by the *MET3* 5' and 3' UTR, which are likely to contain the elements that mediate stage-specific expression. A neomycin phosphotransferase gene could be used for selection, and one of the two drug selectable marker genes already integrated in the *MET3* locus could serve as a targeting sequence.

At present, we cannot exclude the possibility that integration of the knockout construct may have affected the expression of genes flanking *MET3*. This could also explain why ectopic expression of *MET3-9E10* does not complement the phenotype,

174

yet provides no explanation for the differences between dko1 and dko2 cells. At present, nothing is known about the expression of the protein kinase downstream, and the pyrroline-5-carboxylate dehydrogenase (*P5CDH*) upstream of *MET3*. Interestingly, studies on the effect of intermediary metabolites on *T. cruzi* differentiation discovered that substances which inhibit P5CDH also inhibit metacyclogenesis (Krassner et al., 1990, see also section 1.4.2). It might therefore be worth comparing transcript levels between wild type and knockout cells to see if there is any evidence of *P5CDH* down-regulation in the knockout cells.

A number of *T. cruzi* cell lines were generated that constitutively over-expressed c-myc and GFP tagged versions of MET3. The results prove that MET3 can be expressed at high levels in epimastigotes without any apparent deleterious effect on the cells. Over-expression of MET3-9E10 in *MET3* knockout cells clearly did not affect the growth rate and the rate of metacyclogenesis (Figure 4.35). Epimastigotes that express tagged versions of MET3 in a wild-type background appeared no different from wild-type epimastigotes, which do not express any form of MET3. The rate of metacyclogenesis was not investigated systematically, since this cons-titutive expression system was considered suboptimal to conduct such studies. By inserting the MET3 coding sequence into vector pTcINDEX (Figure 1.5), a vector (pTcINDEX-MET3) has now been constructed that facilitates tetracycline-inducible protein expression in a CL Brener cell line (Martin Taylor, submitted for publi-cation). This will provide a system to study the effect of inappropriate MET3 expression in a controlled manner.

4.5.5 Approaches to test the predicted nucleic acid binding activity and identify molecules that interact with MET3

The analysis of the knockout cell lines was complicated by the absence of a hypothe-sis concerning the exact biochemical function of the MET3 protein. The specific subnuclear localisation of MET3 suggests that it can interact with one or several nucleolar components. MET3 may interact with nucleic acids or proteins, either individually, or as part of a larger ribonucleoprotein complex. It will be essential to

identify those interacting factors in order to understand the biochemical function of MET3 and its role during metacyclogenesis.

Based on the results from sequence analysis presented in this study, it seems probable that MET3 could bind RNA or DNA directly. The MET3 sequence has features characteristic of RNA binding proteins, yet it does not belong to any previously characterised protein family. With its basic nature (pI >11), and high proportion of serine and arginine residues, MET3 resembles proteins found in splicing or spliceosome associated factors in metazoan cells, which are rich in SR dipeptides and often contain RRM domains (Bickmore and Sutherland, 2002; Birney et al., 1993). Yet, MET3 is clearly distinct from the small number of trypanosomatid SR proteins with putative roles in *trans*-splicing that have been described to date (Ismaili et al., 1999; Manger and Boothroyd, 1998; Portal et al., 2003). MET3 contains nine SR or RS dipeptides and a domain reminiscent of the RRM submotif RNP-1. It is not clear what constitutes a minimal SR domain, but most previously characterised proteins contain more than nine (Birney et al., 1993). Moreover, structural predictions for MET3 are only partially compatible with the requirements for RRM domains. While their sequence conservation is low, the 3D structures of the RRM domains are highly conserved, and arranged in the order $\beta 1$-$\alpha 1$-$\beta 2$-$\beta 3$-$\alpha 2$-$\beta 4$. RNP1 lies in the $\beta 3$ strand, and the conserved aromatic residues protrude from the β-sheet to interact with the RNA (Birney et al., 1993). The RNP-1-like domain in MET3 is predicted to form a β-sheet, consistent with this conserved structure, but the predicted overall 2D structure predictions for MET3 do not fit the RRM domain structure. However, RNP-1 submotifs are also found in some RNA binding protein that are not of the RRM type (Birney et al., 1993).

To determine whether MET3 represents a novel type of RNA (or DNA) binding protein, it will be necessary to investigate its biochemical properties experimentally. As a result of the present study a number of expression constructs and cell lines are available that will facilitate these studies. Preliminary experiments have shown that when *T. cruzi* protein extracts are prepared at physiological salt concentration, a large proportion of MET3 is associated with the insoluble fraction. At high salt concentrations (0.5-1.2M NaCl) MET3 is completely soluble. This is consistent with

176

nucleic acid binding activity. Simple *in vitro* binding assays, using single-stranded DNA or homoribopolymers (poly(A), poly(C), poly(G), poly(U)) could prove directly whether MET3 can bind to nucleic acids. Such approaches have been used to characterise a number of other trypanosomatid RNA-binding proteins, including Nopp44/46 (Das et al., 1996), p34/37 (Zhang and Williams, 1997) and members of the *T. cruzi* RRM type RNA-binding family, including TcUBP1 (De Gaudenzi et al., 2003). Preliminary experiments to test whether recombinant MET3 (expressed in *E. coli*) binds to ssDNA sepharose have so far provided no evidence for binding activity (data not shown). A wider range of conditions will have to be tested however, before definitive conclusions can be drawn. MET3 may be part of a multi-protein complex or require specific modifications in order to perform its function. It may therefore be necessary to use *T. cruzi* extracts rather than protein purified from *E. coli*.

Provided that nucleic acid binding activity is confirmed *in vitro*, the crucial and possibly most difficult step will be to identify the substrate to which MET3 binds in *T. cruzi* cells. A widely used PCR-based approach to identify RNA substrates that bind to a protein of interest is the SELEX method (Systematic Evolution of Ligands by EXponential enrichment) (Tuerk and Gold, 1990). An alternative approach allows the identification of mRNAs associated with a protein, using cellular extracts. This method combines an mRNA co-purification assay with differential display (Rodgers et al., 2002). Once candidate substrates have been identified, there is a variety of affinity-based methods to test the specificity of the interactions in detail, such as gel shift and RNA protection assays and Southwestern blot. To verify whether an interaction occurs in *in vivo*, RT-PCR can be performed on RNAs co-immunoprecipitated with the protein (see for example De Gaudenzi et al., 2003).

To gain further insight into the function of MET3, it will be important to identify proteins that interact with it. There are a number of methods to identify putative MET3 binding proteins, such as the yeast two-hybrid system, or co-immunoprecipitation. Since Western blot analysis suggests the presence of modified forms of MET3 (Figure 4.24), and a number of potential phosphorylation sites in the MET3 sequence were identified we speculate that MET3 may be phosphorylated. Given the nucleolar localisation of MET3 and the presence of four potential casein kinase 2 (CK2) sites

177

in MET3 (Figure 4.6), CK2 is a possible candidate. In many organisms, multiple nucleolar proteins are phosphorylated by CK2. In *T. brucei*, it has been shown that the CK2 catalytic subunit $CK2\alpha$ accumulates in the nucleolus, and phosphorylates the nucleolar RNA binding Nopp44/46 *in vitro* (Park et al., 2002). Interestingly, it has been reported that in mammalian cells CK2 localises to the nucleolus specifically under conditions of stress (Gerber et al., 2000).

4.5.6 Does the *T. cruzi* nucleolus function as a stress-sensor?

In the course of studying the localisation of MET3, we discovered that the dispersal of nucleolar antigens is a feature common to all non-dividing *T. cruzi* cells, including stationary phase epimastigotes. This finding extends previous observations which reported the dispersal of nuclear antigens in trypomastigote nuclei (Elias et al., 2001b). The results from this study provide no evidence that MET3 plays a causal role in the structural changes that occur in the nucleus during differentiation. Deletion of MET3 did not result in any discernible alteration in nuclear morphology, and it did not affect the reversible dissociation of the nucleolar antigen L1C6 (Figure 4.39). Conversely, overexpression of MET3 had no effect on the epimastigote nucleus, as judged by its appearance under the light microscope.

However, this finding raises interesting questions with regard to the signal that triggers metacyclogenesis, and suggests experiments that may shed new light on the elusive mechanism. Alterations in nucleolar structure are a hallmark of cellular stress responses (Welch and Suhan, 1985). Nucleolar fragmentation occurs for instance in heat shocked (Liu et al., 1996), apoptotic (Wyllie et al., 1981) and senescent cells (Sinclair et al., 1997). A recent study has now proposed a convincing model which postulates that the nucleolus itself functions as a cellular 'stress sensor' (Olson, 2004; Rubbi and Milner, 2003). Cells respond to a variety of cellular stresses with the stabilisation of the tumor suppressor p53, which induces cell cycle arrest or apoptosis. The mechanism by which a diverse range of stimuli can trigger p53 stabilisation has long eluded researchers. Rubbi and Millner (2003) showed that the stabilisation of p53 is correlated with nucleolar disruption and nucleolar disruption

alone is able to stabilise p53. In other words, nucleolar disruption is the common mechanism by which cellular stress triggers the protective p53 response.

Based on this concept of a 'cellular stress sensor' it is tempting to speculate that the disassembly of the nucleolus might provide a crucial signal required to trigger metacyclogenesis. A variety of different conditions has been shown to promote metacyclogenesis (section 1.4.2). Yet, despite thorough investigations in a number of laboratories, a signal that is both necessary and sufficient to trigger differentiation has not been defined. However, the common denominator of metacyclogenesis inducing conditions is that they are highly likely to induce cellular stress. The following observations correlate conditions that promote metacyclogenesis with the impairment of nucleolar structure and function. First, in our studies, the dispersal of nucleolar antigens, and the development of metacyclics coincided at the beginning of stationary phase. Second, a previous study has shown that transcription of ribosomal DNA is strongly down-regulated in all non-dividing *T. cruzi* cells, including metacyclics and cell-derived trypomastigotes (Elias et al., 2001b).

In order to gain mechanistic insight into metacyclogenesis, it will be important to determine whether nucleolar disruption occurs under all conditions that induce metacyclogenesis, and to test whether metacyclogenesis can be induced directly by causing nucleolar disruption. There are a number of physical and chemical agents that disrupt the nucleolus in other cells, including heat shock, the anti tumor drug camptothecin (a topoisomerase I inhibitor), 5,6-dichloro-1-β-D-ribofuranosyl-benzimidazole (DRB, a casein kinase II inhibitor) and bleomycin (Rubbi and Milner, 2003). It will be interesting to test these for their effect on *T. cruzi* epimastigotes.

5. THE *MET2* GENE

The *MET2* gene was identified in a screen for RNA transcripts that are upregulated in metacyclics in the *T. cruzi* strain DM28c (Krieger et al., 1999; Yamada-Ogatta et al., 2004). The corresponding protein appears to be associated with the kinetoplast of metacyclics in this strain (Yamada-Ogatta et al., 2004). Its function is unknown, as there is no apparent relatedness to other genes. Our aim was to determine if *MET2* is also a marker for metacyclics in the CL Brener strain, and to use transfection-based approaches to investigate the function of *MET2*. Section 5.1 describes the identification of *MET2* homologues from *T. cruzi* CL Brener, which are located adjacent to the putative centromere of chromosome 3 (Obado et al., in press). The level of *MET2* RNA increases dramatically during metacyclogenesis (section 5.2). *MET2* therefore represents another useful marker in further analysis of *T. cruzi* differentiation in the CL Brener strain. To address *MET2* function, we aimed to create null mutants. This requires deletion of three copies of *MET2* from the CL Brener genome. Section 5.3 describes the successful generation of *MET2* double knockout cell lines. Preliminary results concerning the localisation of the MET2 protein in CL Brener (Section 5.4) are consistent with an association with the kinetoplast.

5.1 CHARACTERISATION OF *MET2* FROM CL BRENER

5.1.1 Identification of two distinct *MET2* alleles

To identify and analyse the CL Brener homologues of *MET2*, we employed essentially the same strategies as for the identification of *MET3* (see chapter 4). BLAST searches of GenBank and GeneDB using the *MET2* cDNA sequence from *T. cruzi* strain Dm28c (metaciclina II, GenBank: AF312552) found two CL Brener cosmids from the *T. cruzi* genome project that contain the *MET2* locus (Figure 5.1). The cosmids were derived from chromosome 3 (Andersson et al., 1998). Sequence divergence between the two cosmids indicated that they were probably derived from two different homologues of chromosome 3. Throughout this work, the *MET2* sequences from CL Brener are referred to as *MET2-CL1o17* and *MET2-CL39p3*, according to

the identifier of the respective cosmid clone. The *MET2* gene from strain Dm28c is referred to as *MET2-Dm28c*. A pair-wise comparison of *MET2-CL1o17*, *MET2-CL39p3* and *MET2-Dm28c* revealed a nucleotide identity of 96% between each pair. This level of heterogeneity between allelic copies was also found in the *MET3* gene (section 4.1.1), and similar values have been reported for a number of other CL Brener genes (Bromley et al., 2004; Machado and Ayala, 2001; Tran et al., 2003). In addition to nucleotide mismatches scattered throughout the *MET2* gene, the inter-genic sequences also contain frequent deletions and insertions (Figure 5.1), including a 0.3 kb insertion in the *MET2-39p3* locus, close to the 5' UTR (Figure 5.3B). These sequence differences allowed the design of allele-specific targeting vectors for gene deletion by homologous recombination (section 5.3). More significantly, we found a frameshift in the *MET2-Dm28c* ORF relative to the ORF common to both CL Brener alleles. The frameshift is caused by the insertion of a thymine at position 666 in *MET2-Dm28c*, which is not present in either of the two CL Brener alleles (Figure 5.1). This was confirmed experimentally by sequencing PCR fragments derived from CL Brener genomic DNA. Downstream of the frameshift, the *MET2-Dm28c* sequence is still highly similar to the CL Brener sequences, including a TAG codon at position 884, which aligns with the predicted stop codon in the CL Brener sequences. There is a strong possibility that the *MET2-Dm28c* sequence in the database contains a sequencing error.

```
1o17                             GACACATTCAGTATCAGTAAGAGAGAATTCGTGTGTTTTCA  -1
39p9      ........................................................  -1
Dm28c     ACAGTTTCTGTACTATATT..................................C.....  60

1o17      ATGATTCCCTGCCGCTCCAACCAACAACAGCATCAAAACATCTCCAAAACGAGGCGAAGC  60
39p9      ATG.......................................T.....G.....  60
Dm28c     ATG.......................................T......G.....  120
PCR       ........................................................

1o17      GCTTATAACTTTACGGACTTTTGCGACGGACGAAATTCATCTGCACCGATGCGTTCTGTC  120
39p9      .A...........A..A...............................G.........  120
Dm28c     .A...............................................A.......  180
PCR       ........................................................

1o17      GTTGCTGGCGACGCTGGAAAAGGAGCAATGTTCACAGCAGCAGAATGGTGATTTGGTTGTC  180
39p9      ............A.................................GC.........  180
Dm28c     ............T.............T.........A.............GC....A..  240
PCR       ........................................................

1o17      ATGACTGATATCATGGAAAGAACAGGTGCATCCACATCTGCCGAACGTGAAACGAATTGC  240
39p9      ........................................................  240
Dm28c     ......G....A.................................T.........  300
PCR       ........................................................

1o17      AGTGAACGTTCTTTGCTAGAGGACAATATAATTCCGGTGTTAAGCATGAGCGAACTTACA  300
39p9      ...........A.........A...........G.............T.....  300
Dm28c     ...........A.....G.........................T.....  360
PCR       ........................................................

1o17      ACAGTGTGTCCACAATTCATCGCTACACCGACCATCGCCACCCCAAATACCCCCAATTGT  360
39p9      ...............G...T...........G.....................T.  360
Dm28c     ...............G...T.................................  420
PCR       ........................................................

1o17      TCTCATTTTAATGCTGCCACTCCTTGTTTGGAGAAATCGCCATTATTATCCTCTTCTTTT  420
39p9      ...T...........................G.........................  420
Dm28c     ...T...................T...........................C.....  480
PCR       ........................................................

1o17      TCTGCTTCGCCGTCAGGATGCGAAGATGAAAAAAAGTCTGGGATATTCTGCAGCTGTCTTT  480
39p9      .A......A...............................................  480
Dm28c     ...........................G............................  540
PCR       ........................................................

1o17      GAATCCGATGGAGAGAAGGATGATGACTGTGATTCCTTGAAAAAGCGCTGACAAAACTAAA  540
39p9      ..............A...............T.........................  540
Dm28c     ..........................T...................C.........  600
PCR       ........................................................

1o17      CTCAAGGCACCTTATCTCATTAACAAATCCCCTGGCTCCATTGCTGCAATTGAATTGGCG  600
39p9      .......................C........T.....................  600
Dm28c     .......................C........T.....................  660
PCR       ........................................................

1o17      TTGTT-GACAAACTCAAAGGAGGATGAGCATACGCAGTCGGATGGGGCATGCTCGATTGA  659
39p9      .....-...G..............................................G  659
Dm28c     .....T.................TGA.................................  720
PCR       .....-...............................................

1o17      CAAGGATTTGGAGTCTGCACTGGAGGAACAAGGGGGTTACGGCACCTGGAACTTGGTGAC  719
39p9      ................................................C.........  719
Dm28c     .......................................T.....C.........  780

1o17      GTCGCGTCTGAAAGCAGTGCATCTTCGTCACATGGCGCACAAGTATGCATTGGAAGTTGT  779
39p9      ........................................................  779
Dm28c     ...................C.........T...........................  840
```

```
1o17    TCGTCAACGCTATGGAGGCAGGATTATAGATGATTCATCTAAGTAGCTTTTAAAGGACAT 839
39p9    .................G............................TAGT............ 839
Dm28c   .............................................TAGT............ 900

1o17    TGTGATTTGTGCTATGGAGGAGGATTCTTCTCCCTTTTCCCTTTTGTTTGAAAAGAAAAG 899
39p9    ..........................................................-- 897
Dm28c   ...........................A...................AAAAAAAAAA 960

1o17    AAAAAAAACACGTGCCTTGAGAG---------TTTTTTTTTTTTTTTTTGCGTGTCTTTTG 950
39p9    .......................GTTGTTTTT.....CC........C............ 957
Dm28c   AAAAAAAAAA                                                   970

1o17    ATTTTTATCTTGGCTGTATGTGAGTGTCTTTCTGTTTTTTTTTTTTTTTTTTT-TTTTTT 1009
39p9    ......G.................T.........G....................GG...... 1017

1o17    GTTTGCGTGTTTGTGCCTTTCCCTCCTTTTTTTTTTTTTTTTTTTTTTTGGGGGAA-ATAT 1068
39p9    ..............G.........GG.......................TT.A...T.... 1077

1o17    ATATTTATATTTAAATATTTTTATATAGTAAAATAAAAAAATATTATTATTATTATTATT 1128
39p9    .............T..................T...T.C.T.................... 1137
```

Figure 5.1: Alignment of MET2 *DNA sequences*

The *MET2* gene was found on two cosmids derived from *T. cruzi* CL Brener chromosome 3: cosmids 1o17 (GenBank:AF052831, Andersson et al., 1998) and 39p3 (GenBank: AC096913). Sequence divergence indicates that they are derived from two different homologues of chromosome 3. This sequence alignment compares the CL Brener genomic sequences with the cDNA sequence of *MET2-Dm28c* (*metaciclina II*, GenBank:AF312552) and a fragment obtained by sequencing a cloned PCR fragment from CL Brener (PCR). The predicted start and stop codons of the *MET2* gene are marked with bold letters. This alignment reveals a frameshift at position 666 in *MET2-Dm28c* (highlighted). The spliced leader sequence at the 5' end, and the putative poly-A tail at the 3' end of the *MET2-Dm28c* cDNA sequence are underlined. Sequencing of a RT-PCR derived cDNA fragment of *MET2-CL1* (section 2.3.3) confirmed the splice acceptor site. Note how the putative poly-A tail in *MET3-Dm28c* aligns with an A-rich stretch in the genomic sequences from CL Brener. A dot (.) denotes nucleotides identical to the *MET2CL-1o17* sequence. Gaps are marked with a dash (-). Two potential regulatory ARE motifs downsstream of *MET2* are highlighted in grey.

5.1.2 The *MET2* gene is located adjacent to the putative
centromere of chromosome 3

Both cosmids spanning the *MET2* genomic locus are derived from chromosome 3. (Andersson et al., 1998). In CL Brener, this chromosome is present in three copies, which differ in size. A smaller homologue of approximately 0.6 Mb is present in two copies. CL Brener cells harbour an additional larger homologue of approximately 1 Mb, which is present in only one copy (Obado et al., in press). The nature of the 400 kb insertion remains to be characterised. The size difference between homologous chromosomes may reflect the fact that CL Brener is a hybrid strain (Machado and Ayala, 2001).

To confirm the chromosomal location of *MET2* experimentally, Southern blots of CL Brener chromosomes separated by pulse field gel electrophoresis, were hybridised with a *MET2* probe. The result shows that both the 0.6 Mb and 1 Mb homologues of chromosome 3 contain the *MET2* gene and that no closely related sequences are found on other chromosomes (Figure 5.2). In *T. cruzi* Dm28c, *MET2* hybridised only to a 0.65 Mb chromosome (Yamada-Ogatta et al., 2004). Southern blot analysis of genomic DNA, probed with the *MET2* coding sequence, shows that *MET2* is a 'single copy gene', with three allelic copies in CL Brener (Figure 5.3A). The *MET2* probe detects two bands in genomic DNA digested with restriction enzymes that do not cut within the *MET2* coding sequence. A fainter band corresponds to the *MET2* gene on the 1 Mb chromosome, and a stronger band corresponds to the *MET2* gene on the 0.6 Mb chromosome, which is present in two copies. This interpretation is consistent with the restriction maps derived from genomic sequence information (Figure 5.3B and data not shown). It is also supported by analysis of DNA from single knockout mutants (section 5.3), where deletion of *MET2* from the 1 Mb chromosome results in the disappearance of the weaker band from genomic Southern blots (Figure 5.9).
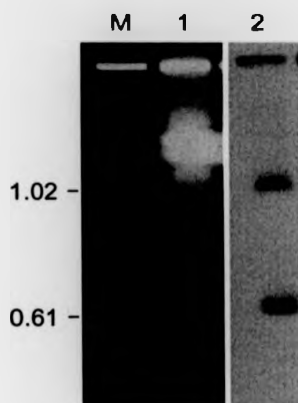
Figure 5.2: Localisation of MET2 *on chromosome 3 (CHEFE)*

Chromosomes of *T. cruzi* CL Brener were separated by CHEFE (lane 1). *Saccharomyces cerevisiae* chromosomes were run as size markers (lane M). A blot of the gel, probed with the *MET2* coding sequence. This shows two bands that correspond to the 0.6 Mb and 1 Mb homologues of chromosome 3 (lane 2). Numbers on the left indicate size in megabases (Mb).

Figure 5.3 (overleaf): Restriction map of the MET2 *locus*

Genomic DNA from *T. cruzi* CL Brener was digested with restriction enzymes as indicated above each lane. (**A**) Southern blot probed with *MET2*. The banding pattern is consistent with a single copy gene with three alleles. The fainter band is derived from *MET2-CL39p3* (present in one copy), the stronger band from *MET2-CL1o17* (two copies). Numbers on the left indicate size in kb. (**B**) Restriction map of the *MET2-CL1o17* locus, derived from genomic sequence information and confirmed by the Southern blot. Numbers indicate nucleotide positions. The first nucleotide in the *MET2* start codon was designated position '1'. The site of a 0.3 kb insertion in the *MET2-CL39p3* locus is indicated. An *Xho* I site contained within this insertion is marked (X).

Figure 5.3

A

10		Bam HI
8		Eco RI
		Eco RV
6		Hind IIII
5		Nco I
4		Sac II
		Sal I
3		Spe I
2.5		Xba I
2		Xho I
1.5		
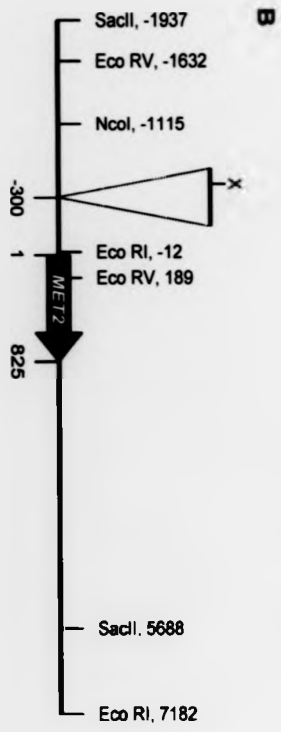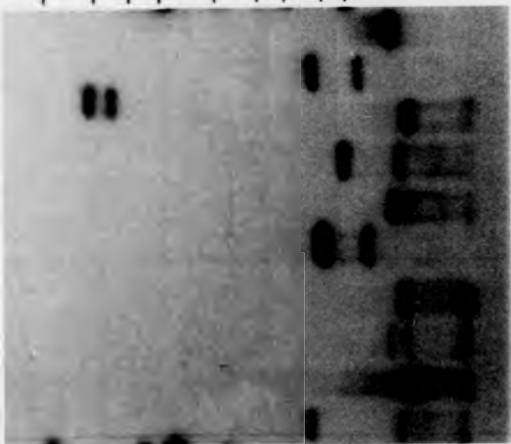
Analysis of a 93.4 kb genomic contig that spans the *MET2* locus (Andersson et al., 1998) revealed an interesting location of the *MET2* gene (Figure 5.4), in close proximity to the putative centromere (Obado et al., in press). The arrangement of genes in large polycistronic transcription units is a peculiar feature of kinetoplastid genomes (section 1.6.1), which may be related to the mechanisms by which gene expression is controlled. The coding sequences contained within the 93.4 kb contig are arranged in two directional clusters, which are transcribed in opposite directions (Andersson et al., 1998). The point at which the transcription units diverge is referred to as the strand switch region. Interestingly, the *MET2* gene lies very close to this strand switch (Figure 5.4). *MET2* is the second gene in the cluster on the shorter arm of the chromosome, and the *MET2* ORF is flanked on either side by long intergenic regions of low sequence complexity. The intergenic sequence downstream of *MET2* is characterised by repeats consisting mainly of A and T. These include single nucleotide runs (mainly T), runs of AT dinucleotides, and repeated units of patterns of up to 8 nucleotides were identified using the program Repeat Finder (Benson, 1999).

The strand switch region (approx 20 kb) is made up largely of degenerate retro-elements and the DNA sequence is characterised by a high GC content (Andersson et al., 1998; Ghedin et al., 2004). This region has recently been shown to comprise a putative centromere, i.e. the element that confers mitotic stability (Obado et al., in press).
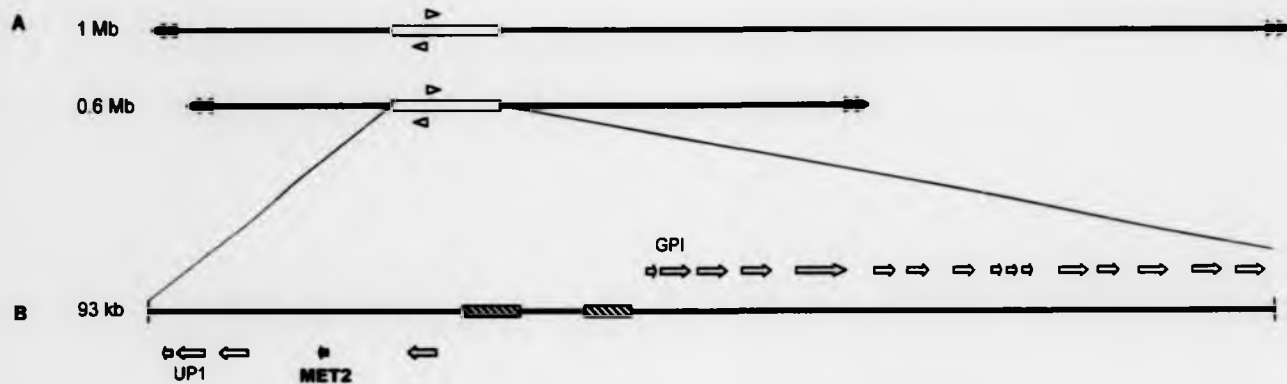
Figure 5.4: Map of chromosome 3

(A) Schematic representation of chromosome 3 from *T. cruzi* CL Brener. Chromosome 3 is trisomic in this strain. A smaller homologue of 0.6 Mb is present in two copies, and a larger homologue of 1 Mb is present in one copy. Black double arrows represent telomeric repeats. Open arrows indicate the position of the strand switch, and show the direction of transcription. The 16 kb strand switch region is composed predominantly of degenerate retroelements (hatched boxes) belonging to the VIPER/SIRE and LITc families (Obado et al., in press). (B) Enlarged view of the strand switch region of chromosome 3 which contains the putative centromere (Obado et al., in press). Coding sequences (blue arrows) cluster on the same strand, and the strand switch separates two divergent clusters. *MET2* is the second gene on the short arm of chromosome 3. The *MET2* ORF is flanked on either side by approximately 5 kb of intergenic sequence. The chromosome 3 specific genes *GPI* and *UP1* were used as probes in Southern blots.

5.1.3 *MET2* is only found in *T. cruzi*

The gene order in trypanosomatids is highly conserved (Bringaud et al., 1998; Ghedin et al., 2004). The first large-scale comparative study of kinetoplastid chromosomes found a high level of synteny between *T. cruzi*, *T. brucei* and *L. major* chromosomes, and identified chromosomal segments homologous to the *T. cruzi* chromosome 3 strand switch region on *L. major* chromosome 12 and *T. brucei* chromosome 1 (Ghedin et al., 2004). Despite this high conservation of genes between kinetoplastid species, we found no sequences related to *MET2* in the syntenic loci of *T. brucei* and *L. major*. *MET2* is located on the shorter arm of *T. cruzi* chromosome 3 (to the left of the strand switch, see Figure 5.4), where more divergence was observed between the species than on the longer arm (Ghedin et al., 2004). Searches of kinetoplastid genomes in GeneDB have identified no sequences that appear to be related to *MET2*, other than the CL Brener alleles. Similarly, BLAST searches of GenBank have identified no sequences similar to *MET2* in any other organisms. *MET2* is thus another *T. cruzi*-specific gene.

5.1.4 MET2 protein sequence analysis

The derived sequence of the CL Brener MET2 protein is 274 amino acids long, with a calculated mass of 29.8 kDa, and a pI of 5.13. The protein sequence derived from *MET2-Dm28c* (GenBank: AAK37448) contains only 208 amino acids. When the reading frame of the *MET2-Dm28c* gene is shifted at position 666 (where this sequence contains an additional thymine not present in CL Brener; Figure 5.1), translation of nucleotides 677 to 886 results in a protein sequence nearly identical to the predicted CL Brener MET2 protein (Figure 5.5). This is a further indication that the frameshift in the MET2-Dm28c cDNA sequence may be due to a sequencing error.

Bioinformatic analysis of the MET2 amino acid sequence shows that the protein is rich in serines (12.8%). There is a strikingly regular spacing of cysteines (H2N-3-C-23-C-21-C-29-C-22-C-16-C-8-C-17-C-22-C-46-C-57-COOH). These are predicted to form five cysteine bonds (Vullo and Frasconi, 2004). Since cysteine bonds are usually found in extracellular and membrane proteins, and MET2 contains no high-

189

scoring hydrophobic or transmembrane segments, the relevance of this prediction is uncertain. A Prosite scan identified only patterns that occur frequently in any protein, and no conserved motifs. PSORT II predicts mitochondrial or nuclear localisation with similar probabilities. A cleavage site for mitochondrial pre-sequence is found at position 29 in MET2-CL1o17 (RRS|AY) and at position 28 in MET2-CL39p3 and MET2-Dm28c (TRG|SD). Mitochondrial localisation is also predicted by the program TargetP (Emanuelsson et al., 2000). MET2 contains no predicted nuclear localisation signals. In short, sequence analysis supports a possible mitochondrial localisation of MET2. The likely biological function of MET2 cannot be easily inferred from sequence analysis.

```
CL1o17     MIPCRSNQQQHQNISKTRRSAYNFTDFCDGRNSSAPMRSVVAGDAGKEQCSQQQNGDLVV  60
CL39p3     ...............N..G.D....E..........V.......T.............A..  60
Dm28c      ...............N..G.D................S.............F.......A.I 60

CL1o17     MTDIMERTGASTSAERETNCSERSLLEDNIIPVLSMSELTTVCPQFIATPTIATPNTPNC 120
CL39p3     ...........................K..K........F.......VV.....A.....F 120
Dm28c      ...............D...........K.S.........F.......VV............ 120

CL1o17     SHFNAATPCLEKSPLLSSSFSASPSGCEDEKSLGYSAAVFESDGEKDDDCDSLKSADKTK 180
CL39p3     .Y......C..R........Y.............................V......... 180
Dm28c      .Y......F.........P..............................V......G... 180

CL1o17     LKAPYLINKSPGSIAAIELALLTNSKEDEHTQSDGACSIDKDLESALEEQGGYGTWNLVT 240
CL39p3     ........T.............................G................H... 240
Dm28c      ........T............FDKL.GG*                                208
Dm28c(2)                     ............................S.H... 240

CL1o17     SRLKAVHLRHMAHKYALEVVRQRYGGRIIDDSSK* 274
CL39p3     .......................D........* 274
Dm28c(2)   ...............................* 274
```

Figure 5.5: Alignment of predicted MET2 protein sequences

CL1o17 and CL39p3 are translations of two distinct MET2 alleles from *T. cruzi* CL Brener. Dm28c corresponds to GenBank entry AAK37448 (translation of nucleotides 1-687 of the MET2-Dm28c cDNA, reading frame 1). Dm28c (2) is the translation of nucleotides 667-886 of the MET2-Dm28c cDNA (reading frame 2). This alignment suggests that the truncation of MET2-Dm28c relative to the CL Brener sequences could be due to a simple sequencing error. A dot (.) denotes amino acids identical to the CL-1o17 sequence. A star (*) denotes a stop codon.

scoring hydrophobic or transmembrane segments, the relevance of this prediction is uncertain. A Prosite scan identified only patterns that occur frequently in any protein, and no conserved motifs. PSORT II predicts mitochondrial or nuclear localisation with similar probabilities. A cleavage site for mitochondrial pre-sequence is found at position 29 in MET2-CL1o17 (RRS|AY) and at position 28 in MET2-CL39p3 and MET2-Dm28c (TRG|SD). Mitochondrial localisation is also predicted by the program TargetP (Emanuelsson et al., 2000). MET2 contains no predicted nuclear localisation signals. In short, sequence analysis supports a possible mitochondrial localisation of MET2. The likely biological function of MET2 cannot be easily inferred from sequence analysis.

```
CL1o17    MIPCRSNQQQHQNISKTRRSAYNFTDFCDGRNSSAPMRSVVAGDAGKEQCSQQQNGDLVV  60
CL39p3    ................N..G.D....E..........V.......T............A..  60
Dm28c     ................N..G.D................S..........F.......A.I  60

CL1o17    MTDIMERTGASTSAERETNCSERSLLEDNIIPVLSMSELTTVCPQFIATPTIATPNTPNC  120
CL39p3    .........................K..K........F.......VV.....A.....F  120
Dm28c     .............D...........K.S.........F.......VV............  120

CL1o17    SHFNAATPCLEKSPLLSSSFSASPSGCEDEKSLGYSAAVFESDGEKDDDCDSLKSADKTK  180
CL39p3    .Y......C..R........Y.................................V.....  180
Dm28c     .Y......F........P...................................V.....G...  180

CL1o17    LKAPYLINKSPGSIAAIELALLTNSKEDEHTQSDGACSIDKDLESALEEQGGYGTWNLVT  240
CL39p3    ........T.................................G................H...  240
Dm28c     ........T............FDKL.GG*                                208
Dm28c(2)              ...........................S.H...  240

CL1o17    SRLKAVHLRHMAHKYALEVVRQRYGGRIIDDSSK*  274
CL39p3    .........................D........*  274
Dm28c(2)  ...............................*  274
```

Figure 5.5: Alignment of predicted MET2 protein sequences

CL1o17 and CL39p3 are translations of two distinct MET2 alleles from *T. cruzi* CL Brener. Dm28c corresponds to GenBank entry AAK37448 (translation of nucleotides 1-687 of the MET2-Dm28c cDNA, reading frame 1). Dm28c (2) is the translation of nucleotides 667-886 of the MET2-Dm28c cDNA (reading frame 2). This alignment suggests that the truncation of MET2-Dm28c relative to the Cl. Brener sequences could be due to a simple sequencing error. A dot (.) denotes amino acids identical to the CL-1o17 sequence. A star (*) denotes a stop codon.

5.2 *MET2* RNA TRANSCRIPTS ARE UPREGULATED DURING METACYCLOGENESIS

MET2 was initially identified in a screen for RNA transcripts that are enriched in metacyclics of *T. cruzi* strain Dm28c (Yamada-Ogatta et al., 2004). We established by Northern blot analysis that *MET2* is also upregulated in CL Brener cells in stationary phase (Figure 5.6). The stage-specificity of *MET2* RNA transcripts is rather striking. A band of approximately 5 kb hybridises strongly with RNA samples from cultures containing 24% metacyclics but only very weakly with those containing 6% metacyclics, and no signal is detected in epimastigote RNA. This resembles closely the pattern observed in strain Dm28c, where *MET2* RNA detected exclusively in fully developed metacyclics (Yamada-Ogatta et al., 2004). The *MET2* RNA transcript thus represents a very specific marker for *T. cruzi* metacyclics.

Interestingly, the estimated size of the *MET2* transcript detected in Northern blots of CL Brener RNA is considerably larger than the size of the *MET2-Dm28c* cDNA deposited in GenBank, which is only 970 bp. The *MET2-Dm28c* cDNA from the database contains both the spliced leader sequence at the 5' end and a poly A tail at the 3' end, suggesting that it represents the entire mRNA sequence. However, when this cDNA sequence is aligned with the CL Brener genomic sequences spanning the *MET2* locus, the putative poly-A tail of MET2-Dm28c aligns with an A-rich stretch of genomic sequence (Figure 5.1). The method used to clone the *MET2* cDNA from the Dm28c strain involved synthesis of the first cDNA strand using an oligo-dT$_{15}$ primer (Krieger and Goldenberg, 1998). Since this primer will bind to stretches of A nucleotides as present in the 3'UTR of *MET2*, it is very likely that the reported *MET2-Dm28c* 'cDNA' sequence represents a truncated fragment.

The factors regulating the stage-specificity of *MET2* RNA levels remain to be determined. It is interesting to note that the derived 3' UTR of MET2-1o17 contains several AU rich sequences in its 3'UTR. In the intergenic sequence downstream of the MET2 stop codon, a total of nine AU-rich sequences were found that correspond either to the AUUUA motif or the UUAUUUA(U/A)(U/A) motif (Chen and Shyu, 1995). The two AUUUA motifs closest to the stop codon are highlighted in Figure

5.1. AU-rich elements (ARE) are *cis-* acting elements in the 3'UTR that regulate gene expression in a number of eukaryotes, including *T. cruzi* (Chen and Shyu, 1995; Di Noia et al., 2000). Whether they are involved in the stabilisation of the *MET2* transcript is an open question.
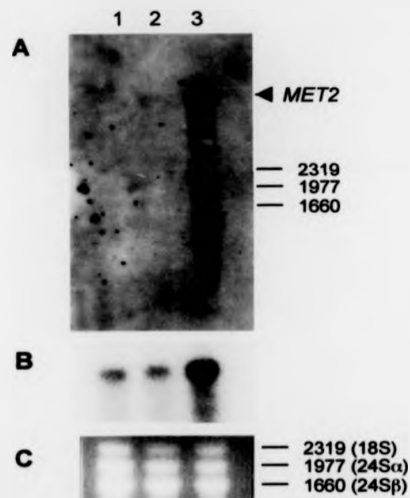
Figure 5.6: Stage-specific expression of MET2 *RNA*

Northern blot analysis of RNA extracted from cultures containing different proportions of meta-cyclics. Lane 1: log-phase epimastigotes; lane 2: 6% metacyclics; lane 3: 24% metacyclics. (A) A transcript of approximately 5 kb is detected with a *MET2* probe in samples containing 24% meta-cyclics. A faint 5 kb band is also seen in the sample containing 6% metacyclics. No *MET2* transcript is detected in epimastigote RNA. (B) *gp82*, a gene known to be upregulated in metacyclics. (C) RNA loading was controlled by ethidium bromide staining of the ribosomal RNA bands (18S, 24Sα and 24Sβ) prior to blotting. 10μg RNA was loaded per lane. Numbers on the right indicate size in bases.

5.3 GENERATION OF *MET2* DOUBLE KNOCKOUT CELL LINES

To address the function of *MET2* in metacyclics, we aimed to generate *MET2* null mutants. The only available method to generate a loss of function mutant of an un-characterised gene in *T. cruzi* is by gene deletion, which is a very slow process in this organism. Since *MET2* is not expressed in epimastigotes, it seems unlikely that this gene will be required for cell survival and replication in the epimastigote stage. It was therefore deemed feasible to delete all three alleles from the genome. This section describes the successful generation of several double knockout clones. In these cells, two copies of *MET2* were deleted from the genome of *T. cruzi* CL Brener by successive replacement with drug selectable marker genes (Figures 2.2 and 5.7). The sequences flanking the *MET2* ORF are divergent between the different alleles, and each construct was designed to target a specific allele.
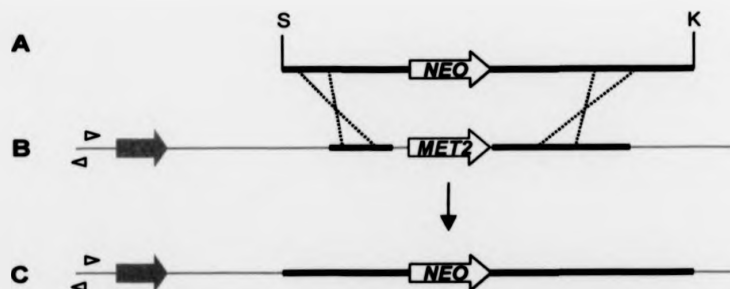
Figure 5.7: Strategy used to delete the MET2 *gene*
(A) The selectable marker gene (e.g. *NEO*) is flanked by *GAPDH* intergenic sequences (black lines) containing the splice acceptor site and polyadenylation signals for processing of the RNA, and by sequences homologous to the *MET2* locus (blue lines). The targeting construct is a linear fragment obtained by restriction digest of pkoMET2(39p3)-NEO with *Sac* I (S) and *Kpn* I (K). (B) The endo-genous *MET2* locus. The open arrows on the left indicate the location of the strand switch and show the direction of transcription. Crosses symbolise the crossover event. (C) Integration by homologous recombination has replaced the *MET2* gene with the targeting construct. Cells where integration of *NEO* has occurred are selected using the drug G418.

Figure 5.8: Generation of* MET2 *single knockout clone B

Southern blot of *T. cruzi* CL Brener chromosomes, separated by CHEFE. WT: wild-type. Lanes A and B: G418-resistant *T. cruzi* clones A and B, derived from transfection with pkoMET2-NEO. The blot was hybridised successively with probes to genes present on chromosome 3: *MET2*, *GPI*, and *UP1*, and with a *NEO* probe, as indicated. The expected 1 Mb and 0.6 Mb bands are detected with chromosome 3 specific genes in WT. In clone B (sko-B), the *MET2* probe fails to hybridise with the 1 Mb chromosome (*), indicating that integration of the *NEO* gene has resulted in deletion of one copy of *MET2*. The *UP1* and *GPI* probes detect no difference between WT and sko-B. Clone A shows an unexpected chromosome pattern. Failure to detect the 1 Mb band with any of the chromosome 3 specific genes (#) indicates that the 1 Mb chromosome is missing. A strong signal in the well is detected with *MET2* and *UP1*, but not with *GPI*. This suggests that the missing 1 Mb chromosome may have undergone a rearrangement, and the resulting entity shows altered mobility in pulse field gels reminiscent of circular DNA fragments.
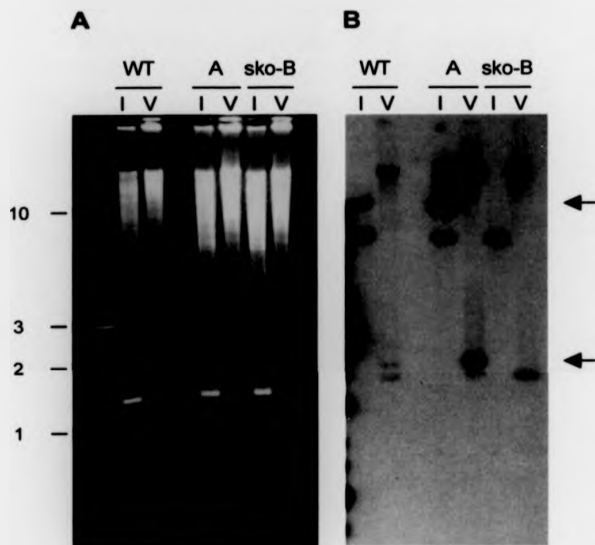
Figure 5.9: Genomic Southern blot of MET2 *single knockout clone*

(A) Ethidium bromide stained genomic DNA of wild-type *T. cruzi* CL Brener (WT), and two clones derived from transfection with pkoMET2-NEO (A and sko-B). The DNA was digested with *Eco* RI (I) or *Eco* RV (V) as indicated above each lane, and resolved on a 0.8% agarose gel. (B) Southern blot probed with *MET2*. The probe detects two fragments in WT DNA that are absent in sko-B (>10 kb *Eco* RI, and 2.1 kb *Eco* RV; arrows). This demonstrates that in sko-B the *MET2-39p3* allele has been deleted. In clone A, deletion of *MET2* has not occurred. An increase in signal intensity in the >10 kb *Eco* RI, and 2.1 kb *Eco* RV fragment indicates that in these cells the copy number of the *MET2-39p3* allele has increased. Numbers on the left indicate size in kb.
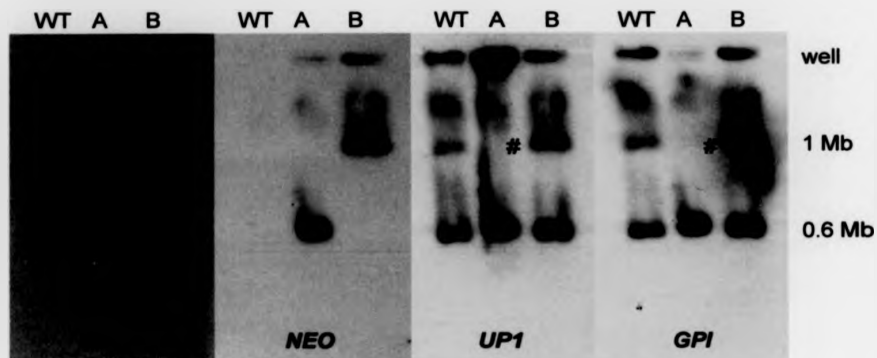
Figure 5.10: Generation of MET2 *single knockout clones C and D*

Southern blot of *T. cruzi* CL Brener chromosomes, separated by CHEFE. WT: wild-type. Lanes C and D: *HYG*-resistant *T. cruzi* clones C and D, derived from transfection with pkoMET2-HYG. The *HYG* probe hybridises to the 0.6 Mb chromosome in clones C and D, suggesting correct integration of the targeting construct. The *MET2* probe hybridises with the two remaining gene copies on the 1 Mb and 0.6 Mb chromosome.

5.3.1 Deletion of the first *MET2* allele

After the first round of transfection, drug resistant clones were analysed for integration of the targeting construct into the *MET2* locus. Chromosome Southern blots showed that in three clones integration of the drug selectable marker gene into the expected chromosome 3 homologue had occurred, resulting in the deletion of one copy of *MET2*. These single knockout clones were named sko-B, sko-C and sko-D. In sko-B (Figure 5.8), a *NEO* probe hybridises with the 1 Mb chromosome. The *MET2* probe only hybridises to the 0.6 Mb homologue, indicating that the allele has been deleted from the 1 Mb homologue. This was confirmed by a genomic Southern blot of sko-B (Figure 5.9), where one of the *MET2* bands found in wild-type cells had disappeared. In sko-C and sko-D, a *HYG* probe hybridises with the 0.6 Mb chromosome (Figure 5.10). Because the 0.6 Mb homologue is present in two copies, the *MET2* probe still hybridises with both chromosome 3 homologues.

Two cell lines (NEO resistant clone A and HYG resistant ko-HYG1) showed unexpected alterations in chromosome numbers. This involved the disappearance of the 1 Mb band from chromosome Southern blots, and evidence of a circular amplicon derived from this chromosome, as discussed in section 5.4

5.3.2 Deletion of a second *MET2* allele

The single knockout cells were subsequently transfected with pko-MET2-PAC. Four of the resulting double drug resistant cell lines were analysed by probing chromosome Southern blots with *PAC* and *MET2* probes (Figure 5.11). The *PAC* probe hybridises to the 0.6 Mb homologue, as expected, in all four clones. Hybridisation with the *MET2* probe shows that only one copy of the gene remains in each clone. Thus, four *MET2* double knockout cell lines were obtained: those derived from sko-B were named dko-BC and dko-BE; those derived from sko-C were named dko-CA and dko-CD. A third round of transfection was done to delete the remaining copy of *MET2*. These cells are currently undergoing selection. Once the null mutants are available, their ability to differentiate to metacyclics will be tested.
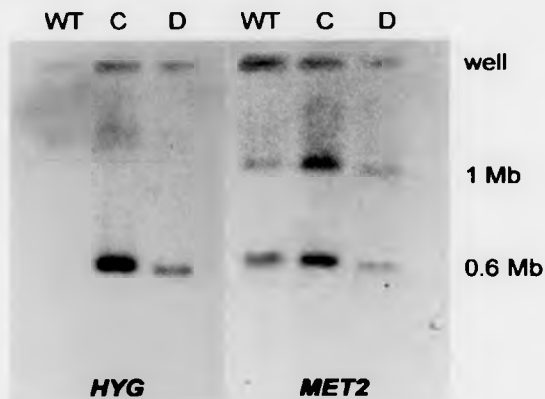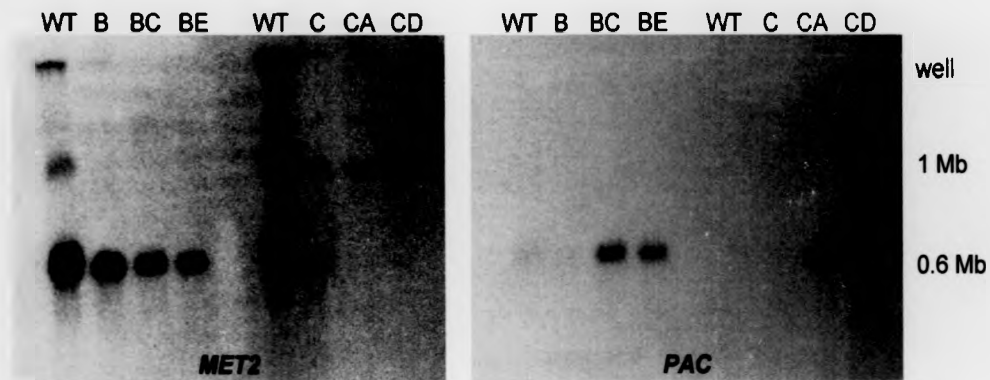
Figure 5.11: Generation of MET2 double knockout clones

Southern blot of *T. cruzi* CL Brener chromosomes, separated by CHEFE. WT: wild-type. Lanes B and C: single knockout clones. Lanes BC, BE, CA and CD: double knockout clones. Hybridisation of the blot with *PAC* demonstrates integration of the targeting construct into the 0.6 Mb chromosome. Hybridisation of the blot with *MET2* confirms deletion of a second *MET2* allele in the double knockout clones. In dko-BC and dko-BE, one of the *MET2* alleles on the 0.6 Mb homologue remains. In dko-CA and dko-CD, the *MET2* allele on the 1 Mb homologue remains.

5.4 TARGETING THE *MET2* LOCUS PRODUCED INTERESTING CHROMOSOMAL ABNORMALITIES

Targeting of pkoMET2-NEO to the *MET2* locus on the 1 Mb homologue yielded one drug resistant cell line (clone A), which displayed unexpected alterations to chromosome 3. In Southern blot of chromosomes from clone A (Figure 5.9), the *NEO* probe hybridises to a chromosome of 0.6 Mb, and not as expected to one of 1 Mb. More unexpectedly, three probes specific to chromosome 3 (*MET2*, *UP1* and *GP1*) all fail to detect the 1 Mb chromosome in this clone. In addition, to the expected signal from the 0.6 Mb homologue, *MET2* and *UP1* (but not *GP1*) hybridise strongly with the well. The strong hybridisation signal in the well is seen neither in the wild type controls, nor in sko-B where the *NEO* gene has correctly integrated into the 1 Mb chromosome (Figure 5.9). In clone A, the signal in the well may be due to the presence of a circular amplicon, that has arisen from parts of the 1 Mb homologue. The suggestion that the 1 Mb chromosome has undergone a rearrangement, and has not been entirely lost from the cells is supported by Southern blot analysis of clone A (Figure 5.10). A *MET2* probe detects the same restriction fragments in a Southern blot of clone A genomic DNA and wild type DNA. However, in clone A, the restriction fragment derived from the *MET2* gene on the 1 Mb chromosome shows a strong increase in signal intensity. Taken together, the data from clone A suggest the following: the 1 Mb chromosome has undergone a change which (i) alters its mobility in pulse field gels and (ii) has led to an increase in the copy number of the *MET2* gene.

Disappearance of the 1 Mb homologue of chromosome 3 was also observed in one cell line transfected with the construct pkoMET2-HYG, which was designed to integrate into the 0.6 Mb chromosome (data not shown). More significantly, in experiments unrelated to *MET2*, other manipulations of the centromeric region of chromosome 3 have resulted in clones with a chromosome pattern very similar to clone A (Sam Obado and John Kelly, unpublished data). Experiments are underway to verify whether exactly the same rearrangement has occurred in completely independent experiments, and to determine the nature of the putative amplicon.

5.5 EXPRESSION OF RECOMBINANT MET2

In order to investigate expression and localisation of MET2 protein, two strategies were pursued: (i) using recombinant MET2 purified from a bacterial expression system to raise specific antisera in mice, and (ii) expression of epitope tagged MET2 in *T. cruzi*.

5.5.1 Expression of recombinant MET2 in *E. coli*

A bacterial expression system was chosen for production of recombinant MET2. Expression vector pTrcHis-met2 was generated by cloning a fragment encoding amino acids 2-213 of the MET2 coding sequence into vector pTrcHis C to produce pTrcHis-met2 (section 2.3.4). *E. coli* BL21 were transformed with pTrcHis-met2 and expression of the recombinant protein was induced with IPTG, as described in section 2.4.2. Western blots of *E. coli* lysates were screened for expression of the 27 kDa recombinant proteins. Recombinant MET2 was expressed at high levels 2 hours after induction, and found predominantly in the insoluble fraction. Expression of the protein at 18°C or 4°C instead of 37°C did not yield higher proportions of soluble protein (data not shown). After addition of 1% Triton X-100 to the lysis buffer, a small part of protein was found in the soluble fractions. We were thus able to purify soluble MET2 under native conditions on a Ni-NTA agarose column (Figure 5.12) as described in section 2.4.3.

5.5.2 Antisera from mice immunised with recombinant MET2
fail to detect the protein

Attempts to raise MET2 antibodies (section 2.4.4) yielded negative results. Antisera from three mice did not detect the 27 kDa recombinant protein in Western blots (Figure 5.13A). A cross-reacting band of 50 kDa present in Western blots was found to be non-specific, as it was also detected by the pre-immune sera (Figure 5.13B). An additional 75 kDa band detected with serum from mouse 2 only, also appears to be unrelated to MET2. The antibody against the Xpress epitope tag detects low levels of higher molecular weight bands, which could represent multimeric forms of MET2.

However, these bands run slightly lower than the 75 kDa band detected with serum from mouse 2. None of the mouse sera detected any proteins in *T. cruzi* extracts that migrate at the size of MET2, or that show any stage-specificity (not shown).
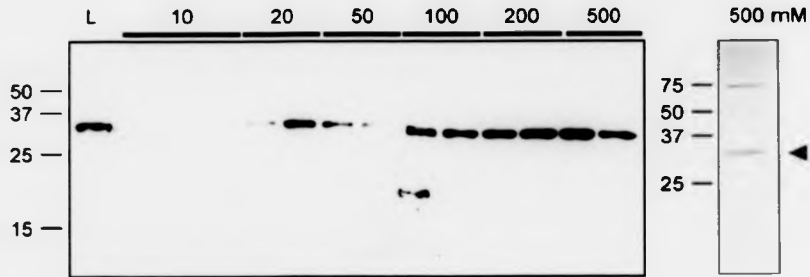
Figure 5.12: Expression and purification of recombinant MET2

A fragment of *MET2* (coding for amino acids 2-213) was cloned into the expression vector pTrcHis, to yield a recombinant protein 6xHis-MET2 fused in-frame with an N-terminal 6x His tag and an Xpress epitope tag (6xHis-MET2). Recombinant protein was expressed in *E. coli* BL21, and purified under native conditions using a Ni-NTA agarose column. Left panel: Western blot of eluted protein probed with the anti-Xpress antibody. The protein was eluted from the column using increasing concentrations of imidazole, as indicated above each lane. L: crude *E. coli* extract. Right panel: Coomassie-stained 6xHis-MET2 (arrow), eluted with 500 mM imidazole. Numbers on the left indicate molecular size in kilo Daltons.
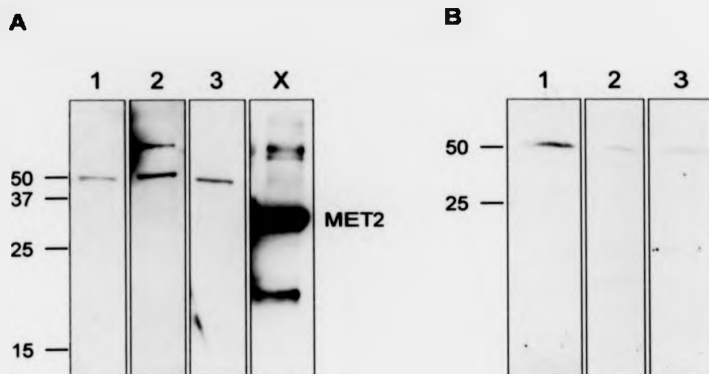
Figure 5.13: Immunisation of mice yields no MET2 specific antiserum

Purified 6xHis-MET2 was used to immunise three mice. Serum from these mice was tested on Western blots of the purified protein (*E. coli* lysates fractionated on Ni-NTA columns). (**A**) Lanes 1-3: sera from three different mice fail to detect purified 6xHis-MET2 on a Western blot. An unidentified 50 kDa band is detected with all sera. A band of approximately 75 kDa is only detected with serum from mouse 2. These unidentified bands could represent *E. coli* proteins that were eluted together with 6xHis-MET2 from the Ni-NTA column. Lane X: The anti-Xpress antibody detects 6xHis-MET2 in these samples (positive control). (**B**) Pre-immune sera from the same three mice also detect a 50 kDa band in Western blots of recombinant 6xHis-MET2. This shows that these bands are not related to MET2. Numbers on the left indicate molecular sizes in kilo Daltons.

5.5.3 Localisation of epitope tagged MET2 *T. cruzi* epimastigotes

In order to determine the subcellular location of MET2 protein in *T. cruzi* without the aid of specific antibodies, we pursued an epitope tagging strategy. Episomal expression vector pTEX-MET2-9E10 was constructed for this purpose (section 2.3.4; Figure 2.1), and transfected into *T. cruzi* epimastigotes. Cells resistant to G418 were readily obtained. Southern blot analysis confirmed that these cells harbour intact pTEX-MET2-9E10 plasmids (not shown). Paraformaldehyde fixed cells were stained with a monoclonal anti c-myc antibody for immunofluorescence detection of MET2-9E10. Only a small subset of cells stained positive with the antibody. The signal in

these cells is concentrated adjacent to, or overlapping with the kinetoplast (Figure 5.14). Unfortunately, none of the metacyclics that developed in these cultures stained positive for c-myc. The c-myc antibody failed to detect MET2-9E10 in Western blots (not shown). A likely explanation for this is that due to the small proportion of cells that expressed the protein, the amount present in the sample remained below the detection limit.

These results show that MET2-9E10 associates with the kinetoplast (or structures adjacent to the kinetoplast), when expressed in epimastigotes. The MET2-Dm28c protein was found associated with the metacyclic kinetoplast in electron microscopy studies (Yamada-Ogatta et al., 2004). Our results are consistent with this. A possible reason for the low levels of MET2-9E10 expression in transfected epimastigotes could be that MET2 expression is incompatible with cell proliferation. Alternatively, it may be that the C-terminus, including the c-myc tag is removed by processing. Native MET2 in Dm28c detected on a Western blot is only 14 kDa. Future, localisation studies will therefore utilise an inducible expression system (Figure 1.5), which has only recently become available (Martin Taylor, unpublished). This will also allow us to determine the effect of MET2 overexpression in a controlled manner.

Figure 5.14: Localisation of MET2-9E10 in epimastigotes

T. cruzi epimastigotes transfected with pTEX-MET2-9E10 were fixed and labelled with anti-c-myc antibody to detect recombinant MET2-9E10. A signal adjacent to the kinetoplast is detected in a small subset of cells. Kinetoplast (K) and nuclear DNA (N) is stained with TOTO-3. Scale bar = 2 μm.

5.6 DISCUSSION

5.6.1 *MET2* RNA transcripts are upregulated during metacyclogenesis

The results of this study show that expression of *MET2* in *T. cruzi* CL Brener is stage-specific, and follows a similar pattern to that of the *T. cruzi* Dm28c homologue (Yamada-Ogatta et al., 2004). Northern blot analysis showed that the level of MET2 RNA transcripts was strongly upregulated in stationary cultures of *T. cruzi* CL Brener, and that it correlates with a high number of metacyclics in the sample. Interestingly, the up-regulation of *MET2* transcripts, was much more pronounced than that of *MET3* (chapter 4). While low levels of *MET3* RNA were readily detected in Northern blots of epimastigote RNA, neither CL Brener (Figure 5.6) nor Dm28c epimastigotes (Yamada-Ogatta et al., 2004), contained detectable levels of the *MET2* transcript.

We could not test directly whether the *MET2* transcripts in RNA from stationary phase cells were derived exclusively from metacyclics. This is because attempts to separate CL Brener metacyclics from stationary epimastigotes and intermediate forms failed to provide pure samples of either cell type (see section 3.3.4). However, in *T. cruzi* Dm28c, the *MET2* transcript was only detected in RNA from metacyclics, but not in differentiating epimastigotes (Yamada-Ogatta et al., 2004). It is therefore likely that the *MET2* RNA transcript detected in the CL Brener samples is also derived from metacyclics. As yet, the levels of *MET2* transcripts in amastigotes and trypomastigotes have not been analysed in either *T. cruzi* strain. It will be straight forward to close this gap by hybridising Northern blots of amastigote and trypomastigote RNA with a *MET2* probe.

Based on the results of the Northern blot analysis, the 3'UTR of *MET2* from CL Brener is estimated to extend over 4 kb. We would speculate that elements that control stage-specificity of *MET2* expression are contained within this long 3'UTR. Examination of the sequence downstream of the *MET2* stop codon identified nine AU-rich elements (ARE) similar to those that have been shown to control RNA stability in a number of eukaryotes including *T. cruzi* (Di Noia et al., 2000). ARE-

binding proteins which recognise AUUUA repeats in cytokine RNAs can exert both stabilising and destabilising effects (Chen et al., 2002). *T. cruzi* contains at least one developmentally regulated RNA binding protein (TcUBP-1) that regulates expression of ARE containing genes. TcUBI-1 recognises the AU-rich instability element in the *T. cruzi* small mucin gene family *TcSMUG*, and overexpression of TcUBP-1 decreases the half life of *TcSMUG* derived transcripts (D'Orso and Frasch, 2001b). However, in *TcSMUG*, the AU-rich element results in destabilisation and decreased translation efficiency of the transcript in metacyclics (Di Noia et al., 2000). This is the opposite of the metacyclic stage-specific up-regulation of the *MET2* transcript. To test whether the AU-rich sequences in the *MET2* 3'UTR play a role in mediating stage-specificity, mutational analysis of the 3'UTR in conjunction with a reporter gene assay may be undertaken (as discussed in section 4.5.1). Such a strategy could also identify other determinants of *MET2* RNA stability. Whether ARE elements can stabilise RNA transcripts in metacyclics at all, and whether they are involved specifically in the stabilisation of the MET2 transcript remains to be shown.

The genomic location of the *MET2* gene may also affect its expression. *MET2* is located on chromosome 3, close to a strand switch (Andersson et al., 1998), and flanked on either side by nearly 5 kb of low complexity sequence. This strand switch region has been the subject of intense investigation. A chromosome fragmentation approach has demonstrated that it contains the elements that mediate mitotic stability of this chromosome. The evidence thus far suggests that this putative *T. cruzi* centromere is a regional feature, and not determined by DNA sequence alone (Obado et al., in press). This would imply some sort of epigenetic modification of the chromatin in this region. A detailed dissection of the *T. cruzi* centromere is currently underway. These studies should provide information as to the nature of these modifications, and how they affect the expression of *MET2* and other genes within this region.

It has been speculated that regulatory elements that control transcription are located within this strand switch region of *T. cruzi* chromosome 3 (Andersson et al., 1998). One strand switch region on a trypanosomatid chromosome that has been analysed experimentally is that of *L. major* chromosome 1 (Martinez-Calvillo et al., 2003).

The results of nuclear run-on assays suggest that pol II transcription initiates in a single region of <100 bp, located within the strand switch region between two divergent gene clusters. However, the sequence of this <100 bp area that contains a putative bidirectional promoter is not particularly conserved in the homologous regions in the genome of other *Leishmania* species. Similarly, no sequence conservation is evident between the chromosome 1 strand switch region as a whole, and other strand switch regions in *L. major* chromosomes (Martinez-Calvillo et al., 2003). This lack of a consensus sequence makes the prediction of putative transcription initiation sites in *T. cruzi* chromosomes difficult. It will be important to test experimentally whether such a transcriptional start site is located within the strand switch of *T. cruzi* chromosome 3.

Since *MET2* is only the second gene in the polycistronic transcription unit, it will be of interest to determine where transcription is initiated in relation to *MET2*. We have reported in the present study the generation of several *MET2* gene deletion mutants. The targeting constructs that were used to delete copies of *MET2* contained no promoter to drive transcription of the drug selectable marker gene, yet drug resistant epimastigotes were readily obtained. This indicates that in replicating cells, the *MET2* locus is accessible to the transcriptional machinery, and we would speculate that transcription of the marker gene occurs as a result of run through pol II transcription. However, this will need to be tested experimentally. Nuclear run-on assays may be used to measure transcription of the genes within or in close proximity to the strand switch, similar to the analysis of L. major chromosome 1 (Martinez-Calvillo et al., 2003) and should determine if there is any transcription from the *MET2* locus in wild-type epimastigotes.

5.6.2 Localisation of epitope tagged MET2 in epimastigotes

Two completely independent lines of evidence suggest that the MET2 protein is associated with the kinetoplast. In EM studies conducted by Yamada-Ogatta (2004) native MET2 was found associated with the kinetoplast in metacyclics of *T. cruzi* Dm28c. In this study, an epitope tagging strategy and immunofluorescence microscopy were used to determine the localisation of c-myc tagged MET2 (MET2-9E10)

207

in epimastigotes. The protein was found adjacent to, or partially overlapping with, the kinetoplast DNA. However, the fact that the MET2-9E10 protein was detected in only a very small subset of epimastigotes transfected with pTEX-MET2-9E10 makes it difficult to base firm conclusions on these data.

A number of reasons may account for the failure to detect recombinant MET2-9E10 using antibodies against the epitope tag. First, the pTEX vector which was used in the present study for overexpression of MET2 in epimastigotes is designed for constitutive overexpression (Kelly et al., 1992), which is not optimal to study a developmentally regulated protein. If MET2 is detrimental to epimastigotes, then constitutive expression is likely to result in the selection of mutants that can down-regulate MET2 expression. Previous examples of proteins that could not be expressed in epimastigotes using the pTEX vector include the mucin surface proteins (Allen, 2000) and the mitochondrial protein BPP1 (Bromley, 2002; Bromley et al., 2004). In these studies, detailed investigations were undertaken to determine the RNA levels of the drug selectable marker gene and the gene of interest. Such studies were not pursued here, because this would not have contributed directly to the aim of this study, which was to gain information about the localisation and function of the MET2 protein. Furthermore, the availability of an inducible expression system will now allow us to conduct better controlled studies.

Second, the coding sequence of MET2 may contain information that prevents the protein from being expressed in the epimastigote stage. However, if this was the case, metacyclics that develop from pTEX-MET2-9E10 transfected epimastigotes should express the recombinant protein. Yet none of the metacyclics that developed in these populations expressed MET2-9E10 at detectable levels. Finally, we cannot completely exclude the possibility that the recombinant protein was expressed but could not be detected under the conditions used. The protein could have been excreted from the cell and was therefore not detected in fixed cells or Western blots of cell extracts. Alternatively, the C-terminus of the protein, including the epitope tag may be removed in the mature protein. The *T. cruzi* Dm28c wild type MET2 protein detected in a Western blot was only 14 kDa, rather than 27kDa as expected from the full-length MET2 sequence (Yamada-Ogatta et al., 2004). This suggests that proces-

sing of the protein occurs, but currently no further information is available as to the mechanisms involved.

Taken together, the data obtained in the present study, and the data from Yamada-Ogatta (2004) suggest association of MET2 with the kinetoplast. This is compatible with bioinformatic analysis, which predicts mitochondrial localisation. This will have to be substantiated with further evidence, however. It will be a priority to generate antibodies against MET2 in order to study the expression and localisation of the wild-type protein. Inducible expression systems for *T. cruzi* have recently become available (DaRocha et al., 2004, Martin Taylor, unpublished), which will be used to address any further questions concerning the localisation of MET2 and the effects of overexpression of the protein.

5.6.3 Functional analysis of MET2

Because such a low proportion of epimastigotes expressed detectable levels of MET2-9E10, these cells were not used for detailed phenotypic analysis. The inducible expression system will be used to test the effect of MET2 induction on replicating epimastigotes.

To determine whether *MET2* is essential for development of metacyclics, we are in the process of generating null mutants. We have reported here the generation of cell lines that had two copies of the MET2 gene deleted. Gene deletion by homologous recombination is a slow process in *T. cruzi*. This is further complicated here by the fact that the *MET2* gene is present in three copies in the CL Brener genome. Yet, this is the only method by which loss of *MET2* function can be tested. Since expression of *MET2* is highly stage-specific, it is unlikely that the gene is essential in epimastigotes. It was therefore considered feasible to generate viable knockouts. The results presented here demonstrate that all three chromosomal loci can be targeted using the specific constructs built as part of this work. It is therefore expected that a third round of transfection will yield null mutants.

As the function of *MET2* is unknown, the analysis of prospective null mutants poses a challenge. The obvious first test will be to analyse whether these cells can differentiate to metacyclics, and whether they can complete the entire life-cycle.

6. CONCLUDING REMARKS

MET2 and MET3 represent two novel metacyclic-specific markers that will facilitate dissection of *T. cruzi* metacyclogenesis. The precise function of these genes remains elusive, but a number of interesting questions have arisen as a result of this study. Three main areas to follow up are outlined below.

1. Further investigation of the function and regulation of MET2 and MET3.

To gain insight into the function of *MET2* it will be a priority to delete the third allele (section 5.3), and analyse the resulting null mutants for their ability to develop into metacyclic trypomastigotes and progress through the life-cycle, similar to the analysis of the *MET3* knockouts (section 4.4). A challenge of this work, has been to design informative experiments to analyse genes of unknown function. As any genome sequenced to date contains a high proportion of novel genes, similar challenges are bound to arise frequently in post genomics research. What we present here is a first description of *MET3* knockout mutants. Clearly, more information on the biochemical properties of the encoded proteins will inform the design of more sophisticated experiments to characterise the knockouts. To gain such additional information, and ultimately determine the function of MET3 in the metacyclic nucleus, it will be essential to identify putative nucleolar interaction partners. Specifically, as outlined in section 4.5.5, the hypothesis may be tested that MET3 is a nucleic acid binding protein.

No attempts were made to study the mechanism by which *MET2* and *MET3* are regulated. To do this, transfection-based approaches, using reporter gene constructs can be used to test the hypothesis that elements within the 3'UTR of these genes may determine stage specificity, as outlined in sections 4.5.1 and 5.6.1. In addition, substances that inhibit transcription and protein synthesis can be tested on their effect on *MET2* and *MET3* expression. Specifically, treatment of cells with actinomycin D, which has previously been used to blocks transcription in *T. cruzi* (Abuin et al., 1999; Coughlin et al., 2000) can be used to assess the relative stability of the *MET2*

and *MET3* transcripts in different life-cycle stages. To determine whether there are any labile protein factors in epimastigotes that destabilise the RNA transcripts, the cells could be treated with the protein synthesis inhibitor cycloheximide. Such experiments should determine whether similar or different mechanisms control *MET2* and *MET3* expression, and they may provide insight into mechanism of metacyclic-specific gene expression.

2. The T. cruzi *nucleolus and its role in metacyclogenesis.*

In order to gain insight into the structural changes that occur in the *T. cruzi* nucleus during differentiation, it will be essential to analyse cells by electron microscopy. Immunogold labelling of thin sections can be used to visualise the precise location of the MET3 protein. This may determine whether the protein is associated with particular nucleolar sub-structures. In addition, it will be important to study the fate of the rDNA during dispersal. This can be achieved by in situ hybridisation with rDNA probes to assess the extent to which dispersal of antigens actually reflects disruption of the actual nucleolar structure. Is nucleolar dispersal a signal required to trigger metacyclogenesis? Manipulation of the nucleolar structure using chemical agents (as discussed in section 4.5.6) may test this hypothesis.

3. A robust in vitro metacyclogenesis assay.

Finally, it remains a priority to improve the experimental system to study metacyclogenesis *in vitro*. A main consideration in future experiments will be the choice of *T. cruzi* strain. A conclusion from this work is that the CL Brener strain is not an ideal model. The fact that the CL Brener strain is partially trisomic complicates genetic manipulation, as evident in the generation of the *MET2* knockout cell lines that requires deletion of three alleles. It will be important to choose a strain that consistently produces high numbers of metacyclics, as has been reported for example for the Dm28c strain in TAU medium (section 3.1.3). In addition, the catalogue of metacyclic-specific markers needs to be further expanded to facilitate a molecular description of the regulatory mechanisms that control this important transition in the *T. cruzi* life-cycle.

212

REFERENCES

Aboagye-Kwarteng, T., ole-MoiYoi, O. K., and Lonsdale-Eccles, J. D. (1991). Phosphorylation diffe-rences among proteins of bloodstream developmental stages of *Trypanosoma brucei brucei*. Biochem J *275 (Pt 1)*, 7-14.

Abuin, G., Freitas-Junior, L. H., Colli, W., Alves, M. J., and Schenkman, S. (1999). Expression of trans-sialidase and 85-kDa glycoprotein genes in *Trypanosoma cruzi* is differentially regulated at the post-transcriptional level by labile protein factors. J Biol Chem *274*, 13041-13047.

Adroher, F. J., Osuna, A., and Lupianez, J. A. (1990). Differential energetic metabolism during *Trypanosoma cruzi* differentiation. II. Hexokinase, phosphofructokinase and pyruvate kinase. Mol Cell Biochem *94*, 71-82.

Aguero, F., Verdun, R. E., Frasch, A. C., and Sanchez, D. O. (2000). A random sequencing approach for the analysis of the *Trypanosoma cruzi* genome: general structure, large gene and repetitive DNA families, and gene discovery. Genome Res *10*, 1996-2005.

Ahmad, S., Gromiha, M. M., and Sarai, A. (2004). Analysis and Prediction of DNA-binding proteins and their binding residues based on composition, sequence and structure information. Bioinformatics *20*, 477-486.

Alexandre, S., Paindavoine, P., Hanocq-Quertier, J., Paturiaux-Hanocq, F., Tebabi, P., and Pays, E. (1996). Families of adenylate cyclase genes in *Trypanosoma brucei*. Mol Biochem Parasitol *77*, 173-182.

Allaoui, A., Francois, C., Zemzoumi, K., Guilvard, E., and Ouaissi, A. (1999). Intracellular growth and metacyclogenesis defects in *Trypanosoma cruzi* carrying a targeted deletion of a Tc52 protein-encoding allele. Mol Microbiol *32*, 1273-1286.

Allen, C. L. (2000) The organisation, diversity and expression of the mucin-like glycoprotein genes of *Trypanosoma cruzi*, PhD Thesis, University of London, London.

Allen, C. L., and Kelly, J. M. (2001). *Trypanosoma cruzi*: mucin pseudogenes organized in a tandem array. Exp Parasitol *97*, 173-177.

Alsford, S., and Horn, D. (2004). Trypanosomatid histones. Mol Microbiol *53*, 365-372.

Alvarenga, N. J., and Brener, Z. (1979). Isolation of pure metacyclic trypomastigotes of *Trypanosoma cruzi* from triatomine bugs by use of a DEAE-cellulose column. J Parasitol *65*, 814-815.

Andersen, J. S., Lyon, C. E., Fox, A. H., Leung, A. K., Lam, Y. W., Steen, H., Mann, M., and Lamond, A. I. (2002). Directed proteomic analysis of the human nucleolus. Curr Biol *12*, 1-11.

Andersson, B. (2004). Analysis of the complete genome of *Trypanosoma cruzi*. Paper presented at: British Society for Parasitology Trypanosomiasis and Leishmaniasis Seminar (Ceske Budejovice).

Andersson, B., Aslund, L., Tammi, M., Tran, A. N., Hoheisel, J. D., and Pettersson, U. (1998). Complete sequence of a 93.4-kb contig from chromosome 3 of *Trypanosoma cruzi* containing a strand-switch region. Genome Res *8*, 809-816.

Andrews, N. W., Abrams, C. K., Slatin, S. L., and Griffiths, G. (1990). A *T. cruzi*-secreted protein immunologically related to the complement component C9: evidence for membrane pore-forming activity at low pH. Cell *61*, 1277-1287.

Anonymous (1999). Recommendations from a Satellite Meeting. Mem Inst Oswaldo Cruz *Suppl. 1*, 429-432.

Araya, J. E., Cano, M. I., Yoshida, N., and da Silveira, J. F. (1994). Cloning and characterization of a gene for the stage-specific 82-kDa surface antigen of metacyclic trypomastigotes of *Trypanosoma cruzi*. Mol Biochem Parasitol *65*, 161-169.

Arney, K. L., and Fisher, A. G. (2004). Epigenetic aspects of differentiation. J Cell Sci *117*, 4355-4363.

Avila, A. R., Dallagiovanna, B., Yamada-Ogatta, S. F., Monteiro-Goes, V., Fragoso, S. P., Krieger, M. A., and Goldenberg, S. (2003). Stage-specific gene expression during *Trypanosoma cruzi* meta-cyclogenesis. Genet Mol Res *2*, 159-168.

Avila, A. R., Yamada-Ogatta, S. F., da Silva Monteiro, V., Krieger, M. A., Nakamura, C. V., de Souza, W., and Goldenberg, S. (2001). Cloning and characterization of the metacyclogenin gene, which is specifically expressed during *Trypanosoma cruzi* metacyclogenesis. Mol Biochem Parasitol *117*, 169-177.

Bakalara, N., Seyfang, A., Baltz, T., and Davis, C. (1995). *Trypanosoma brucei* and *Trypanosoma cruzi*: life cycle-regulated protein tyrosine phosphatase activity. Exp Parasitol *81*, 302-312.

Barrett, M. P., Burchmore, R. J., Stich, A., Lazzari, J. O., Frasch, A. C., Cazzulo, J. J., and Krishna, S. (2003). The trypanosomiases. Lancet *362*, 1469-1480.

Bastin, P., Ellis, K., Kohl, L., and Gull, K. (2000a). Flagellum ontogeny in trypanosomes studied via an inherited and regulated RNA interference system. J Cell Sci *113*, 3321-3328.

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

Batista, J. A., Teixeira, S. M., Donelson, J. E., Kirchhoff, L. V., and de Sa, C. M. (1994). Character-ization of a *Trypanosoma cruzi* poly(A)-binding protein and its genes. Mol Biochem Parasitol *67*, 301-312.

Beard, C. A., Wrightsman, R. A., and Manning, J. E. (1988). Stage and strain specific expression of the tandemly repeated 90 kDa surface antigen gene family in *Trypanosoma cruzi*. Mol Biochem Parasitol *28*, 227-234.

Belli, S. I. (2000). Chromatin remodelling during the life cycle of trypanosomatids. Int J Parasitol *30*, 679-687.

Bellofatto, V., and Cross, G. A. (1989). Expression of a bacterial gene in a trypanosomatid protozoan. Science *244*, 1167-1169.

Benson (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acid Research *27*, 573-580.

Bentley, J. K., and Beavo, J. A. (1992). Regulation and function of cyclic nucleotides. Curr Opin Cell Biol *4*, 233-240.

Berberof, M., Vanhamme, L., Tebabi, P., Pays, A., Jefferies, D., Welburn, S., and Pays, E. (1995). The 3'-terminal region of the mRNAs for VSG and procyclin can confer stage specificity to gene expression in *Trypanosoma brucei*. EMBO J *14*, 2925-2934.

Beverley, S. M. (2003). Protozomics: trypanosomatid parasite genetics comes of age. Nat Rev Genet *4*, 11-19.

Beverley, S. M., Akopyants, N. S., Goyard, S., Matlib, R. S., Gordon, J. L., Brownstein, B. H., Stormo, G. D., Bukanova, E. N., Hott, C. T., Li, F., *et al.* (2002). Putting the *Leishmania* genome to work: functional genomics by transposon trapping and expression profiling. Philos Trans R Soc Lond B Biol Sci *357*, 47-53.

Bickmore, W. A., and Sutherland, H. G. (2002). Addressing protein localization within the nucleus. EMBO J *21*, 1248-1254.

Bieger, B., and Essen, L. O. (2001). Structural analysis of adenylate cyclases from *Trypanosoma brucei* in their monomeric state. EMBO J *20*, 433-445.

Birney, E., Kumar, S., and Krainer, A. R. (1993). Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. Nucleic Acids Res *21*, 5803-5816.

Blattner, J., and Clayton, C. E. (1995). The 3'-untranslated regions from the *Trypanosoma brucei* phosphoglycerate kinase-encoding genes mediate developmental regulation. Gene *162*, 153-156.

Bogitsh, B. J., and Cheng, T. C. (1990). Blood and tissue protozoa I: Hemoflagellates. In Human Parasitology (Saunders College Publishing).

Bonaldo, M. C., d'Escoffier, L. N., Salles, J. M., and Goldenberg, S. (1991). Characterization and expression of proteases during *Trypanosoma cruzi* metacyclogenesis. Exp Parasitol *73*, 44-51.

Bonaldo, M. C., Souto-Padron, T., de Souza, W., and Goldenberg, S. (1988). Cell-substrate adhesion during *Trypanosoma cruzi* differentiation. J Cell Biol *106*, 1349-1358.

Borst, P. (2002). Antigenic variation and allelic exclusion. Cell *109*, 5-8.

Borst, P., and Fairlamb, A. H. (1998). Surface receptors and transporters of *Trypanosoma brucei*. Annu Rev Microbiol *52*, 745-778.

Boshart, M., and Mottram, J. C. (1997). Protein Phosphorylation and Protein Kinases in Trypano-somatids. In Trypanosomiasis and Leishmaniasis, Hyde, ed. (CAB International), pp. 227-244.

Brack, C. (1968). [Electron microscopic studies on the life cycle of *Trypanosoma cruzi* with special reference to developmental forms in the vector *Rhodnius prolixus*]. Acta Trop *25*, 289-356.

Brener, Z. (1973). Biology of *Trypanosoma cruzi*. Annu Rev Microbiol *27*, 347-382.

Bringaud, F., Vedrenne, C., Cuvillier, A., Parzy, D., Baltz, D., Tetaud, E., Pays, E., Venegas, J., Merlin, G., and Baltz, T. (1998). Conserved organization of genes in trypanosomatids. Mol Biochem Parasitol *94*, 249-264.

Bromley, E. V. (2002) Characterisation of a novel protein belonging to the WD repeat family in the protozoan parasite *Trypanosoma cruzi*, PhD Thesis, University of London, London.

Bromley, E. V., Taylor, M. C., Wilkinson, S. R., and Kelly, J. M. (2004). The amino terminal domain of a novel WD repeat protein from *Trypanosoma cruzi* contains a non-canonical mitochondrial targeting signal. Int J Parasitol *34*, 63-71.

Burd, C. G., and Dreyfuss, G. (1994). Conserved structures and diversity of functions of RNA-binding proteins. Science *265*, 615-621.

Burleigh, B. A., and Andrews, N. W. (1995). The mechanisms of *Trypanosoma cruzi* invasion of mammalian cells. Annu Rev Microbiol *49*, 175-200.

Burleigh, B. A., and Andrews, N. W. (1998). Signaling and host cell invasion by *Trypanosoma cruzi*. Curr Opin Microbiol *1*, 461-465.

Caler, E. V., Vaena de Avalos, S., Haynes, P. A., Andrews, N. W., and Burleigh, B. A. (1998). Oligopeptidase B-dependent signaling mediates host cell invasion by *Trypanosoma cruzi*. EMBO J *17*, 4975-4986.

Carmo, M. S., Araya, J. E., Ramirez, M. I., Boscardin, S., Cano, M. I., Baida, R. P., Ruiz, R. C., Santos, M. R., Chiurillo, M. A., Ramirez, J. L., *et al.* (1999). Organization and expression of a multi-gene family encoding the surface glycoproteins of *Trypanosoma cruzi* metacyclic trypomastigotes involved in the cell invasion. Mem Inst Oswaldo Cruz *94*, 169-171.

Carmo-Fonseca, M., Mendes-Soares, L., and Campos, I. (2000). To be or not to be in the nucleolus. Nat Cell Biol *2*, E107-112.

Chagas, C. (1909). Über eine neue Trypanosomiasis des Menschen. Archiv für Schiffs- und Tropen-Hygiene *13*, 351-353.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. Science *263*, 802-805.

Chaves, I., Zomerdijk, J., Dirks-Mulder, A., Dirks, R. W., Raap, A. K., and Borst, P. (1998). Subnuclear localization of the active variant surface glycoprotein gene expression site in *Trypanosoma brucei*. Proc Natl Acad Sci U S A *95*, 12328-12333.

Chen, C. Y., and Shyu, A. B. (1995). AU-rich elements: characterisation and importance in mRNA degradation. TIBS *20*, 465-470.

Chen, C. Y., Xu, N., and Shyu, A. B. (2002). Highly selective actions of HuR in antagonizing AU-rich element-mediated mRNA destabilization. Mol Cell Biol *22*, 7268-7278.

Chung, S. H., and Swindle, J. (1990). Linkage of the calmodulin and ubiquitin loci in *Trypanosoma cruzi*. Nucleic Acids Res *18*, 4561-4569.

Clayton, C., Adams, M., Almeida, R., Baltz, T., Barrett, M., Bastien, P., Belli, S., Beverley, S., Biteau, N., Blackwell, J., *et al.* (1998). Genetic nomenclature for *Trypanosoma* and *Leishmania*. Mol Biochem Parasitol *97*, 221-224.

Clayton, C. E. (1999). Genetic manipulation of kinetoplastida. Parasitol Today *15*, 372-378.

Clayton, C. E. (2002). Life without transcriptional control? From fly to man and back again. EMBO J *21*, 1881-1888.

Cokol, M., Nair, R., and Rost, B. (2000). Finding nuclear localization signals. EMBO Rep *1*, 411-415.

Contreras, V. T., Morel, C. M., and Goldenberg, S. (1985a). Stage specific gene expression precedes morphological changes during *Trypanosoma cruzi* metacyclogenesis. Mol Biochem Parasitol *14*, 83-96.

Contreras, V. T., Salles, J. M., Thomas, N., Morel, C. M., and Goldenberg, S. (1985b). In vitro differentiation of *Trypanosoma cruzi* under chemically defined conditions. Mol Biochem Parasitol *16*, 315-327.

Coughlin, B. C., Teixeira, S. M., Kirchhoff, L. V., and Donelson, J. E. (2000). Amastin mRNA abundance in *Trypanosoma cruzi* is controlled by a 3'-untranslated region position-dependent cis-element and an untranslated region-binding protein. J Biol Chem *275*, 12051-12060.

Cross, G. A., and Takle, G. B. (1993). The surface trans-sialidase family of *Trypanosoma cruzi*. Annu Rev Microbiol *47*, 385-411.

D'Angelo, M. A., Sanguineti, S., Reece, J. M., Birnbaumer, L., Torres, H. N., and Flawia, M. M. (2004). Identification, characterization and subcellular localization of TcPDE1, a novel cAMP-specific phosphodiesterase from *Trypanosoma cruzi*. Biochem J *378*, 63-72.

D'Orso, I., De Gaudenzi, J. G., and Frasch, A. C. (2003). RNA-binding proteins and mRNA turnover in trypanosomes. Trends Parasitol *19*, 151-155.

D'Orso, I., and Frasch, A. C. (2001a). Functionally different AU- and G-rich cis-elements confer developmentally regulated mRNA stability in *Trypanosoma cruzi* by interaction with specific RNA-binding proteins. J Biol Chem *276*, 15783-15793.

D'Orso, I., and Frasch, A. C. (2001b). TcUBP-1, a developmentally regulated U-rich RNA-binding protein involved in selective mRNA destabilization in trypanosomes. J Biol Chem *276*, 34801-34809.

D'Orso, I., and Frasch, A. C. (2002). TcUBP-1, an mRNA destabilizing factor from trypanosomes, homodimerizes and interacts with novel AU-rich element- and Poly(A)-binding proteins forming a ribonucleoprotein complex. J Biol Chem *277*, 50520-50528.

Dallagiovanna, B., Plazanet-Menut, C., Ogatta, S. F., Avila, A. R., Krieger, M. A., and Goldenberg, S. (2001). *Trypanosoma cruzi*: a gene family encoding chitin-binding-like proteins is posttranscriptionally regulated during metacyclogenesis. Exp Parasitol *99*, 7-16.

DaRocha, W. D., Otsu, K., Teixeira, S. M., and Donelson, J. E. (2004). Tests of cytoplasmic RNA interference (RNAi) and construction of a tetracycline-inducible T7 promoter system in *Trypanosoma cruzi*. Mol Biochem Parasitol *133*, 175-186.

Das, A., Park, J. H., Hagen, C. B., and Parsons, M. (1998). Distinct domains of a nucleolar protein mediate protein kinase binding, interaction with nucleic acids and nucleolar localization. J Cell Sci *111 (Pt 17)*, 2615-2623.

Das, A., Peterson, G. C., Kanner, S. B., Frevert, U., and Parsons, M. (1996). A major tyrosine-phosphorylated protein of *Trypanosoma brucei* is a nucleolar RNA-binding protein. J Biol Chem *271*, 15675-15681.

Davis, R. E. (1996). Spliced leader RNA trans-splicing in metazoa. Parasitol Today *12*, 33-40.

De Gaudenzi, J. G., D'Orso, I., and Frasch, A. C. (2003). RNA recognition motif-type RNA-binding proteins in *Trypanosoma cruzi* form a family involved in the interaction with specific transcripts in vivo. J Biol Chem *278*, 18884-18894.

de Isola, E. L., Lammel, E. M., and Gonzalez Cappa, S. M. (1986). *Trypanosoma cruzi*: differentiation after interaction of epimastigotes and *Triatoma infestans* intestinal homogenate. Exp Parasitol *62*, 329-335.

de Isola, E. L., Lammel, E. M., Katzin, V. J., and Gonzalez Cappa, S. M. (1981). Influence of organ extracts of *Triatoma infestans* on differentiation of *Trypanosoma cruzi*. J Parasitol *67*, 53-58.

de Jesus, A. R., Cooper, R., Espinosa, M., Gomes, J. E., Garcia, E. S., Paul, S., and Cross, G. A. (1993). Gene deletion suggests a role for *Trypanosoma cruzi* surface glycoprotein GP72 in the insect and mammalian stages of the life cycle. J Cell Sci *106 (Pt 4)*, 1023-1033.

de Rooij, J., Zwartkruis, F. J., Verheijen, M. H., Cool, R. H., Nijman, S. M., Wittinghofer, A., and Bos, J. L. (1998). Epac is a Rap1 guanine-nucleotide-exchange factor directly activated by cyclic AMP. Nature *396*, 474-477.

de Souza, W. (1999). A short review on the morphology of *Trypanosoma cruzi*: from 1909 to 1999. Mem Inst Oswaldo Cruz *94 Suppl 1*, 17-36.

de Souza, W. (2002). Special organelles of some pathogenic protozoa. Parasitol Res *88*, 1013-1025.

de Souza, W. (2003). Novel Cell Biology of *Trypanosoma cruzi*. In American Trypanosomiasis, K. M. Tyler, and M. A. Miles, eds. (Boston, Kluwer Academic Publishers), pp. 13-24.

de Souza, W., Carreiro, I. P., Miranda, K., and Silva, N. L. (2000). Two special organelles found in *Trypanosoma cruzi*. An Acad Bras Cienc *72*, 421-432.

de Souza, W., and Meyer, H. (1974). On the fine structure of the nucleus in *Trypanosoma cruzi* in tissue culture forms. Spindle fibers in the dividing nucleus. J Protozool *21*, 48-52.

Dell, K. R., and Engel, J. N. (1994). Stage-specific regulation of protein phosphorylation in *Leishmania major*. Mol Biochem Parasitol *64*, 283-292.

Devera, R., Fernandes, O., and Coura, J. R. (2003). Should *Trypanosoma cruzi* be called "cruzi" complex? A review of the parasite diversity and the potential of selecting population after in vitro culturing and mice infection. Mem Inst Oswaldo Cruz *98*, 1-12.

Di Noia, J. M., Buscaglia, C. A., De Marchi, C. R., Almeida, I. C., and Frasch, A. C. (2002). A *Trypanosoma cruzi* small surface molecule provides the first immunological evidence that Chagas' disease is due to a single parasite lineage. J Exp Med *195*, 401-413.

Di Noia, J. M., D'Orso, I., Sanchez, D. O., and Frasch, A. C. (2000). AU-rich elements in the 3'-untranslated region of a new mucin-type gene family of *Trypanosoma cruzi* confers mRNA instability and modulates translation efficiency. J Biol Chem *275*, 10218-10227.

Diehl, S., Diehl, F., El-Sayed, N. M., Clayton, C., and Hoheisel, J. D. (2002). Analysis of stage-specific gene expression in the bloodstream and the procyclic form of *Trypanosoma brucei* using a genomic DNA-microarray. Mol Biochem Parasitol *123*, 115-123.

Docampo, R., Scott, D. A., Vercesi, A. E., and Moreno, S. N. (1995). Intracellular Ca2+ storage in acidocalcisomes of *Trypanosoma cruzi*. Biochem J *310 (Pt 3)*, 1005-1012.

Dundr, M., and Misteli, T. (2001). Functional architecture in the cell nucleus. Biochem J *356*, 297-310.

Durand-Dubief, M., and Bastin, P. (2003). TbAGO1, an Argonaute protein required for RNA interference, is involved in mitosis and chromosome segregation in *Trypanosoma brucei*. BMC Biol *1*, 2.

Elias, M. C., Faria, M., Mortara, R. A., Motta, M. C., de Souza, W., Thiry, M., and Schenkman, S. (2002). Chromosome localization changes in the *Trypanosoma cruzi* nucleus. Eukaryot Cell *1*, 944-953.

Elias, M. C., Marques Porto, R., Faria, M., and Schenkman, S. (2001a). Distinct mechanisms operate to control stage-specific and cell-cycle dependent gene expression in *Trypanosoma cruzi*. In Molecular Mechanisms in the Pathogenesis of Chagas Disease, J. M. Kelly, ed. (New York, Eurekah.com and Kluwer Academic/Plenum Publishers), pp. 16-24.

Elias, M. C., Marques-Porto, R., Freymuller, E., and Schenkman, S. (2001b). Transcription rate modulation through the *Trypanosoma cruzi* life cycle occurs in parallel with changes in nuclear organisation. Mol Biochem Parasitol *112*, 79-90.

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

Ersfeld, K., Melville, S. E., and Gull, K. (1999). Nuclear and genome organization of *Trypanosoma brucei*. Parasitol Today *15*, 58-63.

Estevez, A. M., and Simpson, L. (1999). Uridine insertion/deletion RNA editing in trypanosome mitochondria - a review. Gene *240*, 247-260.

Evers, R., Hammer, A., Kock, J., Jess, W., Borst, P., Memet, S., and Cornelissen, A. W. (1989). *Trypanosoma brucei* contains two RNA polymerase II largest subunit genes with an altered C-terminal domain. Cell *56*, 585-597.

Fantoni, A., Dare, A. O., and Tschudi, C. (1994). RNA polymerase III-mediated transcription of the trypanosome U2 small nuclear RNA gene is controlled by both intragenic and extragenic regulatory elements. Mol Cell Biol *14*, 2021-2028.

Favoreto, S., Jr., Dorta, M. L., and Yoshida, N. (1998). *Trypanosoma cruzi* 175-kDa protein tyrosine phosphorylation is associated with host cell invasion. Exp Parasitol *89*, 188-194.

Ficarro, S. B., McCleland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., and White, F. M. (2002). Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. Nat Biotechnol *20*, 301-305.

Fimia, G. M., and Sassone-Corsi, P. (2001). Cyclic AMP signalling. J Cell Sci *114*, 1971-1972.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature *391*, 806-811.

Fragoso, S. P., Plazanet-Menut, C., Carreira, M. A., Motta, M. C., Dallagiovana, B., Krieger, M. A., and Goldenberg, S. (2003). Cloning and characterization of a gene encoding a putative protein associated with U3 small nucleolar ribonucleoprotein in *Trypanosoma cruzi*. Mol Biochem Parasitol *126*, 113-117.

Fraidenraich, D., Pena, C., Isola, E. L., Lammel, E. M., Coso, O., Anel, A. D., Pongor, S., Baralle, F., Torres, H. N., and Flawia, M. M. (1993). Stimulation of *Trypanosoma cruzi* adenylyl cyclase by an alpha D-globin fragment from *Triatoma* hindgut: effect on differentiation of epimastigote to trypomastigote forms. Proc Natl Acad Sci U S A *90*, 10140-10144.

Franco, F. R., Paranhos-Bacalla, G. S., Yamauchi, L. M., Yoshida, N., and da Silveira, J. F. (1993). Characterization of a cDNA clone encoding the carboxy-terminal domain of a 90-kilodalton surface antigen of *Trypanosoma cruzi* metacyclic trypomastigotes. Infect Immun *61*, 4196-4201.

Furger, A., Schurch, N., Kurath, U., and Roditi, I. (1997). Elements in the 3' untranslated region of procyclin mRNA regulate expression in insect forms of *Trypanosoma brucei* by modulating RNA stability and translation. Mol Cell Biol *17*, 4372-4380.

Furuya, T., Kashuba, C., Docampo, R., and Moreno, S. N. (2000). A novel phosphatidylinositol-phospholipase C of *Trypanosoma cruzi* that is lipid modified and activated during trypomastigote to amastigote differentiation. J Biol Chem *275*, 6428-6438.

Gale, M., Jr., Carter, V., and Parsons, M. (1994). Cell cycle-specific induction of an 89 kDa serine/threonine protein kinase activity in *Trypanosoma brucei*. J Cell Sci *107*, 1825-1832.

Garcia, E. S., Gonzalez, M. S., de Azambuja, P., Baralle, F. E., Fraidenraich, D., Torres, H. N., and Flawia, M. M. (1995). Induction of *Trypanosoma cruzi* metacyclogenesis in the gut of the hemato-phagous insect vector, *Rhodnius prolixus*, by hemoglobin and peptides carrying alpha D-globin sequences. Exp Parasitol *81*, 255-261.

Garcia, S. N., and Pillus, L. (1999). Net results of nucleolar dynamics. Cell *97*, 825-828.

Gerber, D. A., Souquere-Besse, S., Puvion, F., Dubois, M. F., Bensaude, O., and Cochet, C. (2000). Heat-induced relocalization of protein kinase CK2. Implication of CK2 in the context of cellular stress. J Biol Chem *275*, 23919-23926.

Gerbi, S. A., Borovjagin, A. V., and Lange, T. S. (2003). The nucleolus: a site of ribonucleoprotein maturation. Curr Opin Cell Biol *15*, 318-325.

Ghedin, E., Bringaud, F., Peterson, J., Myler, P., Berriman, M., Ivens, A., Andersson, B., Bontempi, E., Eisen, J., Angiuoli, S., *et al.* (2004). Gene synteny and evolution of genome architecture in trypanosomatids. Mol Biochem Parasitol *134*, 183-191.

Gibson, W. C., Swinkels, B. W., and Borst, P. (1988). Post-transcriptional control of the differential expression of phosphoglycerate kinase genes in *Trypanosoma brucei*. J Mol Biol *201*, 315-325.

Gilinger, G., and Bellofatto, V. (2001). Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. Nucleic Acids Res *29*, 1556-1564.

Gomez, M. L., Erijman, L., Arauzo, S., Torres, H. N., and Tellez-Inon, M. T. (1989). Protein kinase C in *Trypanosoma cruzi* epimastigote forms: partial purification and characterization. Mol Biochem Parasitol *36*, 101-108.

Gong, K. W., Kunz, S., Zoraghi, R., Kunz Renggli, C., Brun, R., and Seebeck, T. (2001). cAMP-specific phosphodiesterase TbPDE1 is not essential in *Trypanosoma brucei* in culture or during midgut infection of tsetse flies. Mol Biochem Parasitol *116*, 229-232.

Gonzales-Perdomo, M., Romero, P., and Goldenberg, S. (1988). Cyclic AMP and adenylate cyclase activators stimulate *Trypanosoma cruzi* differentiation. Exp Parasitol *66*, 205-212.

Greig, S., and Ashall, F. (1990). Electrophoretic detection of *Trypanosoma cruzi* peptidases. Mol Biochem Parasitol *39*, 31-37.

Gruszynski, A. E., DeMaster, A., Hooper, N. M., and Bangs, J. D. (2003). Surface coat remodeling during differentiation of *Trypanosoma brucei*. J Biol Chem *278*, 24665-24672.

Gunzl, A., Ullu, E., Dorner, M., Fragoso, S. P., Hoffmann, K. F., Milner, J. D., Morita, Y., Nguu, E. K., Vanacova, S., Wunsch, S., *et al.* (1997). Transcription of the *Trypanosoma brucei* spliced leader RNA gene is dependent only on the presence of upstream regulatory elements. Mol Biochem Parasitol *85*, 67-76.

Hall, N., Berriman, M., Lennard, N. J., Harris, B. R., Hertz-Fowler, C., Bart-Delabesse, E. N., Gerrard, C. S., Atkin, R. J., Barron, A. J., Bowman, S., *et al.* (2003). The DNA sequence of chromosome I of an African trypanosome: gene content, chromosome organisation, recombination and polymorphism. Nucleic Acids Res *31*, 4864-4873.

Hariharan, S., Ajioka, J., and Swindle, J. (1993). Stable transformation of *Trypanosoma cruzi*: inactivation of the PUB12.5 polyubiquitin gene by targeted gene disruption. Mol Biochem Parasitol *57*, 15-30.

Hayward, R. E., Derisi, J. L., Alfadhli, S., Kaslow, D. C., Brown, P. O., and Rathod, P. K. (2000). Shotgun DNA microarrays and stage-specific gene expression in *Plasmodium falciparum* malaria. Mol Microbiol *35*, 6-14.

Heath, S., Hieny, S., and Sher, A. (1990). A cyclic AMP inducible gene expressed during the development of infective stages of *Trypanosoma cruzi*. Mol Biochem Parasitol *43*, 133-141.

Hendriks, E., van Deursen, F. J., Wilson, J., Sarkar, M., Timms, M., and Matthews, K. R. (2000). Life-cycle differentiation in *Trypanosoma brucei*: molecules and mutants. Biochem Soc Trans *28*, 531-536.

Hendriks, E. F., Robinson, D. R., Hinkins, M., and Matthews, K. R. (2001). A novel CCCH protein which modulates differentiation of *Trypanosoma brucei* to its procyclic form. EMBO J *20*, 6700-6711.

Hoek, M., Engstler, M., and Cross, G. A. (2000). Expression-site-associated gene 8 (ESAG8) of *Trypanosoma brucei* is apparently essential and accumulates in the nucleolus. J Cell Sci *113 (Pt 22)*, 3959-3968.

Homsy, J. J., Granger, B., and Krassner, S. M. (1989). Some factors inducing formation of metacyclic stages of *Trypanosoma cruzi*. J Protozool *36*, 150-153.

Hotchkiss, T. L., Nerantzakis, G. E., Dills, S. C., Shang, L., and Read, L. K. (1999). *Trypanosoma brucei* poly(A) binding protein I cDNA cloning, expression, and binding to 5 untranslated region sequence elements. Mol Biochem Parasitol *98*, 117-129.

Hulo, N., Sigrist, C. J., Le Saux, V., Langendijk-Genevaux, P. S., Bordoli, L., Gattiker, A., De Castro, E., Bucher, P., and Bairoch, A. (2004). Recent improvements to the PROSITE database. Nucleic Acids Res *32 Database issue*, D134-137.

Hurley, J. H. (1998). The adenylyl and guanylyl cyclase superfamily. Curr Opin Struct Biol *8*, 770-777.

Iseki, M., Matsunaga, S., Murakami, A., Ohno, K., Shiga, K., Yoshida, K., Sugai, M., Takahashi, T., Hori, T., and Watanabe, M. (2002). A blue-light-activated adenylyl cyclase mediates photoavoidance in *Euglena gracilis*. Nature *415*, 1047-1051.

Ismaili, N., Perez-Morga, D., Walsh, P., Mayeda, A., Pays, A., Tebabi, P., Krainer, A. R., and Pays, E. (1999). Characterization of a SR protein from *Trypanosoma brucei* with homology to RNA-binding cis-splicing proteins. Mol Biochem Parasitol *102*, 103-115.

Janz, L., and Clayton, C. (1994). The PARP and rRNA promoters of *Trypanosoma brucei* are composed of dissimilar sequence elements that are functionally interchangeable. Mol Cell Biol *14*, 5804-5811.

Jensen, B. C., Wang, Q., Kifer, C. T., and Parsons, M. (2003). The NOG1 GTP-binding protein is required for biogenesis of the 60 S ribosomal subunit. J Biol Chem *278*, 32204-32211.

Kamath, R. S., and Ahringer, J. (2003). Genome-wide RNAi screening in *Caenorhabditis elegans*. Methods *30*, 313-321.

Keith, K., Hide, G., and Tait, A. (1990). Characterisation of protein kinase C like activities in *Trypanosoma brucei*. Mol Biochem Parasitol *43*, 107-116.

Kelley, L. A., MacCallum, R. M., and Sternberg, M. J. (2000). Enhanced genome annotation using structural profiles in the program 3D-PSSM. J Mol Biol *299*, 499-520.

Kelly, J. M. (1993). Isolation of DNA and RNA from *Leishmania*. In Protocols in Molecular Parasitology, J. E. Hyde, ed. (New Jersey, Humana Press), pp. 123-131.

Kelly, J. M., Das, P., and Tomas, A. M. (1994). An approach to functional complementation by introduction of large DNA fragments into *Trypanosoma cruzi* and *Leishmania donovani* using a cosmid shuttle vector. Mol Biochem Parasitol *65*, 51-62.

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 Acids Res *20*, 3963-3969.

Kleffmann, T., Schmidt, J., and Schaub, G. A. (1998). Attachment of *Trypanosoma cruzi* epimastigotes to hydrophobic substrates and use of this property to separate stages and promote metacyclogenesis. J Eukaryot Microbiol *45*, 548-555.

Kock, J., Evers, R., and Cornelissen, A. W. (1988). Structure and sequence of the gene for the largest subunit of trypanosomal RNA polymerase III. Nucleic Acids Res *16*, 8753-8772.

Kohl, L., and Gull, K. (1998). Molecular architecture of the trypanosome cytoskeleton. Mol Biochem Parasitol *93*, 1-9.

Kollien, A. H., and Schaub, G. A. (2000). The development of *Trypanosoma cruzi* in triatominae. Parasitol Today *16*, 381-387.

Krassner, S. M., Chang, J., Pak, S., Luc, K. O., and Granger, B. (1993). Absence of transitory [Ca2+]i flux during early in vitro metacyclogenesis of *Trypanosoma cruzi*. J Eukaryot Microbiol *40*, 224-230.

Krassner, S. M., Granger, B., Lee, P., Guerra, C., Le, T., and Luc, K. O. (1991). Action of exogenous potassium and calcium ions on in vitro metacyclogenesis in *Trypanosoma cruzi*. J Protozool *38*, 602-608.

Krassner, S. M., Granger, B., Phermsangngnam, P., Le, T., and Linden, V. (1990). Further studies on substrates inducing metacyclogenesis in *Trypanosoma cruzi*. J Protozool *37*, 128-132.

Krieger, M. A., Avila, A. R., Ogatta, S. F., Plazanet-Menut, C., and Goldenberg, S. (1999). Differential gene expression during *Trypanosoma cruzi* metacyclogenesis. Mem Inst Oswaldo Cruz *94*, 165-168.

Krieger, M. A., and Goldenberg, S. (1998). Representation of Differential Expression: A new approach to study differential gene expression in trypanosomatids. Parasitology Today *14*, 163-166.

Kunz, S., Kloeckner, T., Essen, L. O., Seebeck, T., and Boshart, M. (2004). TbPDE1, a novel class I phosphodiesterase of *Trypanosoma brucei*. Eur J Biochem *271*, 637-647.

LaCount, D. J., Bruse, S., Hill, K. L., and Donelson, J. E. (2000). Double-stranded RNA interference in *Trypanosoma brucei* using head-to-head promoters. Mol Biochem Parasitol *111*, 67-76.

Lambert, C., Leonard, N., De Bolle, X., and Depiereux, E. (2002). ESyPred3D: Prediction of proteins 3D structures. Bioinformatics *18*, 1250-1256.

Lanar, D. E., and Manning, J. E. (1984). Major surface proteins and antigens on the different in vivo and in vitro forms of *Trypanosoma cruzi*. Mol Biochem Parasitol *11*, 119-131.

Lander, E. S. (1996). The new genomics: global views of biology. Science *274*, 536-539.

Levin, L. R., and Reed, R. R. (1995). Identification of functional domains of adenylyl cyclase using in vivo chimeras. J Biol Chem *270*, 7573-7579.

Li, J., Ruyechan, W. T., and Williams, N. (2003). Stage-specific translational efficiency and protein stability regulate the developmental expression of p37, an RNA binding protein from *Trypanosoma brucei*. Biochem Biophys Res Commun *306*, 918-923.

Lippman, Z., and Martienssen, R. (2004). The role of RNA interference in heterochromatic silencing. Nature *431*, 364-370.

Liu, Y., Liang, S., and Tartakoff, A. M. (1996). Heat shock disassembles the nucleolus and inhibits nuclear protein import and poly(A)+ RNA export. EMBO J *15*, 6750-6757.

Lodish, H., Baltimore, D., Berk, A., Zipursky, S. L., Matsudaira, P., Darnell, J. (1995). Molecular Cell Biology, Third edn (New York, Scientific American Books).

Machado, C. A., and Ayala, F. J. (2001). Nucleotide sequences provide evidence of genetic exchange among distantly related lineages of *Trypanosoma cruzi*. Proc Natl Acad Sci U S A *98*, 7396-7401.

Mair, G., Shi, H., Li, H., Djikeng, A., Aviles, H. O., Bishop, J. R., Falcone, F. H., Gavrilescu, C., Montgomery, J. L., Santori, M. I., *et al.* (2000). A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. Rna *6*, 163-169.

Maldonado, R. A., Linss, J., Thomaz, N., Olson, C. L., Engman, D. M., and Goldenberg, S. (1997). Homologues of the 24-kDa flagellar Ca(2+)-binding protein gene of *Trypanosoma cruzi* are present in other members of the Trypanosomatidae family. Exp Parasitol *86*, 200-205.

Manger, I. D., and Boothroyd, J. C. (1998). Identification of a nuclear protein in *Trypanosoma brucei* with homology to RNA-binding proteins from cis-splicing systems. Mol Biochem Parasitol *97*, 1-11.

Manning-Cela, R., Cortes, A., Gonzalez-Rey, E., Van Voorhis, W. C., Swindle, J., and Gonzalez. A. (2001). LYT1 protein is required for efficient in vitro infection by *Trypanosoma cruzi*. Infect Immun *69*, 3916-3923.

Manuelidis, L. (1990). A view of interphase chromosomes. Science *250*, 1533-1540.

Marchetti, M. A., Tschudi, C., Kwon, H., Wolin, S. L., and Ullu, E. (2000). Import of proteins into the trypanosome nucleus and their distribution at karyokinesis. J Cell Sci *113 (Pt 5)*, 899-906.

Marques Porto, R., Amino, R., Elias, M. C., Faria, M., and Schenkman, S. (2002). Histone H1 is phosphorylated in non-replicating and infective forms of *Trypanosoma cruzi*. Mol Biochem Parasitol *119*, 265-271.

Martinez-Calvillo, S., Lopez, I., and Hernandez, R. (1997). pRIBOTEX expression vector: a pTEX derivative for a rapid selection of *Trypanosoma cruzi* transfectants. Gene *199*, 71-76.

Martinez-Calvillo, S., Nguyen, D., Stuart, K., and Myler, P. J. (2004). Transcription initiation and termination on *Leishmania major* chromosome 3. Eukaryot Cell *3*, 506-517.

Martinez-Calvillo, S., Yan, S., Nguyen, D., Fox, M., Stuart, K., and Myler, P. J. (2003). Transcription of *Leishmania major* Friedlin chromosome 1 initiates in both directions within a single region. Mol Cell *11*, 1291-1299.

Matthews, K. R. (1999). Developments in the differentiation of *Trypanosoma brucei*. Parasitol Today *15*, 76-80.

Matthews, K. R., Ellis, J. R., and Paterou, A. (2004). Molecular regulation of the life cycle of African trypanosomes. Trends Parasitol *20*, 40-47.

Matthews, K. R., Sherwin, T., and Gull, K. (1995). Mitochondrial genome repositioning during the differentiation of the African trypanosome between life cycle forms is microtubule mediated. J Cell Sci *108 (Pt 6)*, 2231-2239.

Matthews, K. R., Tschudi, C., and Ullu, E. (1994). A common pyrimidine-rich motif governs trans-splicing and polyadenylation of tubulin polycistronic pre-mRNA in trypanosomes. Genes Dev *8*, 491-501.

McBride, A. E., and Silver, P. A. (2001). State of the arg: protein methylation at arginine comes of age. Cell *106*, 5-8.

McGuffin, L. J., Bryson, K., and Jones, D. T. (2000). The PSIPRED protein structure prediction server. Bioinformatics *16*, 404-405.

Meister, G., and Tuschl, T. (2004). Mechanisms of gene silencing by double-stranded RNA. Nature *431*, 343-349.

Melese, T., and Xue, Z. (1995). The nucleolus: an organelle formed by the act of building a ribosome. Curr Opin Cell Biol *7*, 319-324.

Miles, M. A. (1993). Culturing and biological cloning of *Trypanosoma cruzi*. Methods Mol Biol *21*, 15-28.

Miles, M. A., Feliciangeli, M. D., and de Arias, A. R. (2003). American trypanosomiasis (Chagas' disease) and the role of molecular epidemiology in guiding control strategies. BMJ *326*, 1444-1448.

Misteli, T. (2000). Cell biology of transcription and pre-mRNA splicing: nuclear architecture meets nuclear function. J Cell Sci *113 (Pt 11)*, 1841-1849.

Moncayo, A. (1999). Progress towards interruption of transmission of Chagas disease. Mem Inst Oswaldo Cruz *94*, 401-404.

Moore, M. S., DeZazzo, J., Luk, A. Y., Tully, T., Singh, C. M., and Heberlein, U. (1998). Ethanol intoxication in *Drosophila*: Genetic and pharmacological evidence for regulation by the cAMP signaling pathway. Cell *93*, 997-1007.

Moreira-Leite, F. F., Sherwin, T., Kohl, L., and Gull, K. (2001). A trypanosome structure involved in transmitting cytoplasmic information during cell division. Science *294*, 610-612.

Morel, C., Chiari, E., Camargo, E. P., Mattei, D. M., Romanha, A. J., and Simpson, L. (1980). Strains and clones of *Trypanosoma cruzi* can be characterized by pattern of restriction endonuclease products of kinetoplast DNA minicircles. Proc Natl Acad Sci U S A *77*, 6810-6814.

Morking, P. A., Dallagiovanna, B. M., Foti, L., Garat, B., Picchi, G. F., Umaki, A. C., Probst, C. M., Krieger, M. A., Goldenberg, S., and Fragoso, S. P. (2004). TcZFP1: a CCCH zinc finger protein of *Trypanosoma cruzi* that binds poly-C oligoribonucleotides in vitro. Biochem Biophys Res Commun *319*, 169-177.

Mynott, T. L., Ladhams, A., Scarmato, P., and Engwerda, C. R. (1999). Bromelain, from pineapple stems, proteolytically blocks activation of extracellular regulated kinase-2 in T cells. J Immunol *163*, 2568-2575.

Nakai, K., and Kanehisa, M. (1992). A knowledge base for predicting protein localization sites in eukaryotic cells. Genomics *14*, 897-911.

Naula, C., Schaub, R., Leech, V., Melville, S., and Seebeck, T. (2001). Spontaneous dimerization and leucine-zipper induced activation of the recombinant catalytic domain of a new adenylyl cyclase of *Trypanosoma brucei*, GRESAG4.4B. Mol Biochem Parasitol *112*, 19-28.

Naula, C., and Seebeck, T. (2000). Cyclic AMP signaling in trypanosomatids. Parasitol Today *16*, 35-38.

Navarro, M., and Gull, K. (2001). A pol I transcriptional body associated with VSG mono-allelic expression in *Trypanosoma brucei*. Nature *414*, 759-763.

Ngo, H., Tschudi, C., Gull, K., and Ullu, E. (1998). Double-stranded RNA induces mRNA degradation in *Trypanosoma brucei*. Proc Natl Acad Sci U S A *95*, 14687-14692.

Nilsen, T. W. (1995). Trans-splicing: an update. Mol Biochem Parasitol *73*, 1-6.

Nishizuka, Y. (1992). Intracellular signaling by hydrolysis of phospholipids and activation of protein kinase C. Science *258*, 607-614.

Nitz, N., Gomes, C., de Cassia Rosa, A., D'Souza-Ault, M. R., Moreno, F., Lauria-Pires, L., Nascimento, R. J., and Teixeira, A. R. (2004). Heritable integration of kDNA minicircle sequences from *Trypanosoma cruzi* into the avian genome: insights into human Chagas disease. Cell *118*, 175-186.

Nogueira, N., Bianco, C., and Cohn, Z. (1975). Studies on the selective lysis and purification of *Trypanosoma cruzi*. J Exp Med *142*, 224-229.

Nozaki, T., and Cross, G. A. (1995). Effects of 3' untranslated and intergenic regions on gene expression in *Trypanosoma cruzi*. Mol Biochem Parasitol *75*, 55-67.

Nunes, L. R., Carvalho, M. R., Shakarian, A. M., and Buck, G. A. (1997). The transcription promoter of the spliced leader gene from *Trypanosoma cruzi*. Gene *188*, 157-168.

Obado, S. O., Taylor, M. C., Wilkinson, S. R., Bromley, E. V., and Kelly, J. M. Functional mapping of a trypanosome centromere by chromosome fragmentation identifies a 16 kb GC-rich transcriptional "strand-switch" domain as a major feature. Genome Research, in press.

Ochatt, C. M., Ulloa, R. M., Torres, H. N., and Tellez-Inon, M. T. (1993). Characterization of the catalytic subunit of *Trypanosoma cruzi* cyclic AMP-dependent protein kinase. Mol Biochem Parasitol *57*, 73-81.

Oda, Y., Nagasu, T., and Chait, B. T. (2001). Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. Nat Biotechnol *19*, 379-382.

Ogbadoyi, E., Ersfeld, K., Robinson, D., Sherwin, T., and Gull, K. (2000). Architecture of the *Trypanosoma brucei* nucleus during interphase and mitosis. Chromosoma *108*, 501-513.

Olson, M. O. (2004). Sensing cellular stress: another new function for the nucleolus? Sci STKE *2004*, pe10.

Orr, G. A., Werner, C., Xu, J., Bennett, M., Weiss, L. M., Takvorkan, P., Tanowitz, H. B., and Wittner, M. (2000). Identification of novel serine/threonine protein phosphatases in *Trypanosoma cruzi*: a potential role in control of cytokinesis and morphology. Infect Immun *68*, 1350-1358.

Ouaissi, A., Cornette, J., Schoneck, R., Plumas-Marty, B., Taibi, A., Loyens, M., and Capron, A. (1992). Fibronectin cleavage fragments provide a growth factor-like activity for the differentiation of *Trypanosoma cruzi* trypomastigotes to amastigotes. Eur J Cell Biol *59*, 68-79.

Ouaissi, M. A., Dubremetz, J. F., Schoneck, R., Fernandez-Gomez, R., Gomez-Corvera, R., Billaut-Mulot, O., Taibi, A., Loyens, M., Tartar, A., Sergheraert, C., and et al. (1995). *Trypanosoma cruzi*: a 52-kDa protein sharing sequence homology with glutathione S-transferase is localized in parasite organelles morphologically resembling reservosomes. Exp Parasitol *81*, 453-461.

Ouali, M., and King, R. D. (2000). Cascaded multiple classifiers for secondary structure prediction. Protein Science *9*, 1162-1176.

Paindavoine, P., Rolin, S., Van Assel, S., Geuskens, M., Jauniaux, J. C., Dinsart, C., Huet, G., and Pays, E. (1992). A gene from the variant surface glycoprotein expression site encodes one of several transmembrane adenylate cyclases located on the flagellum of *Trypanosoma brucei*. Mol Cell Biol *12*, 1218-1225.

Pandey, A., and Mann, M. (2000). Proteomics to study genes and genomes. Nature *405*, 837-846.

Park, J. H., Brekken, D. L., Randall, A. C., and Parsons, M. (2002). Molecular cloning of *Trypanosoma brucei* CK2 catalytic subunits: the alpha isoform is nucleolar and phosphorylates the nucleolar protein Nopp44/46. Mol Biochem Parasitol *119*, 97-106.

Park, J. H., Jensen, B. C., Kifer, C. T., and Parsons, M. (2001). A novel nucleolar G-protein conserved in eukaryotes. J Cell Sci *114*, 173-185.

Parsons, M. (2004). Glycosomes: parasites and the divergence of peroxisomal purpose. Mol Microbiol *53*, 717-724.

Parsons, M., and Ruben, L. (2000). Pathways involved in environmental sensing in trypanosomatids. Parasitol Today *16*, 56-62.

Parsons, M., Valentine, M., and Carter, V. (1993). Protein kinases in divergent eukaryotes: identification of protein kinase activities regulated during trypanosome development. Proc Natl Acad Sci U S A *90*, 2656-2660.

Parsons, M., Valentine, M., Deans, J., Schieven, G. L., and Ledbetter, J. A. (1991). Distinct patterns of tyrosine phosphorylation during the life cycle of *Trypanosoma brucei*. Mol Biochem Parasitol *45*, 241-248.

Pasion, S. G., Hines, J. C., Ou, X., Mahmood, R., and Ray, D. S. (1996). Sequences within the 5' untranslated region regulate the levels of a kinetoplast DNA topoisomerase mRNA during the cell cycle. Mol Cell Biol *16*, 6724-6735.

Pedroso, A., Cupolillo, E., and Zingales, B. (2003). Evaluation of *Trypanosoma cruzi* hybrid stocks based on chromosomal size variation. Mol Biochem Parasitol *129*, 79-90.

Pitula, J., Ruyechan, W. T., and Williams, N. (2002a). Two novel RNA binding proteins from *Trypanosoma brucei* are associated with 5S rRNA. Biochem Biophys Res Commun *290*, 569-576.

Pitula, J. S., Park, J., Parsons, M., Ruyechan, W. T., and Williams, N. (2002b). Two families of RNA binding proteins from *Trypanosoma brucei* associate in a direct protein-protein interaction. Mol Biochem Parasitol *122*, 81-89.

Pollastri, G., Przybylski, D., Rost, B., and Baldi, P. (2002). Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles. Proteins *47*, 228-235.

Porcel, B. M., Tran, A. N., Tammi, M., Nyarady, Z., Rydaker, M., Urmenyi, T. P., Rondinelli, E., Pettersson, U., Andersson, B., and Aslund, L. (2000). Gene survey of the pathogenic protozoan *Trypanosoma cruzi*. Genome Res *10*, 1103-1107.

Portal, D., Espinosa, J. M., Lobo, G. S., Kadener, S., Pereira, C. A., De La Mata, M., Tang, Z., Lin, R. J., Kornblihtt, A. R., Baralle, F. E., *et al.* (2003). An early ancestor in the evolution of splicing: a *Trypanosoma cruzi* serine-arginine-rich protein (TcSR) is functional in cis-splicing. Mol Biochem Parasitol *127*, 37-46.

Predic, J., Soskic, V., Bradley, D., and Godovac-Zimmermann, J. (2002). Monitoring of gene expression by functional proteomics: response of human lung fibroblast cells to stimulation by endothelin-1. Biochemistry *41*, 1070-1078.

Rangel-Aldao, R., Allende, O., and Cayama, E. (1985). A unique type of cyclic AMP-binding protein of *Trypanosoma cruzi*. Mol Biochem Parasitol *14*, 75-81.

Rangel-Aldao, R., Allende, O., Triana, F., Piras, R., Henriquez, D., and Piras, M. (1987). Possible role of cAMP in the differentiation of *Trypanosoma cruzi*. Mol Biochem Parasitol *22*, 39-43.

Rangel-Aldao, R., Comach, G., Allende, O., Cayama, E., Delgado, V., Piras, R., Piras, M., Henriquez, D., and Negri, S. (1986). *Trypanosoma cruzi*: polypeptide markers of epimastigotes and trypomastigotes. Mol Biochem Parasitol *20*, 25-32.

Rangel-Aldao, R., Tovar, G., and Ledezma de Ruiz, M. (1983). The cAMP receptor protein of *Trypanosoma cruzi*. J Biol Chem *258*, 6979-6983.

Robinson, K. A., and Beverley, S. M. (2003). Improvements in transfection efficiency and tests of RNA interference (RNAi) approaches in the protozoan parasite *Leishmania*. Mol Biochem Parasitol *128*, 217-228.

Rodgers, N. D., Jiao, X., and Kiledjian, M. (2002). Identifying mRNAs bound by RNA-binding proteins using affinity purification and differential display. Methods *26*, 115-122.

Roditi, I., Schwarz, H., Pearson, T. W., Beecroft, R. P., Liu, M. K., Richardson, J. P., Buhring, H. J., Pleiss, J., Bulow, R., Williams, R. O., and et al. (1989). Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of *Trypanosoma brucei*. J Cell Biol *108*, 737-746.

Rodriques Coura, J., and de Castro, S. L. (2002). A critical review on Chagas disease chemotherapy. Mem Inst Oswaldo Cruz *97*, 3-24.

Rolin, S., Paindavoine, P., Hanocq-Quertier, J., Hanocq, F., Claes, Y., Le Ray, D., Overath, P., and Pays, E. (1993). Transient adenylate cyclase activation accompanies differentiation of *Trypanosoma brucei* from bloodstream to procyclic forms. Mol Biochem Parasitol *61*, 115-125.

Rout, M. P., and Field, M. C. (2001). Isolation and characterization of subnuclear compartments from *Trypanosoma brucei*. Identification of a major repetitive nuclear lamina component. J Biol Chem *276*, 38261-38271.

Rubbi, C. P., and Milner, J. (2003). Disruption of the nucleolus mediates stabilization of p53 in response to DNA damage and other stresses. EMBO J *22*, 6068-6077.

Ruben, L., Kelly, J. M., and Chakrabarti, D. (2002). Intracellular Signaling. In Molecular and Medical Parasitology, R. Koumuniecki, Nilsen, T. and Marr, J.J. eds., (London, Academic Press).

Rudenko, G., Chung, H. M., Pham, V. P., and Van der Ploeg, L. H. (1991). RNA polymerase I can mediate expression of CAT and neo protein-coding genes in *Trypanosoma brucei*. EMBO J *10*, 3387-3397.

Ruepp, S., Furger, A., Kurath, U., Renggli, C. K., Hemphill, A., Brun, R., and Roditi, I. (1997). Survival of *Trypanosoma brucei* in the tsetse fly is enhanced by the expression of specific forms of procyclin. J Cell Biol *137*, 1369-1379.

Ruiz, R. C., Favoreto, S., Jr., Dorta, M. L., Oshiro, M. E., Ferreira, A. T., Manque, P. M., and Yoshida, N. (1998). Infectivity of *Trypanosoma cruzi* strains is associated with differential expression of surface glycoproteins with differential Ca2+ signalling activity. Biochem J *330*, 505-511.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., and Barrell, B. (2000). Artemis: sequence visualization and annotation. Bioinformatics *16*, 944-945.

232

Salazar, N. A., Mondragon, A., and Kelly, J. M. (1996). Mucin-like glycoprotein genes are closely linked to members of the trans-sialidase super-family at multiple sites in the *Trypanosoma cruzi* genome. Mol Biochem Parasitol *78*, 127-136.

Sambrook, J., Fitsch, R. F., Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, NY, Cold Spring Harbor Laboratory Press).

Sanchez, G., Wallace, A., Olivares, M., Diaz, N., Aguilera, X., Apt, W., and Solari, A. (1990). Biological characterization of *Trypanosoma cruzi* zymodemes: in vitro differentiation of epimastigotes and infectivity of culture metacyclic trypomastigotes to mice. Exp Parasitol *71*, 125-133.

Schaub, G. A. (1989). *Trypanosoma cruzi*: quantitative studies of development of two strains in small intestine and rectum of the vector *Triatoma infestans*. Exp Parasitol *68*, 260-273.

Scheer, U., and Hock, R. (1999). Structure and function of the nucleolus. Curr Opin Cell Biol *11*, 385-390.

Schenkman, S., Eichinger, D., Pereira, M. E., and Nussenzweig, V. (1994). Structural and functional properties of *Trypanosoma* trans-sialidase. Annu Rev Microbiol *48*, 499-523.

Schenkman, S., Elias, M. C., Dossin, F. M., and Chagas da Cunha, J. P. (2004). Orchestration of replication and transcription in the nucleus of *Trypanosoma cruzi*. Paper presented at: British Society for Parasitology Trypanosomiasis and Leishmaniasis Seminar (Ceske Budejovice).

Scherl, A., Coute, Y., Deon, C., Calle, A., Kindbeiter, K., Sanchez, J. C., Greco, A., Hochstrasser, D., and Diaz, J. J. (2002). Functional proteomic analysis of human nucleolus. Mol Biol Cell *13*, 4100-4109.

Schurch, N., Furger, A., Kurath, U., and Roditi, I. (1997). Contributions of the procyclin 3' untranslated region and coding region to the regulation of expression in bloodstream forms of *Trypanosoma brucei*. Mol Biochem Parasitol *89*, 109-121.

Schurch, 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. Mol Cell Biol *14*, 3668-3675.

Seebeck, T., Gong, K., Kunz, S., Schaub, R., Shalaby, T., and Zoraghi, R. (2001). cAMP signalling in *Trypanosoma brucei*. Int J Parasitol *31*, 491-498.

Shalaby, T., Liniger, M., and Seebeck, T. (2001). The regulatory subunit of a cGMP-regulated protein kinase A of *Trypanosoma brucei*. Eur J Biochem *268*, 6197-6206.

233

Shi, H., Djikeng, A., Tschudi, C., and Ullu, E. (2004a). Argonaute protein in the early divergent eukaryote *Trypanosoma brucei*: control of small interfering RNA accumulation and retroposon transcript abundance. Mol Cell Biol *24*, 420-427.

Shi, H., Ullu, E., and Tschudi, C. (2004b). Function of the trypanosome argonaute 1 protein in RNA interference requires the N-terminal RGG domain and arginine 735 in the piwi domain. J Biol Chem.

Shi, J., Blundell, T. L., and Mizuguchi, K. (2001). FUGUE: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties. J Mol Biol *310*, 243-257.

Sinclair, D. A., Mills, K., and Guarente, L. (1997). Accelerated aging and nucleolar fragmentation in yeast sgs1 mutants. Science *277*, 1313-1316.

Smith, J. L., Levin, J. R., and Agabian, N. (1989). Molecular characterization of the *Trypanosoma brucei* RNA polymerase I and III largest subunit genes. J Biol Chem *264*, 18091-18099.

Soares, M. J. (1999). The reservosome of *Trypanosoma cruzi* epimastigotes: an organelle of the endocytic pathway with a role on metacyclogenesis. Mem Inst Oswaldo Cruz *94*, 139-141.

Soskic, V., Gorlach, M., Poznanovic, S., Boehmer, F. D., and Godovac-Zimmermann, J. (1999). Functional proteomics analysis of signal transduction pathways of the platelet-derived growth factor beta receptor. Biochemistry *38*, 1757-1764.

Spector, D. L. (2003). The dynamics of chromosome organization and gene regulation. Annu Rev Biochem *72*, 573-608.

Stover, N. A., and Steele, R. E. (2001). Trans-spliced leader addition to mRNAs in a cnidarian. Proc Natl Acad Sci U S A *98*, 5693-5698.

Strickler, J. E., and Patton, C. L. (1975). Adenosine 3',5'-monophosphate in reproducing and diffe-rentiated trypanosomes. Science *190*, 1110-1112.

Stuart, K., and Panigrahi, A. K. (2002). RNA editing: complexity and complications. Mol Microbiol *45*, 591-596.

Sturm, N. R., Vargas, N. S., Westenberger, S. J., Zingales, B., and Campbell, D. A. (2003). Evidence for multiple hybrid groups in *Trypanosoma cruzi*. Int J Parasitol *33*, 269-279.

Sullivan, J. J. (1982). Metacyclogenesis of *Trypanosoma cruzi* in vitro: a simplified procedure. Trans R Soc Trop Med Hyg *76*, 300-303.

Sutton, R. E., and Boothroyd, J. C. (1986). Evidence for trans splicing in trypanosomes. Cell *47*, 527-535.

Tabara, H., Sarkissian, M., Kelly, W. G., Fleenor, J., Grishok, A., Timmons, L., Fire, A., and Mello, C. C. (1999). The rde-1 gene, RNA interference, and transposon silencing in *C. elegans*. Cell *99*, 123-132.

Tan, H., and Andrews, N. W. (2002). Don't bother to knock - the cell invasion strategy of *Trypanosoma cruzi*. Trends Parasitol *18*, 427-428.

Tarleton, R. L. (2003). Chagas disease: a role for autoimmunity? Trends Parasitol *19*, 447-451.

Tarleton, R. L., Zhang, L., and Downs, M. O. (1997). "Autoimmune rejection" of neonatal heart transplants in experimental Chagas disease is a parasite-specific response to infected host tissue. Proc Natl Acad Sci U S A *94*, 3932-3937.

Taylor, M. C., and Kelly, J. M. (2002). Functional dissection of the *Trypanosoma cruzi* genome: New approaches in a new era. In World Class Parasites. Vol. IV: American Trypanosomiasis, K. Tyler, Miles, M.A., eds. (New York, Kluwer Academic Publishers).

Taylor, M. C., Muhia, D. K., Baker, D. A., Mondragon, A., Schaap, P. B., and Kelly, J. M. (1999). *Trypanosoma cruzi* adenylyl cyclase is encoded by a complex multigene family. Mol Biochem Parasitol *104*, 205-217.

Teixeira, S. M. (1998). Control of gene expression in Trypanosomatidae. Braz J Med Biol Res *31*, 1503-1516.

Teixeira, S. M., Kirchhoff, L. V., and Donelson, J. E. (1995). Post-transcriptional elements regulating expression of mRNAs from the amastin/tuzin gene cluster of *Trypanosoma cruzi*. J Biol Chem *270*, 22586-22594.

Teixeira, S. M., Russell, D. G., Kirchhoff, L. V., and Donelson, J. E. (1994). A differentially expressed gene family encoding "amastin," a surface protein of *Trypanosoma cruzi* amastigotes. J Biol Chem *269*, 20509-20516.

Tomas, A. M., and Kelly, J. M. (1996). Stage-regulated expression of cruzipain, the major cysteine protease of *Trypanosoma cruzi* is independent of the level of RNA. Mol Biochem Parasitol *76*, 91-103.

Tomas, A. M., Miles, M. A., and Kelly, J. M. (1997). Overexpression of cruzipain, the major cysteine proteinase of *Trypanosoma cruzi*, is associated with enhanced metacyclogenesis. Eur J Biochem *244*, 596-603.

Tomas, A. M. L. R. (1995) Functional Analysis of the major cysteine protease of *Trypanosoma cruzi* using genetic approaches, PhD Thesis, Univerity of London, London.

235

Tovar, J., and Fairlamb, A. H. (1996). Extrachromosomal, homologous expression of trypanothione reductase and its complementary mRNA in *Trypanosoma cruzi*. Nucleic Acids Res *24*, 2942-2949.

Tran, A. N., Andersson, B., Pettersson, U., and Aslund, L. (2003). Trypanothione synthetase locus in *Trypanosoma cruzi* CL Brener strain shows an extensive allelic divergence. Acta Trop *87*, 269-278.

Tschudi, C., and Ullu, E. (1988). Polygene transcripts are precursors to calmodulin mRNAs in trypanosomes. EMBO J *7*, 455-463.

Tuerk, C., and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science *249*, 505-510.

Tyler, K. M., and Engman, D. M. (2000). Flagellar elongation induced by glucose limitation is preadaptive for *Trypanosoma cruzi* differentiation. Cell Motil Cytoskeleton *46*, 269-278.

Tyler, K. M., and Engman, D. M. (2001). The life cycle of *Trypanosoma cruzi* revisited. Int J Parasitol *31*, 472-481.

Ucros, H., Granger, B., and Krassner, S. M. (1983). *Trypanosoma cruzi*: effect of pH on *in vitro* formation of metacyclic trypomastigotes. Acta Trop *40*, 105-112.

Uliel, S., Liang, X. H., Unger, R., and Michaeli, S. (2004). Small nucleolar RNAs that guide modification in trypanosomatids: repertoire, targets, genome organisation, and unique functions. Int J Parasitol *34*, 445-454.

Ulloa, R. M., Mesri, E., Esteva, M., Torres, H. N., and Tellez-Inon, M. T. (1988). Cyclic AMP-dependent protein kinase activity in *Trypanosoma cruzi*. Biochem J *255*, 319-326.

Ullu, E., Tschudi, C., and Chakraborty, T. (2004). RNA interference in protozoan parasites. Cell Microbiol *6*, 509-519.

Urbina, J. A. (1994). Intermediary Metabolism of *Trypanosoma cruzi*. Parasitology Today *10*, 107-110.

van Deursen, F. J., Thornton, D. J., and Matthews, K. R. (2003). A reproducible protocol for analysis of the proteome of *Trypanosoma brucei* by 2-dimensional gel electrophoresis. Mol Biochem Parasitol *128*, 107-110.

Vandenberghe, A. E., Meedel, T. H., and Hastings, K. E. (2001). mRNA 5'-leader trans-splicing in the chordates. Genes Dev *15*, 294-303.

Vanhamme, L., Berberof, M., Le Ray, D., and Pays, E. (1995). Stimuli of differentiation regulate RNA elongation in the transcription units for the major stage-specific antigens of *Trypanosoma brucei*. Nucleic Acids Res *23*, 1862-1869.

Vanhamme, L., and Pays, E. (1995). Control of gene expression in trypanosomes. Microbiol Rev *59*, 223-240.

Vassella, E., Kramer, R., Turner, C. M., Wankell, M., Modes, C., van den Bogaard, M., and Boshart, M. (2001). Deletion of a novel protein kinase with PX and FYVE-related domains increases the rate of differentiation of *Trypanosoma brucei*. Mol Microbiol *41*, 33-46.

Vassella, E., Reuner, B., Yutzy, B., and Boshart, M. (1997). Differentiation of African trypanosomes is controlled by a density sensing mechanism which signals cell cycle arrest via the cAMP pathway. J Cell Sci *110*, 2661-2671.

Vazquez, M. P., and Levin, M. J. (1999). Functional analysis of the intergenic regions of TcP2beta gene loci allowed the construction of an improved *Trypanosoma cruzi* expression vector. Gene *239*, 217-225.

Vercesi, A. E., Moreno, S. N., and Docampo, R. (1994). Ca^{2+}/H^+ exchange in acidic vacuoles of *Trypanosoma brucei*. Biochem J *304 (Pt 1)*, 227-233.

Vickerman, K. (1985). Developmental cycles and biology of pathogenic trypanosomes. Br Med Bull *41*, 105-114.

Volpe, T. A., Kidner, C., Hall, I. M., Teng, G., Grewal, S. I., and Martienssen, R. A. (2002). Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science *297*, 1833-1837.

Vullo, A., and Frasconi, P. (2004). Disulfide connectivity prediction using recursive neural networks and evolutionary information. Bioinformatics *20*, 653-659.

Wang, Z., Morris, J. C., Drew, M. E., and Englund, P. T. (2000). Inhibition of *Trypanosoma brucei* gene expression by RNA interference using an integratable vector with opposing T7 promoters. J Biol Chem *275*, 40174-40179.

Weeks, G. (2000). Signalling molecules involved in cellular differentiation during *Dictyostelium* morphogenesis. Curr Opin Microbiol *3*, 625-630.

Welch, W. J., and Suhan, J. P. (1985). Morphological study of the mammalian stress response: characterization of changes in cytoplasmic organelles, cytoskeleton, and nucleoli, and appearance of intranuclear actin filaments in rat fibroblasts after heat-shock treatment. J Cell Biol *101*, 1198-1211.

WHO (2003). TDR Disease Watch, Chagas Disease. Nature Reviews Microbiology *1*, 14-15.

Wilkinson, S. R., Obado, S. O., Mauricio, I. L., and Kelly, J. M. (2002). *Trypanosoma cruzi* expresses a plant-like ascorbate-dependent hemoperoxidase localized to the endoplasmic reticulum. Proc Natl Acad Sci U S A *99*, 13453-13458.

Wirtz, E., and Clayton, C. (1995). Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. Science *268*, 1179-1183.

Wirtz, E., Hartmann, C., and Clayton, C. (1994). Gene expression mediated by bacteriophage T3 and T7 RNA polymerases in transgenic trypanosomes. Nucleic Acids Res *22*, 3887-3894.

Wirtz, E., Hoek, M., and Cross, G. A. (1998). Regulated processive transcription of chromatin by T7 RNA polymerase in *Trypanosoma brucei*. Nucleic Acids Res *26*, 4626-4634.

Wyllie, A. H., Beattie, G. J., and Hargreaves, A. D. (1981). Chromatin changes in apoptosis. Histochem J *13*, 681-692.

Yamada-Ogatta, S. F., Motta, M. C., Toma, H. K., Monteiro-Goes, V., Avila, A. R., Muniz, B. D., Nakamura, C., Fragoso, S. P., Goldenberg, S., and Krieger, M. A. (2004). *Trypanosoma cruzi*: cloning and characterization of two genes whose expression is up-regulated in metacyclic trypomastigotes. Acta Trop *90*, 171-179.

Zhang, J., Ruyechan, W., and Williams, N. (1998). Developmental regulation of two nuclear RNA binding proteins, p34 and p37, from *Trypanosoma brucei*. Mol Biochem Parasitol *92*, 79-88.

Zhang, J., and Williams, N. (1997). Purification, cloning, and expression of two closely related *Trypanosoma brucei* nucleic acid binding proteins. Mol Biochem Parasitol *87*, 145-158.

Zhou, H., Watts, J. D., and Aebersold, R. (2001). A systematic approach to the analysis of protein phosphorylation. Nat Biotechnol *19*, 375-378.

Ziegelbauer, K., Quinten, M., Schwarz, H., Pearson, T. W., and Overath, P. (1990). Synchronous differentiation of *Trypanosoma brucei* from bloodstream to procyclic forms *in vitro*. Eur J Biochem *192*, 373-378.

Zingales, B., Pereira, M. E., Almeida, K. A., Umezawa, E. S., Nehme, N. S., Oliveira, R. P., Macedo, A., and Souto, R. P. (1997a). Biological parameters and molecular markers of clone CL Brener - the reference organism of the *Trypanosoma cruzi* genome project. Mem Inst Oswaldo Cruz *92*, 811-814.

Zingales, B., Pereira, M. E., Oliveira, R. P., Almeida, K. A., Umezawa, E. S., Souto, R. P., Vargas, N., Cano, M. I., da Silveira, J. F., Nehme, N. S., *et al.* (1997b). *Trypanosoma cruzi* genome project: biological characteristics and molecular typing of clone CL Brener. Acta Trop *68*, 159-173.

Zingales, B., Rondinelli, E., Degrave, W., Franco da Silveira, J., Levin, M., Le Paslier, D., Modabber, F., Dobrokhotov, B., Swindle, J., Kelly, J. M., *et al.* (1997c). The *Trypanosoma cruzi* Genome Initiative. Parasitology Today *13*, 16-22.

Zomerdijk, J. C., Kieft, R., and Borst, P. (1991a). Efficient production of functional mRNA mediated by RNA polymerase I in *Trypanosoma brucei*. Nature *353*, 772-775.

Zomerdijk, J. C., Kieft, R., Shiels, P. G., and Borst, P. (1991b). Alpha-amanitin-resistant transcription units in trypanosomes: a comparison of promoter sequences for a VSG gene expression site and for the ribosomal RNA genes. Nucleic Acids Res *19*, 5153-5158.

Zomerdijk, J. C., Ouellette, M., ten Asbroek, A. L., Kieft, R., Bommer, A. M., Clayton, C. E., and Borst, P. (1990). The promoter for a variant surface glycoprotein gene expression site in *Trypanosoma brucei*. EMBO J *9*, 2791-2801.

Zoraghi, R., Kunz, S., Gong, K., and Seebeck, T. (2001). Characterization of TbPDE2A, a novel cyclic nucleotide-specific phosphodiesterase from the protozoan parasite *Trypanosoma brucei*. J Biol Chem *276*, 11559-11566.

Zufall, F., Firestein, S., and Shepherd, G. M. (1994). Cyclic nucleotide-gated ion channels and sensory transduction in olfactory receptor neurons. Annu Rev Biophys Biomol Struct *23*, 577-607.