Cyclin dependent kinases and cell cycle control in *Trypanosoma brucei*

Jack Ragnar Ford

Wellcome Unit of Molecular Parasitology

Anderson College

University of Glasgow

Glasgow

G11 6NU

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

November 1999

ABSTRACT

In all eukaryotic organisms, from yeast to mammals, progression through the cell cycle is ultimately regulated by the actions of a group of highly conserved serine/threonine protein kinases named the cyclin dependent kinases (CDKs). CDK activity is controlled by the association of CDK subunits with regulatory protein subunits named cyclins to form active CDK-cyclin complexes. The trypanosomatids represent one of the earliest known divergent branches of the eukaryotic phylogenetic tree. At the start of this project several putative CDK and cyclin homologues had been described in *Trypanosoma brucei*, *tbCRK1-3* and *tbCYC1*. During the project *tbCRK2* and *tbCYC1* were analysed and a new *CRK* (*tbCRK4*) isolated.

The function of the tbCRK2 protein and the *tbCRK2* gene was investigated using biochemical and molecular genetic approaches. Use of anti-CRK2 antisera showed that tbCRK2 is expressed in both proliferative and nonproliferative life cycle stages of *T. brucei*. Attempts to create null mutants at the *tbCRK2* locus repeatedly failed, with the wild-type allele preserved in all transgenic lines generated. Ectopic expression of a cloned copy of *tbCRK2* did not allow deletion of wild-type alleles, although the cloned copy of *tbCRK2* was demonstrated to generate abnormally high cellular levels of tbCRK2. These results may suggest that *tbCRK2* is an essential gene in the procyclic form of the parasite; however, control experiments are needed to verify this.

The tbCYC1 protein was previously reported to be a mitotic-type cyclin in *T. brucei*. Biochemical and molecular genetic approaches were taken to investigate tbCYC1 function. A specific anti-CYC1 antiserum was generated, which demonstrated that in *T. brucei* strain STIB 247 tbCYC1 is expressed in the short stumpy bloodstream and procyclic forms, but not in the long slender bloodstream form, suggesting a function in differentiation rather than cell cycle control *per se*. No significant tbCYC1-associated kinase activity was found in procyclic extracts. In addition, no association of tbCYC1 with any of the known *T. brucei* CRKs or with the yeast protein $p13^{SUC1}$ or the leishmanial protein $p12^{CKS1}$ was found in procyclic extracts. These data suggests that tbCYC1 does not fulfil the function of a mitotic-type cyclin. Attempts to generate null mutants for the *tbCYC1* locus failed, which may provide preliminary evidence that *tbCYC1* serves an essential function in the procyclic form. However, further experiments are required to clarify this.

A full-length genomic clone of the *tbCRK4* gene was isolated and demonstrated to be homologous to the *Crithidia fasciculata cfCRK* gene. Computer analysis revealed the presence of two large, highly hydrophilic insert domains between the conserved kinase catalysis domains VIb and VII, and X and XI. Specific anti-CRK4 antisera were generated and used to demonstrate that in strain STIB 247, tbCRK4 is expressed in the short stumpy bloodstream and

procyclic forms, but not in the long bloodstream form. No tbCRK4-associated kinase activity was detected from procyclic extracts, and no interactions with the yeast protein $p13^{SUC1}$ or the leishmanial protein $p12^{CKS1}$ was found from procyclic extracts.

This study has provided evidence that mechanisms of cell cycle control in trypanosomatids may be based upon the functional CDK-cyclin paradigm observed in higher eukaryotes yet have diverged to serve the parasite's unique biology. If sufficient divergence in terms of biochemical function between trypanosome CRKs and cyclins and the CDKs and cyclin of mammals could be demonstrated, they would provide a tempting drug target.

CHAPTER 4: INVESTIGATION OF GENETIC AND BIOCHEMICAL FUNCTION OF THE *tbCYC1* GENE AND tbCYC1 PROTEIN

CHAPTER 5: ANALYSIS OF *tbCRK2* GENE AND tbCRK2 PROTEIN FUNCTION

CHAPTER 6: GENERAL DISCUSSION

LIST OF FIGURES

CHAPTER 1

CHAPTER 3

CHAPTER 4

CHAPTER 5

LIST OF ABBREVIATIONS

ATP	:	Adenosine triphosphate
BSA	:	Bovine serum albumin
CDK	:	Cyclin-dependent kinase
cDNA	:	Complementary DNA
CIP	:	Calf intestinal phosphatase
CRK	:	*CDC2*-related kinase
dATP	:	Deoxyadenosine triphosphate
dCTP	:	Deoxycytidine triphosphate
DEPC	:	Diethyl pyrocarbonate
dGTP	:	Deoxyguanosine triphosphate
DMSO	:	Dimethyl sulphoxide
dNTP	:	Deoxyribonucleotide triphosphate
DTT	:	Dithiothretiol
ECL	:	Enhanced chemiluminescence
EDTA	:	Ethylenediaminetetraacetic acid
EGTA	:	Ethylene-bis[oxyethylenenitrilo]tetraacetic acid
FBS	:	Foetal bovine serum
FSB	:	Four times SDS-PAGE sample buffer
GEB	:	Glutathione elution buffer
GLB	:	DNA gel loading buffer
GS4B	:	Glutathione-Sepharose 4B
GST	:	Glutathione-S-transferase
HRP	:	Horseradish peroxidase
HSLS	:	High salt lysis solution
IgG	:	Immunoglobulin G
IPTG	:	Isopropyl-β-[D]-thiogalactoside
KAB	:	Kinase assay buffer
KAM	:	Kinase assay mix
LB	:	Luria-Bertani medium
LS	:	Lysis solution
LSG	:	Lysis solution with glycerol
LSI	:	Lysis solution with protease inhibitors

LSLS	:	Low salt lysis solution
LS-T	:	Lysis solution without TritonX-100
MOPS	:	3-[N-Morpholino]propanesulfonic acid
NaOV	:	Sodium orthovanadate
Ni-NTA	:	Nickel-charged nitrilo-tri-acetic Sepharose resin
PBS	:	Phosphate buffered saline
PCR	:	Polymerase chain reaction
Pfu	:	*Pyrococcus furiosus* DNA polymerase
PSG	:	Phosphate saline glucose buffer
PVDF	:	Polyvinylidine difluoride
SB	:	Sonication buffer
SBCRK	:	SUC-binding CRK
SDM-79	:	Semi-defined medium 79
SDS	:	Sodium dodecyl sulphate
SDS-PAGE	:	SDS-polyacrylaminde gel electrophoresis
SPRB	:	SDS-PAGE gel running buffer
SSC	:	Salt sodium tricitrate buffer
STIB	:	Swiss Tropical Institue of Biology
TAE	:	TRIS acetate EDTA buffer
Taq	:	*Thermus aquaticus* DNA polymerase
TBE	:	TRIS borate EDTA buffer
TBST	:	TRIS buffered saline with Tween-20
TE	:	TRIS EDTA buffer
TREU	:	Trypanosomiasis Research Edinburgh University
TRIS	:	Tris[hydroxymethyl]aminomethane
TritonX-100	:	t-Octylphenoxypolyethoxyethanol
TSB	:	Three times SDS-PAGE sample buffer
TTE	:	TRIS taurine EDTA buffer
Tween-20	:	Polyoxyethylene sorbitan monolaureate
WUMP	:	Wellcome unit of Molecular Parasitology
X-Gal	:	5-bromo-4-chloro-3-indolyl-β-[D]-galactoside
ZPFM	:	Zimmermans post fusion medium

NOTE ON GENETIC NOMENCLATURE

The new standard for genetic nomenclature as applied to trypanosomatid parasites has been revised recently (Clayton *et al.*, 1998). The nomenclature used in this thesis shall broadly follow those guidelines. The basic principles are as follows, using the *CRK3* gene as an example. Wild-type alleles are italicised in upper case, with lower case prefix denoting species e.g. CRK3, *tbCRK3* denoting the *T. brucei* gene, *lmmCRK3* denoting the *Leishmania mexicana* gene. Multiple related but non-identical genes are distinguished by non-hyphenated letters or numbers e.g. *PKCα*, *PKCδ*, *THT1*, *THT2*; RNA products have the exact same designations. Protein products have the same designations but are not italicised (I have not used italicisation for the species prefix with protein products).

Gene replacements are indicated with a double colon. For example, replacement of one allele of *tbCRK3* with a bleomycin marker *BLE* is denoted by *Δcrk3::BLE/CRK3*. Insertional inactivations are indicated by a circumflex (^) e.g. *^crk3::BLE* for the insertional inactivation of BLE to one allele of *tbCRK3*.

Properly, the system of genetic nomenclature differs between the two species of yeast mentioned in this thesis. *Schizosaccharomyces pombe* genes are italicised in lower case with a plus sign for wild-type allele, and a minus sign for an inactive mutant allele or deleted wild-type allele e.g. *cdc2*$^{+}$, *cdc2*$^{-}$. Protein products are lower case but non-italicised. *Saccharomyces cerevisiae* genes are italicised in upper case for wild-type alleles and lower case for mutant alleles e.g. *CDC28* and *cdc28*, except for dominant mutants which are italicised in upper case e.g. *DAF1-1*.

However, in the interests of continuity I have applied the trypanosomatid system to the yeast genes: This is not the consensually accepted nomenclature for yeast genes. However, as the sections on *S. pombe* and *S. cerevisiae* cell cycle are separate, confusion should be minimised. Outwith these sections (chapter 1.2.1 and 1.2.2) if a yeast gene or protein is mentioned, species is specified. It should be noted that although the CDC system of naming yeast genes is used for both *S. pombe* and *S. cerevisiae*, the two do not correspond i.e. the same number of CDC gene between the two species designates a completely different gene. For genes homologous between the two species that are designated by different numbers/names, the relationship has been emphasised in the text to clarify the relationship and avoid confusion.

LIST OF OLIGONUCLEOTIDES

Primer	Gene/vector	Sequence 5'-3'	Purpose
CRK2N5'	*CRK2*	GCGCATGCGGGAGTGTTGCCCGCAAG	PCR
CRK2N3'	*CRK2*	GCTCTAGACCGTCGTTTACGATAAGC	PCR
CRK2C5'	*CRK2*	CGGGTACCCGCGGGGGATACCGCCAT	PCR
CRK2C3'	*CRK2*	GCGAATTCCCACCTCCTTTCCCCTAAC	PCR
SAT5'	*SAT*	GCAAGCTTATGAAGATTTCGGTGATC	PCR
SAT3'	*SAT*	GCCTGCAGTTAGGCGTCATCCTGTGCT	PCR
OL88	pHD675	TAGGGGTTATCGGGTAGGGATC	Sequencing, PCR
CYCKO55	*CYC1*	GCATGCATGACAAACTTGAATGTGC	PCR
CYCKO53	*CYC1*	TCTAGACTCCCAGCGTCGTATAGCG	PCR
CYCKO35	*CYC1*	GGTACCAGGTGCTGACGAT	PCR
CYCKO33	*CYC1*	GAATTCTCACCTGGGTGCACATAGT	PCR
D9	pQE series	GCGGATAACAATTTCACACAG	Sequencing
OL355	*CRK4*	CGCGGATCCGCTTGGCGAAGGAACCTATGGAG	PCR, sequencing
OL356	*CRK4*	CGCCCCGGGGCGAAGATACAGCCAACCGACCAC	PCR, sequencing
OL348	*CRK4*	GCATCAGCGGAGTGTCGTGCAT	Sequencing
OL349	*CRK4*	GCCCGTATAGAGAGTATCCCAG	Sequencing
OL350	*CRK4*	GTCGTCCCAGTGCCCTCGCG	Sequencing
OL351	*CRK4*	AGAGTCGCACGATATTATTATGC	Sequencing
OL89	*CRK4*	GGAATCTCCCTAATGTGCAAA	Sequencing
OL90	*CRK4*	TTTCCATTTTCTGCAGGTTCCC	Sequencing
OL91	*CRK4*	GTTGAAGGTACTGATTTTCTACG	Sequencing
OL92	*CRK4*	CTCACTGCCGACGATAATGCG	Sequencing
OL143	*CRK4*	GGCAGCATTCATCGCATTTAGCG	Sequencing
OL193	*CRK4*	CATAGAACGTTGGCGAAGTGAGC	Sequencing
OL144	*CRK4*	CATACGAAGAGGGTCAGGAGGATG	Sequencing
OL206	*CRK4*	ACCAGGCGATATTCCCGTTGTT	Sequencing
pGEX5'	pGEX-5X-2	CCGGGAGCTGCATGTGTCAGAGG	Sequencing

LIST OF TRANSGENIC *T. BRUCEI* CELL LINES

WUMP	Parent strain	Genomic arrangement
563	*T. brucei* STIB 247	*Δcrk2::BLE/CRK2*
564	*T. brucei* STIB 247	*Δcrk2::BLE/CRK2*
827	*T. brucei* AnTat 1.1	*CRK2*/pHD430-*CRK2*
833	*T. brucei* STIB 247	*Δcyc1::PAC/CYC1*
834	*T. brucei* STIB 247	*Δcyc1::PAC/CYC1*
835	*T. brucei* AnTat 1.1	*Δcrk2::SAT/CRK2*/pHD430-*CRK2*
836	*T. brucei* AnTat 1.1	*Δcrk2::SAT/CRK2*/pHD430-*CRK2*
837	*T. brucei* STIB 247	*Δcrk2::BLE/Δcrk2::PAC/CRK2*
838	*T. brucei* STIB 247	*Δcrk2::BLE/CRK2*
839	*T. brucei* STIB 247	*Δcrk2::BLE/Δcrk2::PAC/CRK2*
840	*T. brucei* STIB 247	*Δcrk2::BLE/Δcrk2::PAC/CRK2*
877	*T. brucei* STIB 247	*Δcrk2::BLE/Δcrk2::PAC/Δcrk2::SAT/CRK2*?
878	*T. brucei* STIB 247	*Δcyc1::PAC/Δcyc1::BLE/CYC1*?
878	*T. brucei* STIB 247	*Δcyc1::PAC/Δcyc1::BLE/CYC1*?
880	*T. brucei* STIB 247	*Δcyc1::PAC/Δcyc1::BLE/CYC1*?
1163	*T. brucei* Mitat 1	*CRK2*/pHD675-*CRK2His*

ACKNOWLEDGEMENTS

First and foremost, a !thank you! the size of a small planet to my supervisor, Dr. Jeremy Mottram, for the kind of support and patience that would make a saint shudder. Your halo's in the post.

Colleagues through the ages at WUMP (now WCMP) for laughs, advice, chats and the occasional marathon whinge sessions that make it all okay. In particular: Karen Grant, for much invaluable advice on protein techniques, and for taking an even stroppier approach to life than me. Amanda McAie, for things of a molecular biological nature and for periodically creeping up behind me and scaring the bejeezus out of me by screaming, '*RA-RA*!!'. Phil Halford, for technical support in the early years. Darren Brooks, just for letting me know that I'm not alone in my extreme hatred of Vector NTI. Oh, and all those who made working in the issue culture room !such! a joy...

Friends, for the good times, rampageous nights out (and in) and for having the sense to back off when the mere mention of the word 'thesis' provoked only a scream of rage: Kev, Kate, Tara, Fi, Scott, John, Claire, David, Emma T, Emma B, Pete, Andy, Kara.

Finally, my unbelievably long-suffering parents: Breathe out, guys, it's all over.

DECLARATION

I hereby declare that the contents of this thesis are my own work, unless otherwise stated with due acknowledgement.

Jack Ragnar Ford,
November 1999.

CHAPTER 1: INTRODUCTION

1.1: General biology of African trypanosomes

Protozoa were once regarded as a phylum within the kingdom Anamilia, although it is now believed that this classification is not a natural one and that it contains members that could be classified as animals, plants or fungi. Accordingly, the concept of Protozoa as a taxon has been abandoned and the term is applied to animal-like members of the kingdom Protista. From a historical perspective classification of such single-celled organisms has been problematic, based upon a wide and complex range of features such as morphology, cell ultrastructure, variations in life cycle, and increasingly on biochemical and molecular differences. There continues to be arguments about the relationship between various taxonomic groups within the kingdom Protista, many of which will doubtless be clarified by advances in molecular systematics and increased genetic information (Coombs *et al.*, 1998). The protozoa have traditionally been divided into four groups, based on mode of locomotion: The flagellates, which move by the use of one or more flagella, the amoeba, which move by pseudopodia, the ciliates, moving by the co-ordinated beating of hundred of cilia, and the sporozoans, which lack any obvious form of locomotion (Sleigh, 1989). Within the flagellates group, phylum Kinetoplastida includes order Trypanosomatida: *Leishmania* and *Trypanosoma* species (Cox, 1992, 1993).

The trypanosomatids are characterised by the presence of the kinetoplast, a unique organelle that is an integral part of the mitochondrion and which contains DNA in the form of maxicircles and minicircles. The kinetoplast is located near the base of the flagellum, the exact location of which changes with respect to the rest of the cell in different life cycle stages and is associated with changes in the mitochondrial system. The life cycle of *T. b. brucei* has been described (Vickerman, 1985). The cycle is complex, with alternating proliferative and cell cycle arrested forms (Fig. 1.1), and as the host environments of mammals and tsetse flies are very different the patterns of gene expression between forms vary. The cell cycle arrested short stumpy form is ingested when a tsetse fly takes a bloodmeal from an infected mammal (*Glossina* species; *T. b. brucei* is transmitted by many *Glossina* species, *T. b. rhodesiense* mainly by the *G. morsitans* group, *T. b. rhodesiense* by *G. palpalis* and *G. tachinoides*). The stumpy form enters the midgut and transforms to the procyclic, which is a proliferative form that establishes the infection in the fly. The tsetse itself has a lectin defence mechanism (Maudlin and Welburn, 1987); susceptibility to trypanosome infection is maternally inherited and associated with the presence of a rickettsia-like symbiont in the midgut cells (Maudlin and Ellis, 1985). The symbiont produces a chitinase activity that releases D+ glucosamine, which may interfere with lectin binding, and the chitinase itself may render the peritrophic membrane (PTM) more

penetrable. The procyclic form penetrates the PTM, reaches the ectoperitrophic space and migrates to the proventriculus, where it differentiates to the elongate cell cycle arrested mesocyclic form. Mesocyclics retraverse the PTM and migrate via the oesophagus, proboscis lumen and hypopharynx to the salivary glands (Vickerman *et al.*, 1988). Here, they differentiate to the proliferative epimastigote form, which attaches to the epithelia and gland cell microvilli via branched outgrowths of the flagellum. Differentiation to the cell cycle arrested metacyclic form is accompanied by regression of the flagellar outgrowths, although attachment is retained until cell cycle arrest and acquisition of the VSG coat, leading to detachment as the mature, cell cycle arrest metacyclic (Vickerman *et al.*, 1988). Determination of the VSG expressed appears to occur at the last division of the nascent metacyclic (Tetley *et al.*, 1987). The part of the cycle in the fly takes 3-5 weeks, and in many cases the infection aborts, with only a small percentage of tsetse flies eventually producing metacyclics. Metacyclics are injected to the mammalian vascular system upon tsetse feeding, where they differentiate to the proliferative long slender bloodstream form. Parasitaemia increases until differentiation to the cell cycle arrested stumpy form occurs. This differentiation step is thought to be initiated by a low molecular weight molecule, stumpy inducing factor (SIF; Vassella *et al.*, 1997); stumpy form cells are ingested by the tsetse again and the cycle returns to the start.

The basic metabolism of the parasite is therefore quite distinctive between the forms in the two hosts and there is a requirement for rapid switching between these modes when transmission between species occurs. In addition, metabolic processes may vary between different forms of the parasite within the same host, for example between the proliferative procyclic form in the tsetse midgut and the mature metacyclic form, preadapted for transmission to a mammal, in the salivary glands. One of the most striking examples of this is the difference in energy metabolism between the long slender bloodstream form and the procyclic form.

In the long slender bloodstream form the mitochondrion has a regressed form, morphologically appearing as a simple tube with the components of the TCA (tricarboxylic acid) cycle missing, and energy generation being achieved by the process of 'aerobic glycolysis'. The mitochondrion interacts with an organelle called the glycosome (Clayton and Michels, 1996): Glucose is imported to the cell from the mammalian bloodstream by specific glucose transporters (Barrett *et al.*, 1998) and then enters the glycosome. Here it is metabolised to 1,3-diphosphoglycerate and glycerol-3-phosphate, the former of which leaves the glycosome and enters the cytosol where it is converted to pyruvate, the latter of which is transported to the mitochondrion where it is oxidised to dihydroxyacetone phosphate. This then re-enters the glycosome where its electrons are transferred by glycerophosphate oxidase to oxygen; it is unclear whether this step is coupled to the synthesis of ATP but the net effect is to ensure continual reoxidation of NADH. This rapid oxidation of glucose and production of ATP is inefficient, generating only two moles of ATP per mole of glucose utilised, but this is irrelevant

given the high availability of glucose from the host's bloodstream and therefore inefficiency is compensated for by the high throughput of glucose carbon (Cox, 1993).

As the parasite differentiates through the intermediate form to the cell cycle arrested short stumpy form the mitochondrion becomes more complex, developing branches and cristae. It gains the functions of acetate and succinate production from pyruvate, fumarate reductase is present and so are the enzymes of the TCA cycle although the cycle itself does not function normally. The glycosome gains phosphoenolpyruvate carboxykinase and malate dehydrogenase activity, the latter assuming the role of NADH reoxidation as glycolytic activity wanes (Durieux *et al.*, 1991; Priest and Hadjuk, 1994). These changes are geared towards preadapting the parasite for survival in the insect host. The short stumpy form arrives in the tsetse fly midgut, a highly oxidising environment, and subsequently transforms to a procyclic trypomastigote and migrates via the insect's haemolymph to the salivary glands. Glucose ingested in the bloodmeal is very rapidly oxidised in the midgut and the muscles powering the insects' wings rely on oxidation of amino acids, chiefly proline, as the energy source. Consequently, the parasite must adapt quickly to an amino acid based energy generation. In transforming to the procyclic form the partially developed mitochondrion branches further and acquires large numbers of cristae and a fully functional TCA cycle, and is therefore able to at least partially oxidise the proline and other amino acids that it takes up from the insect's haemolymph. In both life cycle stages - indeed in all life cycle stages of *T. b. brucei* - although the specific metabolic pathways may differ they are regulated in general in the same manner as in other organisms, responding to the adenylate energy charge and the redox state, enzymes showing allosteric regulation, metabolic paths subject to both feedback and feedforward control and to product inhibition.

1.2.1: Fission yeast cell cycle

The eukaryotic cell cycle is a complex and highly ordered process by which a cell replicates its DNA and ensures proper segregation of replicated DNA between its daughter cells. The cell cycle is divided into four phases; S-phase, in which the DNA content of the cell doubles, mitosis (M phase) in which DNA is segregated between the forming daughter cells, and two gap phases that separate these other phases, called G1 and G2. The normal progression is therefore that a newly formed daughter cell passes through G1, growing and preparing for S phase in which chromosomes are replicated. Following S phase, the cell moves through G2 prior to entry to M phase and the formation of two new daughter cells. Understanding of the metazoan cell cycle has to a large extent been underpinned by the deconstruction of the cell cycle of two species of yeast, providing a template of eukaryotic cell cycle control that could then be used as a model for elucidating mechanisms of cell cycle control in mammalian cells. The species of

genetically tractable, easily cultured yeast used in these cell cycle analyses are fission yeast, *Schizosaccharomyces pombe*, and budding yeast, *Saccharomyces cerevisiae*. These yeasts are ideal tools for the dissection of cell cycle control as they have small genomes, only 4-5 times larger than the genome of *E. coli*, are easily cultured and have short mean generation times, allowing rapid isolation of mutants altered in the normal cell cycle. Both propagate with stable haploid genomes, although both may conjugate and subsequently either sporulate, undergoing meiosis to produce daughter cells with new haploid genomic arrangements, or propagate as a diploid (Forsburg and Nurse, 1991). Both tend to mate in response to the stress of nutrient limitation.

Schizosaccharomyces pombe is a rod-shaped single-celled organism of constant diameter that grows by extension of the cell ends and which divides by septation and medial binary fission. Microtubule reorganisation and formation of the mitotic spindle occur in M phase, as in other eukaryotic cells (Hagan and Hyams, 1988), but in contrast - and in common with *Saccharomyces cerevisiae* - the nuclear envelope remains intact. The emergent daughter cells are already big enough to support DNA replication in rapidly growing *Schizosaccharomyces pombe*; therefore S phase commences before cytokinesis is complete, with a correspondingly short G1. Cell cycle regulation is chiefly at the G1/S transition where cell size and nutritional status are monitored, thus if nutrient limitations are a factor G1 is extended before the decision to enter S phase is taken (Nasmyth 1979; Nasmyth *et al.*, 1979). The decision point for a cell to replicate its DNA and complete mitosis is called START, defined as the point in G1 in which the cell becomes committed to the cell cycle. *Schizosaccharomyces pombe* grows mitotically until nutrients are limiting when one of two alternate paths is followed. The cells can arrest in G1 or G2 (reviewed by Costello *et al.*, 1986), although they do not enter a state analogous to the G0 state of mammalian cells, or they can conjugate to form a diploid which then sporulates and undergoes meiosis to produce haploid daughter cells. The diploid state of *S. pombe* is highly unstable, although it can be maintained by culturing in rich media.

Delineation of the mechanisms of cell cycle control in fission yeast has largely been reliant on isolation of temperature-sensitive mutants. These are clones which display normal growth and division at the permissive temperature of 25°C but which display an abnormal phenotype with respect to the cell cycle when shifted to the restrictive temperature of 37°C. The first such identified were highly elongated mutants relative to the normal cell shape, therefore cell cycle progression had been blocked but cell growth was unaffected, which could be confirmed by analysis of the DNA content of the mutants. The genes that produced the mutant phenotypes could then be isolated by screening of genomic or cDNA libraries for plasmids containing *S. pombe* DNA that could rescue the phenotype. Genes isolated in this manner were termed *CDC* genes, for cell division cycle, and by this method 26 separate *CDC* mutants were characterised (Hartwell, 1978).

One gene isolated by this approach, and subsequently to be shown to be a key regulator of cell cycle progression, was the *CDC2* gene. Absence of *CDC2* function leads to cell cycle arrest at both the G1/S and G2/M transitions, revealing a role for this gene at both decision points (Nurse and Bisset, 1981). Analysis of the gene sequence proved it to be a serine/threonine protein kinase, although it did not fall into any of the existing kinase families. It is now known that CDC2 is a member of a distinct family of serine/threonine kinases, called the cyclin dependent kinase (CDK) family. CDKs depend on binding to proteins called cyclins for activation and direction of function; cyclins were initially isolated in biochemical analyses of protein oscillation through the cell cycle in sea urchin embryos (Evans *et al.*, 1983). One protein was identified which was present only at very low levels in G1 and S phase, but which showed an increase in abundance through G2 that peaked at M phase. The protein was rapidly destroyed after M phase was underway, and began its gradual accumulation after M phase as daughter cells re-entered G1 (Evans *et al.*, 1983). In *S. pombe*, a gene was isolated that fulfilled this cyclin function. The previously identified *CDC13* gene was discovered to have three temperature-sensitive alleles that acted as suppressors of a cold-sensitive *CDC2* allele (Nurse *et al.*, 1976; Nasmyth and Nurse, 1981); in addition, overexpression of *CDC2* could rescue a *CDC13* temperature-sensitive allele. Deletion of *CDC13* leads to an M phase bypass, with multiple rounds of S phase with no intervening M phase (Hayles *et al.*, 1994; Fisher and Nurse, 1996). The temperature-sensitive allele mentioned above, *CDC13-117* (a conditional point mutant) arrests cells at the restrictive temperature with a mixed G2/M phenotype in which the chromosomes are condensed with an interphase array of microtubules (Nurse *et al.*, 1976; Nasmyth and Nurse, 1981; Hagan *et al.*, 1988), indicating functions for CDC13 within M phase. Binding of CDC13 to CDC2 and subsequent destruction of CDC13 at anaphase are therefore the key regulatory events (Moreno *et al.*, 1989). CDC2-CDC13 binding is stochiometric, but it is not just the ratio of CDC2 to CDC13 subunits that regulate binding; phosphorylation of Thr167 (a residue conserved in an equivalent position in all CDKs) is necessary for cyclin binding (Ducommun *et al.*, 1991). Phosphorylation of Thr167 is performed by CAK (CDC2-activating kinase). CAK has been purified from *Xenopus laevis* oocytes and found to be the previously identified $p40^{MO15}$ protein (Fesquet *et al.*, 1993; Poon *et al.*, 1993). Furthermore, a dephosphorylation event is necessary subsequent to cyclin binding for activity of the complex. The phosphorylated Tyr15 residue of CDC2 that inhibits CDC2-cyclin complexes (Lee *et al.*, 1988; Gould and Nurse, 1989) is dephosphorylated by the CDC25 protein phosphatase. CDC25 mRNA and protein levels increase through the cell cycle (Ducommun *et al.*, 1990), peaking at G2, and dephosphorylation of Tyr15 occurs just before M phase, activating the complex (Draetta and Beach, 1988; Morla *et al.*, 1989; Millar *et al.*, 1991) and driving the cell into mitosis. A mutation in *CDC25* leads to a G2 arrest (Nurse *et al.*, 1976) with an elongated phenotype; overexpression of *CDC25* drives the cell prematurely into M phase (Fantes, 1979;

Russell and Nurse, 1986) thus *CDC25* acts as a dosage-dependent inducer of mitosis. It therefore defines a rate-limiting step in G2/M transition.

Further genetic analysis for rate-limiting genes identified mutants in which the cells moved through the cell cycle with a reduced size, with a truncated or 'wee' phenotype. The wee phenotype maps to two loci, the *CDC2* gene itself and *WEE1*, a gene that sequence analysis predicts to be a serine/threonine kinase. *WEE1* was isolated by complementation of a strain with the wee phenotype (Russell and Nurse, 1987a). While a mutation in *WEE1* gives the phenotype of cell division at reduced size, overexpression of *WEE1* results in abnormally elongated cells; therefore *WEE1* acts as a dosage-dependent inhibitor of mitosis (Russell and Nurse, 1987a) and is directly opposite in effect to *CDC25*. Although by sequence analysis WEE1 is predicted to be a serine/threonine kinase, it has been demonstrated to inhibit CDC2 by phosphorylating Tyr15 (McGowan and Russell, 1993). Cells with a *WEE1* mutation are too small to pass START after M phase and therefore have an extended G1. Overexpression of an allele of *WEE1* with a mutation suppresses the elongated, G2 arrested phenotype of a *CDC25ts* mutant and restores growth (Fantes, 1979; Russell and Nurse, 1986). Evidence that the effect of the genes is additive comes from overexpression of *CDC25* in a *WEE1*-deficient background that leads to an acceleration of mitosis to such a degree that the cells attempt to divide at too small a size (mitotic catastrophe). Additional evidence supporting the view of a bifurcated upstream control of CDC2 ending in the antagonistic effects of CDC25 and WEE1 comes from the wee phenotype that maps to CDC2. These mutants fall into two classes and map to different portions of the CDC2 protein (Nurse and Thuriaux, 1980; Carr *et al.*, 1989); cells that are independent of WEE1 function and cells that are independent of CDC25 function. A regulator of WEE1 activity, *NIM1*, was isolated (Russell and Nurse, 1987b) and later shown to be allelic with a gene called *CDR1*, which is involved in cell size and nutritional control. NIM1/CDR1 acts to inhibit WEE1 activity and thus has a net inducing effect on CDC2 activity. Another protein that interacts with CDC2-CDC13 is the SUC1 protein, the gene of which was isolated as a suppressor of *CDC2ts* (Hayles *et al.*, 1986). Deletion of *SUC1* leads to mitotic bypass with multiple rounds of S phase (Hayles *et al.*, 1994; Fisher and Nurse, 1996); however, as the *CDC2ts* mutants rescued by *SUC1* also arrest at G2/M, SUC1 is also implicated in mitosis. The function of SUC1 is not yet clear, although there is genetic evidence of a direct interaction with CDC2 (Moreno *et al.*, 1989) and biochemical evidence of strong association between SUC1 and CDC2 homologues from a number of species (Hindley *et al.*, 1987; Dunphy *et al.*, 1988; Arion *et al.*, 1988).

Evidence has been accumulating that the RUM1 gene product is an important control factor in the fission yeast cell cycle. *RUM1* null mutants cannot delay entry to S phase and cannot institute a G1 block, and so consequently are sterile (Moreno and Nurse, 1994). Overexpression of *RUM1* gives a phenotype similar to a *CDC13* null (Moreno and Nurse, 1994;

Hayles *et al.*, 1994). Four cyclins in total have been demonstrated to complex with CDC2 - PUC1, CIG1, CIG2 and CDC13 (Fisher and Nurse, 1995). CDC2-CIG2 activity peaks at START (in parallel with the peak of CIG2 expression) implicating it in control of the traverse of START (Obara-Ishihara and Okayama, 1994; Martín-Castellanos *et al.*, 1996; Mondesert *et al.*, 1996). CDC13 is the mitotic cyclin (Booher *et al.*, 1989; Moreno *et al.*, 1989). CIG1 may have a role in the G1/S transition as a CDC13-CIG2 double null can traverse S phase but a triple null CDC13-CIG2-CIG1 cannot (Fisher and Nurse, 1996; Mondesert *et al.*, 1996). The role of PUC1 in the cell cycle is not yet clear (Forsburg and Nurse, 1991, 1994). RUM1 inhibits CDC2-CDC13 and CDC2-CIG2 activity *in vivo* during G1, and also promotes the turnover of CDC13 (Correa-Bordes and Nurse, 1995; Martín-Castellanos *et al.*, 1996; Correa-Bordes *et al.*, 1997). Transient CDC2 inhibition during G1 may be important in setting the minimum cell size to pass START and therefore timing START as well as preventing premature entry to M phase (Labib and Moreno, 1996; Martín-Castellanos *et al.*, 1996). The expression of *RUM1* oscillates, with RUM1 present from anaphase to the end of G1, mirroring the period of mitotic cyclin degradation (Amon et al,. 1994; Brandeis and Hunt, 1996). Not all CDC13 is destroyed at the end of M phase (Benito *et al.*, 1998) therefore there is a requirement for CDC2-CDC13 inhibition and CDC13 destruction in G1, which RUM1 mediates. The peak of CDC2-CIG1 activity at G1/S may therefore act to relieve CDC2-CDC13 and CDC2-CIG2 repression by RUM1 (Benito *et al.*, 1998).

A general schedule of cell cycle progression is shown (Fig. 1.2). As the cell leaves M phase and enters G1, RUM1 inhibits the remaining CDC2-CDC13 and promotes CDC13 turnover, thus preventing entry to M phase from pre-START G1 or premature entry to S phase. Levels of CIG2 and CDC13 rise and the abundance of the CDC2-CIG2 and CDC2-CDC13 complexes increase, held inactive by RUM1 until the minimum size to pass START is achieved. Once a certain level of CDC2-CIG2 is achieved, RUM1 inhibition is bypassed and the active complex drives the cell past START. As RUM1 function is lost, CDC2-CDC13 activity becomes inhibited by phosphorylation on Tyr15 by the WEE1 kinase. CDC2-CIG2 may or may not have functions once START has been passed. As the cells leave S phase the position within the cell cycle is monitored by rising levels of CDC25 and the presence of (inactive) CDC2-CDC13 complex. Inhibition of CDC2-CDC13 is relieved by inhibition of WEE1 kinase by the NIM1/CDR1 kinase, in response to cell size and nutritional status, and peaking levels of CDC25 in late G2 cells activates the CDC2-CDC13 complex by dephosphorylation of Tyr15, driving the cell into M phase. CIG1 expression peaks at M phase and the CDC2-CIG1 complex is active within the progression of M phase, with CDC13 being rapidly destroyed at anaphase as a necessity for completion of and exit from M phase. Given that gene functions at START can be temporally mapped, the decision to enter S phase is not a simple bimodal switch. The 'switch' of G1/S appears to have functions beyond START and likewise, the CDC2-CDC13 complex that

controls entry to M phase appears to have functions beyond passing the restriction point. As is only reasonable, the nutritional status of the cell influences cell cycle progression by impinging ultimately on the effectors of cell cycle control mentioned.

1.2.2: Budding yeast cell cycle

Budding yeast, *Saccharomyces cerevisiae*, unlike fission yeast is stable as both a haploid and a diploid. The bud initiates in late G1 and grows throughout the rest of the cycle, although it does not grow as big as the mother cell. For division, the nucleus migrates to the bud neck but as with *Schizosaccharomyces pombe*, the nuclear membrane stays intact. In contrast to fission yeast there is no chromosome condensation, and due to the unique nature of division, there are certain requirements to be met; for example, the spindle poles must replicate early. Similarly, the interphase microtubule array must reconfigure to the mitotic nuclear spindle early. Consequently, *Saccharomyces cerevisiae* undergoes much of the cell cycle with cytological markers indicative of G2/M and there is no clear division between S phase, G2 and M phase. The point at which M phase can be said to have initiated has been a point of contention (Nurse, 1985). Cell cycle regulation appears to occur mainly in G1, where cell size and nutritional status are monitored; if nutrition is limiting, G1 is consequently extended. Bud growth, DNA replication and spindle reorganisation initiate simultaneously but are genetically distinct processes, and therefore represent three parallel paths co-ordinated with respect to the cell cycle (Pringle and Hartwell, 1981). A *Saccharomyces cerevisiae* cell in G1 has three possible fates: Firstly, if nutrients are not limiting, in can enter S/G2/M and divide. Secondly, under conditions of nutrient limitation it can enter a stationary state which is metabolically quiescent, and which exhibits a lag period before re-entry to G1 after being introduced to a rich medium (Pringle and Hartwell, 1981). Thirdly, an alternative response to nutrient limitation is for haploids to mate; peptide hormones arrest the two haploid cells synchronously in G1, the cells fuse and continue to propagate as a diploid, or under conditions of starvation sporulates. *Saccharomyces cerevisiae* mates easily and can switch mating types, maximising the chances of diploid formation (Thorner, 1981).

As in the case of *Schizosaccharomyces pombe*, CDC mutants were isolated throughout the cell cycle under the criteria of cell cycle arrest without blocking of cell growth (Hartwell 1974, 1978; Pringle 1978). Such mutants were identified visually, for example by arrest with uniform bud phenotype, or with a visual mitotic defect. As with *Schizosaccharomyces pombe*, mutants isolated were conditional and thus suppressor analysis allowed the identification of other cell cycle regulatory genes.

Several genes were identified in *Saccharomyces cerevisiae* whose function mapped to

G1/S. The linchpin gene isolated was *CDC28*, lesions in which resulted in a cell cycle block at G1/S (Hartwell *et al.*, 1973; Reed, 1980) but certain allelic mutations arrested in G2 (Piggot *et al.*, 1982; Reed and Wittenberg, 1990), thus implicating *CDC28* as having functions in both G1/S and G2/M. Analysis of the coding sequence showed *CDC28* to be a homologue of *Schizosaccharomyces pombe CDC2*, a serine/threonine kinase, and complementation studies demonstrated functional equivalence (Beach *et al.*, 1982; Booher and Beach, 1986). CDC28 was demonstrated to have kinase activity (Reed *et al.*, 1985) but that it was inactive as a monomer (Wittenberg and Reed, 1988). Analysis of the *CDC28ts* mutants demonstrated that arrest was most often in G1; suppressor analysis identified three more regulatory genes. The first of these, the *CKS1* gene, was isolated as a high copy suppressor of some *CDC28ts* alleles, although it was not able to completely bypass the requirement for *CDC28* function (Hadwiger *et al.*, 1989a). Deletion of *CKS1* is lethal, resulting in a nonbudded phenotype and arrest is therefore in G1 prior to START. Sequence analysis of the coding sequence revealed *CKS1* to be homologous to the *Schizosaccharomyces pombe SUC1* gene. The other two genes, termed *CLN1* and *CLN2*, were also identified as high copy suppressors of *CDC28ts* mutants, but like *CKS1* they could not bypass the requirement for *CDC28* function (Hadwiger *et al.*, 1989b). Sequence analysis revealed that neither had homology to the A and B type family of cyclins from metazoan cells, and they therefore appeared to form a distinct cyclin subgroup. A third member of this subgroup, the *CLN3* gene, was isolated independently on two separate occasions as dominant mutant alleles: The *WHI1-1* gene (Nash *et al.*, 1988), which accelerated cells through G1 with a reduced cell size, and as the *DAF1-1* gene (Cross, 1988) which resulted in cells insensitive to α-factor mating signal-induced G1 arrest. Both of these dominant mutant alleles were shown to be truncations in the gene that was renamed *CLN3* (Nash *et al.*, 1988; Cross, 1988). Deletion of one or even two of the *CLN* genes is not lethal, although cells exhibit an increase in cell size and an extended G1. However, deletion of all three is lethal and leads to arrest in G1, therefore the functions of these three genes are redundant (Nasmyth, 1993).

START in budding yeast is generally defined as the period when spindle pole body duplication, budding and DNA replication all initiate (Nasmyth, 1993). CDC28 function mapping to START requires CLN1-3, as not only does deletion of all three lead to a G1 arrest (Nasmyth, 1993) but ectopic overexpression of *CLN1-3* results in acceleration through START (Futcher, 1996). The temporal expression pattern of the three genes differs: CLN3 transcript and protein are present throughout the cell cycle but protein levels rise in early G1 (Tyres *et al.*, 1993; McInerny *et al.*, 1997). This is a result of a translational control element in the 5' untranslated region (UTR; Polymenis and Schmidt, 1997) responsive to the RAS-cAMP signalling pathway (Hall *et al.*, 1998). Thus under conditions of slow growth *CLN3* expression is repressed (Gallego *et al.*, 1997) providing a direct coupling of cell growth with the cell cycle. By contrast, the expression of *CLN1* and *CLN2* are cell cycle regulated with a peak in protein

levels and CDC28-associated kinase activity in late G1 (Tyres *et al.*, 1993). *CLN* transcription and entry to S phase are regulated by the SBF and MBF transactivators, which form a program of START-dependent gene activation (Breeden, 1995). CDC28-CLN3 activates SBF at START, inducing expression of *CLN1* and *CLN2* (Tyres *et al.*, 1993). SBF is a substrate for CDC28-CLN1/2, thus forming a positive feedback loop. MBF is activated near S phase and induces the expression of *CLB5* and *CLB6*, as well as other genes involved in DNA replication (Breeden, 1995). Generally speaking, CLNs are though to function by permitting CLB expression and activity at the correct point in the cell cycle (Nasmyth, 1996).

Negative control of CDC28-CLN complexes is regulated by the SIC1 gene product, which inhibits CDC28-CLN complexes (Schwob *et al.*, 1994; Schneider *et al.*, 1996). SIC1 destruction by the ubiquitin proteolytic machinery is regulated by CDC28-CLN (Schneider *et al.*, 1996), therefore once CDC28-CLN activity rises above a threshold, CDC28-CLN phosphorylation of SIC1 occurs, initiating its destruction. As with CDC2 in fission yeast, CDC28 binding to cyclins is positively regulated by phosphorylation on the conserved Thr161 residue (Thr167), and this event is performed by the CAK1/CIV1 kinase. Oddly, this CDK-related kinase is active as a monomer and has never been demonstrated to bind a cyclin partner (Kaldis *et al.*, 1996; Thuret *et al.*, 1996; Espinoza *et al.*, 1996). CDC28 differs from CDC2, however, in that strains harbouring an altered *CDC28* allele such that CDC28 could not be phosphorylated on the Tyr equivalent of fission yeast Tyr15 could grow and divide normally (Sorger and Murray, 1992; Amon *et al.*, 1992). This has been confirmed by studies on strains lacking the budding yeast *WEE1* homologue *SWE1* (Booher *et al.*, 1993); budding yeast lacking SWE1 function can still divide normally and delay entry to M phase in response to DNA damage. The reason for budding yeast possessing an operative WEE1-CDC25 homologous system, SWE1-MIH1 (Russell *et al.*, 1989) is therefore unclear and has been the subject of speculation (Dunphy, 1994).

Six B-type cyclins have been identified in *S. cerevisiae*, *CLB1-6*, which are involved in the promotion of S phase and M phase (reviewed by Nasmyth, 1996). *CLB5* and *CLB6* are expressed in late G1 (Breeden, 1995). Both are required for S phase entry. Deletion of *CLB5* and *CLB6* impairs S phase entry but does not affect START (Epstein and Cross, 1992; Kuhne and Linder, 1993; Schwob and Nasmyth, 1993), presumably due to degeneracy of function on the part of CLB1-4. *CLB3* and *CLB4* are expressed during S phase while *CLB1* and *CLB2* are expressed in G2 (Fitch *et al.*, 1992). The degeneracy of the CLB proteins makes defining precise *in vivo* functions problematic. The temporal pattern of expression would suggest that CLB3 and CLB4 function in S phase progression, and that CLB1 and CLB2 function at M phase. Deletion of *CLB1-4* does not inhibit DNA replication but cells do not form bipolar mitotic spindles and arrest in G2 (Fitch *et al.*, 1992; Amon *et al.*, 1993), suggesting that CLB3/4 are not needed for DNA synthesis but more specifically for correct spindle formation and function. Deletion of

CLB1 and *CLB2* is lethal, leading to a cell cycle arrest with a fully formed mitotic spindle. A triple deletion of *CLB1*, *CLB2* and *CLB3* leads to cells arrested with an immature spindle (Surana *et al.*, 1991; Fitch *et al.*, 1992; Richardson *et al.*, 1992), supporting the theory that CLB3/4 function during early spindle formation. However, a double *CLB3/4* deletion is viable and shows no obvious phenotype (Schwob and Nasmyth, 1993). With this type of analysis it is therefore difficult to determine if CLBs have overlapping functions or only display degeneracy of function in the absence of other CLBs.

Other CDK-cyclin pairs are involved in G1 progression in *Saccharomyces cerevisiae*. The CDK PHO85, which has 51% identity to CDC28, can bind multiple cyclins including PCL1 and PCL2 (Measday, 1997). PHO85-PCL1/2 complexes have potential roles in G1 progression as deletion of *PCL1* and *PCL2* in a *CLN1/CLN2* null results in a G1 arrest (Espinoza *et al.*, 1994; Measday *et al.*, 1994), and thus PHO85 is essential if there is no CDC28 activity present. PHO85-PCL1 has been demonstrated to phosphorylate SIC1 *in vitro* (Sanchez-Diaz *et al.*, 1998; Nishizawa *et al.*, 1998). SIC1 phosphorylation leads to its rapid degradation, which is a requirement for relief of CDC28-CLN1/2 repression and DNA synthesis to occur (Nishizawa *et al.*, 1998). Cyclin *PCL9* expression peaks in late M phase/early G1, and deletion causes bud formation defects (Tennyson *et al.*, 1998). A general schematic of budding yeast cell cycle control is shown (Fig. 1.3).

The first activity identified for PHO85 was in complex with the PHO80 cyclin, in regulation of *S. cerevisiae* phosphate metabolism. In conditions of limited phosphate availability, the PHO2 and PHO4 transcription factors binds DNA and activate the transcription of genes involved in phosphate metabolism, for example the *PHO5* gene which encodes an acid phosphatase (Jayaraman *et al.*, 1994). In conditions of high phosphate availability, PHO4 cannot bind DNA due to phosphorylation by the PHO85-PHO80 complex. In addition, there is control of the PHO85-PHO80 complex by binding to the PHO81 gene product, which inhibits the complex. PHO81 can bind PHO85-PHO80 during conditions of both high and low phosphate availability, but is only inhibitory under conditions where phosphate is low, therefore PHO81 activity is regulated post-translationally (Schneider *et al.*, 1994; Hirst *et al.*, 1994). PHO85 may also be involved in the control of other cellular activities that are not directly concerned with the regulation of cell cycle progression. For example, PHO85 is involved in glycogen metabolism (Timblin *et al.*, 1996) in complex with the PCL8 and PCL10 cyclins (Huang *et al.*, 1996), regulating glycogen biosynthesis by phosphorylation of the glycogen synthetase GSY2 (Huang *et al.*, 1998). PHO85-PCL8/10 negatively regulates GSY2, and as would be expected a double *PCL1/2* deletion accumulates glycogen due to deregulation of PHO85-PCL1/2 control (Huang et al., 1998). The functions of the PHO85 cyclins PCL5, PCL6, PCL7 and CLG1 are currently unknown. Three further CDK-cyclin pairs have been identified. KIN28-CCL1 is the CDK-cyclin component of the basal transcription factor TFIIH, and SRB10-SRB11 forms part of the RNA

polymerase II (RNA pol II) holoenzyme complex (Liao *et al.*, 1995; Valay *et al.*, 1995). Both display RNA pol II C-terminal domain (CTD) kinase activity and are therefore presumably involved in transcriptional regulation. A third RNA pol II CTD activity has been found to consist of the divergent CDK-cyclin pair CTK1-CTK2, in association with a third protein, CTK3, that has no homology to any known protein (Sterner *et al.*, 1995).

1.3: Metazoan cell cycle

The metazoan cell cycle follows the same schedule as the yeast cell cycle: G1, S phase, G2 and M phase. G1, S phase and G2 are often collectively called interphase. M phase is described as a series of subphases: Prophase is the first subphase, in which chromatin condenses and the mitotic spindle forms. It is followed by prometaphase, where the nuclear membrane fragments; metaphase, when the spindle fibres align chromosomes along the middle of the cell; anaphase, when paired chromosomes are separated to opposite poles of the cell; and telophase, when the nuclear envelope reforms around the daughter nuclei and the mitotic spindle disperses. Telophase is followed by cytokinesis, division of the cell to form two autonomous daughter cells.

Early progress in delineating the molecular mechanisms underpinning the metazoan cell cycle came from experiments using nuclear transplantation to challenge nuclei from one position in the cell cycle with factors present in the cytoplasm of cells at another part of the cell cycle (Prescott, 1976). These determined that entry to S phase is controlled by a *trans*-acting cytoplasmic signal which is not present in pre-S phase G1 cells ('S phase promoting factor'). The second fundamental control process discovered was the inhibition of successive rounds of S phase, in that a G2 nucleus transplanted to an S phase cytoplasm will not undergo another S phase (Heichman and Roberts, 1994). Further, that an interphase nucleus challenged with M phase cytoplasm will attempt to segregate its DNA, thus being promoted by a *trans*-acting cytoplasmic signal that was termed maturation promoting factor (MPF; also known as M phase promoting factor). A histone H1 kinase activity that peaked at M phase was identified and partially purified (Arion *et al.*, 1988), and subsequently demonstrated to correlate to MPF (Labbé *et al.*, 1989). MPF contained the metazoan equivalent of fission yeast CDC2; human *CDC2* (later renamed *CDK1* for cyclin dependent kinase 1) had been isolated by complementation of a fission yeast *CDC2ts* mutant (Lee and Nurse, 1987). MPF was demonstrated to be a multiprotein complex, with a 62 kDa protein, subsequently identified as cyclin B (Pines and Hunter, 1989) coprecipitating with the 34 kDa CDK1 from G2 but not G1 HeLa cells (Draetta and Beach, 1988). Cyclins were first identified in sea urchin oocytes as proteins that showed an oscillating pattern of expression, synthesised during interphase and

destroyed at M phase (Evans *et al.*, 1983). Purification of MPF to homogeneity confirmed that MPF consisted of a complex formed of a 1:1 stochiometric association of CDK1 and cyclin B (Labbé *et al.*, 1989). The binding of CDK1-cyclin B to SUC1, the M phase associated histone H1 kinase activity, and the periodic nature of cyclin B expression suggested that the mechanisms of cell cycle control in multicellular eukaryotes might essentially be the same as those observed in yeast, not least because of the remarkable degree of evolutionary conservation between CDC2 and CDK1.

To date, there have been nine genes isolated and classified as *CDK*s in humans, although not all are directly involved in cell cycle control (reviewed by Pines, 1995; Gao and Zelenka, 1997). The CDK proteins associate with members of the cyclin family, of which a number have so far been identified in humans: Cyclin A (Henglein *et al.*, 1994), cyclins B1, B2 and B3 (Gallant and Nigg, 1994), cyclin C (Lew *et al.*, 1991), cyclins D1, D2 and D3 (Xiong *et al.*, 1991; Motokura *et al.*, 1992), cyclin E (Koff *et al.*, 1991), cyclin F (Bai *et al.*, 1994), cyclins G1 and G2 (Tamura *et al.*, 1993; Horne *et al.*, 1996), cyclin H (Mäkelä *et al.*, 1994; Fisher and Morgan, 1994), cyclin I (Nakamura *et al.*, 1995), cyclin K (Edwards *et al.*, 1998) and cyclin T (Wei *et al.*, 1998). The known patterns of interaction of these CDKs and cyclins are shown in table 1. CDK partners have not yet been identified for cyclin F, cyclin I or cyclin K, although cyclin K appears to be associated with RNA pol II C-terminal domain (CTD) kinase activity (Edwards *et al.*, 1998) and cyclin G is implicated in p53-mediated DNA damage arrest (Horne *et al.*, 1996; Reimer *et al.*, 1999).

Studies on the function of CDK1 showed an immediate difference between metazoans and yeast. In yeast, CDC2/CDC28 has functions at both G1/S and G2/M transitions. Evidence from metazoan cells suggested that CDK1 functioned only at G2/M (or at least, had a critical function only at G2/M). A *Xenopus laevis* oocyte extract depleted of CDK1 will carry out DNA replication but will not enter M phase (Fang and Newport, 1991). A temperature sensitive CDK1 mutant has been shown to traverse S phase but not pass G2/M at the restrictive temperature (Hamaguchi *et al.*, 1992). The association of CDK1 with the A and B cyclins (often dubbed 'mitotic cyclins') has been further investigated, and the kind of degeneracy displayed by the yeast CLN and CLB cyclins is not mirrored in metazoans. For example, transgenic mice homozygous for either cyclin B1 or B2 dysfunction have shown that cyclin B1 but not B2 is essential for viability (Brandeis *et al.*, 1998). The subcellular localisation of the B-type cyclins differs: B1 is located in the cytoplasm at interphase and moves to the nucleus at M phase (Ohi and Gould, 1999). Cyclin B2 is always located at the Golgi, presumably having a function in complex with CDK1 in Golgi fragmentation during M phase, and cyclin B3 is always located in the nucleus (Ohi and Gould, 1999). The synthesis of the B cyclins begins in G1 and they must accumulate above a threshold level to be able to promote mitosis (Murray and Kirschner, 1989; Solomon *et al.*, 1990). Threshold is defined as the point beyond which no further protein

synthesis is required for MPF to induce mitosis. Cyclin B starts to associate with CDK1 in late S phase. Binding of cyclin B promotes phosphorylation of Thr161 by the CDK1 activating kinase (CAK) but the complex is held inactive due to phosphorylation of Thr14 and Tyr15 (Pines, 1993; King *et al.*, 1994; Pines, 1995). Phosphorylation of Thr14 is carried out by MYT1 (Booher *et al.*, 1997; Liu *et al.*, 1997). Although MYT1 also appears able to phosphorylate Tyr15, the major Tyr15 phosphorylating kinase is WEE1, which phosphorylates exclusively on this residue (McGowan and Russell, 1993). MYT1 and WEE1 become heavily phosphorylated once the cell enters M phase, a change concomitant with their inactivation (McGowan and Russell, 1995; Mueller *et al.*, 1995a,b; Booher *et al.*, 1997). In parallel, the dual-specificity CDC25 phosphatase, held inactive until then by dephosphorylation becomes phosphorylated and active (Izumi *et al.*, 1992; Izumi and Maller, 1993; Kumagai and Dunphy, 1992, 1996). Active CDC25 dephosphorylates CDK1 on Thr14 and Tyr15 (Gautier *et al.*, 1991; Kumagai and Dunphy, 1991, 1992) and, released from inhibition, the CDK1-cyclin B complex drives the cell into M phase. Active CDK1-cyclin B can phosphorylate CDC25 (Hoffmann *et al.*, 1993; Dunphy, 1994), forming a positive feedback loop that results in rapid activation of the complex and entry to M phase. During metaphase, ubiquitin-mediated proteolytic machinery is activated that quickly destroys cyclin B prior to anaphase and thus inactivates CDK1 (King *et al.*, 1994). However, cyclin B destruction and CDK1 inactivation is not necessary for anaphase as in cells with ectopically expressed mutant forms of cyclin B that are ultrastable, chromatid separation occurs normally and the cells arrest in telophase (Holloway *et al.*, 1993); however, activation of the machinery is required for completion of cytokinesis. Thus as the cell leaves M phase, CDK1 is once more an inactive monomer, and cyclin B accumulation begins anew. Active CDK1-cyclin A is also required for entry to M phase.

The *CDK2* gene was isolated by complementation of an *S. cerevisiae CDC28ts* mutant (Elledge and Spottswood, 1991); only *CDK1*, *CDK2* and *CDK3* can rescue budding yeast *CDC28ts* mutants (Meyerson *et al.*, 1992). The *CDK2* gene is 66% identical to *CDK1*, the highest homology to *CDK1* of all the *CDK* family (Elledge *et al.*, 1991). The involvement of *CDK2* in the cell cycle has been demonstrated by expression of a dominant negative allele, which resulted in a G1 cell cycle arrest (Van den Heuvel and Harlow, 1993). Biochemical evidence for a G1/S involvement came from studies on depletion of CDK2 from *Xenopus laevis* oocyte extracts, which could undergo M phase but not enter S phase (Fang and Newport, 1991). As regards cyclin binding, CDK2 has been shown to be significantly associated with both cyclin E and cyclin A. Cyclin E is expressed during G1/S and forms a complex with CDK2 near the 'restriction' (R) point at which a cell commits to S phase; the activity of the complex is maximal at this point and declines as cyclin E is rapidly destroyed as S phase is entered (Lew *et al.*, 1991; Dulic *et al.*, 1992; Koff *et al.*, 1992). Overexpression of cyclin E causes accelerated progression through G1 and entry to S phase (Ohtsubo and Roberts, 1993), suggesting that premature CDK2

activation mediates this effect. Cyclin A is expressed at the start of S phase (Pines and Hunter 1990; Tsai *et al.*, 1991). CDK2 and cyclin A form a complex which is required for S phase progression (Girard *et al.*, 1991; Elledge *et al.*, 1992; Rosenblatt *et al.*, 1992) and overexpression of cyclin A causes premature entry to S phase (Rosenberg *et al.*, 1995). The CDK2-cyclin E and CDK2-cyclin A complexes are both associated with the retinoblastoma protein (pRB) and members of the E2F family of transactivators (Mudryj *et al.*, 1991; Devoto *et al.*, 1992; Lees *et al.*, 1992; Pagano *et al.*, 1992).

CDK3 is something of a mystery, as it has not been demonstrated to associate with a cyclin. It has been classified as a *CDK* on the basis of high homology to *CDK1* and *CDK2* and also the fact that it can complement a *CDC28ts* mutant (Meyerson *et al.*, 1992) although it cannot complement a *CDC2ts* mutant. A *CDK3* dominant negative mutant showed a G1 block that was specifically released by wild-type *CDK3* but not *CDK2*, implying an essential function in G1/S progression distinct from CDK2 function (Van den Heuvel and Harlow, 1993). Despite the fact that purified recombinant CDK3 can complex with cyclins A and E *in vitro* and shows kinase activity toward the retinoblastoma protein (pRB), no cyclin partner has been identified for CDK3 *in vivo* (Harper *et al.*, 1995; Connel-Crowley *et al.*, 1997). The temporal activity of CDK3 has been mapped (Braun *et al.*, 1998). CDK3 activity is absent in G0 and early G1, peaks in mid G1 and declines during G2; comparative mapping with cyclin D- and cyclin E-associated kinase activity showed that they map before and after the appearance and peak of CDK3 activity, respectively (Braun *et al.*, 1998). As CDK3 protein levels do not fluctuate throughout the cell cycle, the data suggests that CDK3 activity is controlled by as yet unidentified partner molecule(s), although as CDK3 activity overlaps temporally with the cell-cycle fluctuation of cyclin E-associated kinase activity a role for cyclin E in CDK3 function cannot be ruled out (Braun *et al.*, 1998).

The *CDK4* gene was originally isolated in a general screen for serine/threonine kinases (Hanks, 1987; Matsushime *et al*, 1992) and is most closely related to the *CDK6* gene (Meyerson *et al.*, 1992), with approximately 70% identity between them at the amino acid level. CDK2, CDK4, CDK5 and CDK6 have all been demonstrated to bind the D-type cyclins, but the major partners for cyclins D1-3 are CDK4 and CDK6 (Matsushime *et al.*, 1992; Xiong *et al.*, 1992; Bates *et al.*, 1994; Meyerson and Harlow, 1994). Expression of the D-type cyclins does not oscillate to any significant degree during the cell cycle, but their cellular levels are absolutely dependent on the presence of growth factors, peaking at the G1/S boundary, and they are rapidly degraded upon the withdrawal of mitogens (reviewed by Sherr, 1994). Studies with transgenic mice have demonstrated that cyclins D1-3 are for the most part functionally redundant but that each has tissue-specific functions (Sicinski *et al.*, 1995). The primary function for CDK4-cyclin D (and possibly CDK6-cyclin D) is phosphorylation of pRB (Matsushime *et al.*, 1994; Kato *et al.*, 1993). Given the temporal expression and activity of cyclins D1-3, E and A, the general

model of G1/S progression is that mitogenic signals stabilise cyclins D1-3 in the G0-G1 transition (Sherr, 1994). Cyclins D1-3 complex with CDK4/6 to form active kinase complexes which hyperphosphorylate pRB at about the G1/S transition, releasing E2F. E2F induces cyclin E expression, which maintains pRB phosphorylation in complex with CDK2 as the D-type cyclins are destroyed and CDK4/6 inactivated. Later induction of cyclin A directs CDK2 activity through S phase as cyclin E is destroyed.

The *CDK5* gene was isolated by library screening for *CDK1*-related sequences (Meyerson *et al.*, 1992) and independently as a histone HI kinase activity from bovine brain tissue (Lew *et al.*, 1995). It has been demonstrated to be expressed in many different differentiated tissues but is only found to be active in the brain, apparently due to the highly restricted expression of its regulatory partner p35, and two further partners derived from p35 by proteolysis, p25 and p23 (Lew *et al.*, 1995). Although CDK5 is highly homologous to other members of the CDK family, the primary sequence of p35 is unrelated to that of the members of the cyclin family. Substrates of CDK5-p35 include the microtubule-associated protein TAU and neuron-specific intermediate filaments, suggesting a role in regulation of neurite outgrowth (Lew *et al.*, 1995). This is supported by cotransfection of neurons by *CDK5* and *p35*, which results in longer neurite outgrowth, and expression of dominant negative CDK5 or antisense p35, which results in fewer and shorter neurite outgrowth (Nicolic *et al.*, 1996). A homozygous deletion of *CDK5* in mice causes a lethal phenotype with central nervous system abnormalities (Oshima *et al.*, 1996; Gilmore *et al.*, 1998). The *in vivo* association of CDK5 with other proteins has been demonstrated in bovine brain, where three forms appear to exist: monomeric CDK5, CDK5-p35, which is inactive, and CDK5-p25, which is highly active (Lee *et al.*, 1996). Interestingly, gel filtration chromatography indicated CDK5-p35 was part of a large (670 kDa) macromolecular complex (Lee *et al.*, 1996). CDK5 has also been demonstrated to associate with cyclin D2 (Guidato *et al.*, 1998) although the functional significance of this interaction is not yet clear, and it has been suggested that CDK5 may have a role in the control of muscle differentiation and patterning (Philpott *et al.*, 1997).

The function of CDK7 is not yet clear. The CDC activating kinase (CAK) activity that phosphorylates Thr161 of CDK1 was found to be a complex of CDK7 (Fesquet *et al.*, 1993; Poon *et al.*, 1993) and cyclin H (Fisher and Morgan, 1994; Mäkelä *et al.*, 1994). As almost all CDKs require phosphorylation on the conserved Thr161 equivalent for activity (Nigg, 1995) it was surprising that CAK was itself a CDK-cyclin pair. CDK7-cyclin H were demonstrated to be associated with a protein called Menage à trois (MAT1) that appears to act as an assembly factor (Tassan *et al.*, 1995a), and CDK7-cyclin H-MAT1 was shown to be a component of the basal transcription complex TFIIH (Roy *et al.*, 1995). The *Saccharomyces cerevisiae* CAK is a monomer called CAK1/CIV1 (Espinoza *et al.*, 1996; Kaldis *et al.*, 1996), therefore it was unclear as to whether or not CDK7 had a dual function as a component of TFIIH and as a CAK.

CDK7 has been isolated from cell extracts in different forms (Yankulov and Bentley, 1997) and while CDK7-cyclin H-MAT1 has a strong preference for CDK2 as a substrate, as a component of TFIIH it does not possess CAK activity (Rossignol *et al.*, 1997). Furthermore, phosphorylation of CDK7 is necessary for cyclin H binding in the absence of MAT1, but phosphorylation is not necessary for binding in the presence of MAT1 (Martinez *et al.*, 1997). Current evidence leans towards the dual function of CDK7 as a component of CAK that can be recruited to function as a modulator of transcriptional activity by association with TFIIH (Harper and Elledge, 1998).

The *CDK8* gene was isolated by PCR and subsequent library screening (Tassan *et al.*, 1995b). Cyclin C was identified through its ability to rescue *S. cerevisiae CLN* mutants (LaHue *et al.*, 1991; Leopold and O'Farrel, 1991; Lew *et al.*, 1991) and was therefore initially assumed to play a role in G1 progression, although no direct involvement in G1 progression could be demonstrated. CDK8 was demonstrated to associate with cyclin C by coprecipitation, and appeared to be structurally related to the yeast protein pair SRB10-SRB11, which forms part of the RNA pol II complex in *S. cerevisiae* (Tassan *et al.*, 1995; Liao *et al.*, 1995). CDK8-cyclin C has also been shown to associate with RNA pol II *in vivo* and to possess RNA pol II CTD kinase activity *in vitro* (Rickert *et al.*, 1996). CDK8-cyclin C is therefore the second CDK-cyclin pair demonstrated to be involved in transcriptional control.

The most recently identified member of the *CDK* family, *CDK9*, was also isolated by the approach of PCR and library screening (Graña *et al.*, 1994). Although CDK9 was demonstrated to have kinase activity as a monomer, overexpression of CDK9 does not result in an increase in cellular CDK9-associated activity (Garriga *et al.*, 1996a). In addition, it was demonstrated by gel filtration to associate with multimeric protein complexes of approximately 670 kDa and in smaller complexes (Garriga et al., 1996a) and to be capable of phosphorylating pRB (Garriga *et al.*, 1996b; De Luca *et al.*, 1997). It is expressed at higher levels in differentiating tissue and levels are highest in terminally differentiated tissue (Bagella *et al.*, 1998). CDK9 is a required component of the transcriptional elongation factor P-TEFb, which has RNA pol II CTD kinase activity (Marshall *et al.*, 1996; Yang *et al.*, 1996; Mancebo *et al.*, 1997; Yang *et al.*, 1997). A novel cyclin most closely related to cyclin C was identified, named cyclin T, which associates in complex with CDK9 and is required as part of the P-TEFb complex for RNA pol II CTD activity (Wei *et al.*, 1998). It is therefore possible that CDK9 functions in transcriptional control of differentiation.

Thus the cyclins and their CDK partners described represent functional pairings that have diverged further from the CDK1 paradigm, and some of which have evolved to regulate aspects of cellular biology that are only marginally related to cell cycle control. CDK-cyclin functional pairing is a powerful, flexible system that can adapt to serve as regulators and integration points for different systems.

From the delineation of the expression and pairings of the CDKs and cyclins described, a basic picture of metazoan cell cycle has emerged (Fig. 1.4). Newly formed daughter cells leave M phase with inactive CDK1 and no A or B cyclins, as the proteolytic machinery is still active: This activity ensures that CDK1-cyclin B cannot accumulate in G1 and the cell prematurely enter M phase from G1. As the cell nears the restriction point, the D cyclins are synthesised and - in some manner not yet understood, whether directly or indirectly - turn off the proteolytic machinery, thus allowing gradual accumulation of the B cyclins through late G1 and S phase. In this manner, analogous to the situation in budding yeast where CLB1 and CLB2 stability depends on the CLNs, so the stability of cyclins A and B depend on G1 cyclins. The D cyclins complex with CDK4 and CDK6, CDK4-cyclin D (and possibly CDK6-cyclin D) hyperphosphorylates pRB, which releases E2F1-3; these form heterodimers with members of the DP transactivator family (of which there are three documented members), and the resulting complexes induce the expression of genes required for DNA synthesis, including cyclin E. Newly synthesised cyclin E complexes with CDK2, which is synthesised in G1, and the complex maintains pRB in a hyperphosphorylated state. The D cyclins are then destroyed, thus limiting their span of influence to a narrow window that begins just before R and ends at the G1/S transition (Weinberg, 1995). CDK2-cyclin E maintains the induction of genes required in S phase via pRB/E2F until the G1/S boundary is passed, when cyclin E is rapidly destroyed via the ubiquitin-proteasome pathway due to phosphorylation by CDK2. Cyclin A, synthesised in late G1, then complexes with CDK2 and functions throughout S phase. CDK2-cyclin A negatively regulates the E2F transactivators as the cell traverses and prepares to leave S phase. As the cell leaves S phase CDK1 becomes associated with the B cyclins, which promotes phosphorylation on Thr161 by CDK7-cyclin H. The mitotic complexes are held inactive by phosphorylation on Thr14 by MYT1 and Tyr15 by WEE1 until late G2 when CDC25 activity rises due to increased phosphorylation, and it dephosphorylates both these residues, activating the CDK1-cyclin B complexes which drive the cells through the G2/M transition. CDK1-cyclin B functions until the condensed chromosomes are aligned on the metaphase plate, when the proteolytic machinery that has already destroyed cyclin A destroys cyclin B and renders CDK1 inactive. Anaphase proceeds, followed by cytokinesis, and the two new daughter cells re-enter G1.

Two other families of significant players in the regulation of cell cycle progression have been isolated, the cyclin dependent kinase inhibitor (CKI) proteins. The first family comprises three members: $p21^{Cip1}$ (Harper *et al.*, 1993; Xiong *et al.*, 1993), $p27^{Kip1}$ (Toyoshima and Hunter, 1994), and $p57^{Kip2}$ (Lee *et al.*, 1995). The CIP/KIP family act as specific inhibitors of CDK2 complexes and as positive regulators of CDK4/6 complexes (reviewed by Johnson and Walker, 1999; Sherr and Roberts, 1999). There are four members of the second family of CKIs: $p16^{INK4a}$ (Serrano *et al.*, 1993), $p15^{INK4b}$ (Hannon and Beach, 1994), $p18^{INK4c}$ (Guan *et al.*, 1994), and

$p19^{INK4d}$ (Chan *et al.*, 1995). INK4 proteins function to induce cell cycle arrest in a retinoblastoma protein-dependent manner, by inhibition of CDK4/6 activity. INK4 inhibition of CDK4/6 releases p21/p27, which inhibit CDK2 complexes, thereby preventing a bypass of CDK4 function in control of the retinoblastoma protein-mediated control of E2F transactivation of S phase (Sherr and Roberts, 1999).

1.4: Cell signalling in trypanosomes

In many metazoan organisms the methods by which a cell receives extracellular signals, transduces them across the plasma membrane, integrates and transmits that information to the correct effector mechanisms, and the manner in which these webs of signalling pathways communicate with each other have been extensively studied. While the picture is far from complete in most paths of signalling, a great deal of information has been garnered. The same cannot be said of trypanosomatids; ten years ago it was not known whether the same kinds of mechanisms for signal transduction and intracellular signalling existed in trypanosomes. Over the course of the last ten years fragmentary evidence has accumulated that trypanosomes possess many of the same types of signalling mechanisms as metazoan organisms. However, as might be expected evolutionary divergence as well as the trypanosomes' unique biology means that such paths, when finally elucidated, may well throw up unexpected surprises. What is known about signalling mechanisms in trypanosomes is discussed below.

1) Cell surface receptors: The only receptor that to date has been characterised to a molecular level is the transferrin receptor. Transferrin was identified as an essential component of culture medium for the bloodstream form of *T. brucei*, and a transferrin-binding species identified (Schell *et al.*, 1991). This was discovered to be the product of the expression site associated gene number 6 (*ESAG6*). There are approximately 20 telomeric sites for the expression of variant surface glycoprotein genes; each site has a number of other genes associated with it, which are homologous between expression sites. Further analysis revealed that the functional receptor was a heterodimer of ESAG6 and ESAG7, attached to the plasma membrane by a glycosylphosphatidylinositol (GPI) anchor, located at the C-terminal of ESAG6 (Salmon *et al.*, 1994; Steverding *et al.*, 1994; Ligtenberg *et al.*, 1994). Furthermore, although the receptor is located in the flagellar pocket the majority is found in the lumen of the pocket, suggesting that the receptor may be released to bind transferrin and then re-internalised, although the mechanisms involved are unclear (Salmon *et al.*, 1994; Steverding *et al.*, 1994). The two genes that encode the components of the transferrin receptor are not identical between different expression sites (ESs) thus an *in situ* switch between which ES is active results in the

expression of transferrin receptors with different affinities for transferrin from any particular host (Steverding *et al.*, 1995; Borst *et al.*, 1996; Bitter *et al.*, 1998). A striking example is a single amino acid substitution that accounts for a 10-fold difference in the affinity of the receptor for the same transferrin molecule (Salmon *et al.*, 1997). Furthermore, the same receptor isoform has large differences in affinity for transferrins from different mammalian species, and trypanosomes appear to undergo *in situ* ES switching until an optimal isoform is expressed for the particular mammalian transferrin (Borst *et al.*, 1996).

Host epidermal growth factor has also been demonstrated to influence trypanosome growth (Hide *et al.*, 1989) although a receptor remains to be identified. Low density lipoprotein (LDL) influences trypanosome growth (Coppens *et al.*, 1988) and although it is clear that trypanosomes possess an LDL receptor (Coppens *et al.*, 1991) the gene has not yet been identified. The best candidate appears to be a protein of approximately 145 kDa with high antigenic similarity to the host LDL receptor (Bastin *et al.*, 1994). Other host-derived molecules that have been demonstrated to influence trypanosome growth but for which specific receptors have not yet been isolated include interferon γ (INFγ; Olssen *et al.*, 1993), and tumour necrosis factor α (TNFα; Lucas *et al.*, 1993). These two may be considered to be antagonistic growth regulators, as INFγ stimulates proliferation (Olssen *et al.*, 1993) whereas TNFα, secreted from macrophages under stimulation by trypanosome-secreted factors (Tachado and Schofield, 1994), has a lytic effect. The effect of TNFα appears to be dependent on differential sensitivity of trypanosomes during parasitaemia, as TNFα exerts a lytic effect only during late parasitaemia (high parasite density).

Adenylate cyclase activity has been identified in trypanosomes, at least some of which is stage-regulated (Rolin *et al.*, 1990). Four adenylate cyclase genes have also been identified in *T. brucei*, one of which is contained in the polycistronic transcription units of the VSG genes (Paindavoine *et al.*, 1992) and is termed *ESAG4*. The ESAG4 protein is predicted to be located on the flagellum (Paindavoine *et al.*, 1992) and is expressed only in the bloodstream form of the parasite. Genes related to *ESAG4*, termed *GRESAG*s, have been found to be expressed in both bloodstream and procyclic forms of *T. brucei* (*GRESAG4.1*, *4.2* and *4.3*; Alexandre *et al.*, 1990); comparison over the receptor domain of these molecules gives three groups: *ESAG4*, *GRESAG4.1* and *GRESAG4.2/4.3*. Analysis of the genomic organisation of the *GRESAG* genes has shown that *GRESAG4.3* is a single copy gene, whereas *GRESAG4.1* exists as a multigene family of at least 9 members all located on a 3 Mb chromosome (Alexandre *et al.*, 1996). Within the *GRESAG4.1* family three of the genes clustered in a tandem array while the others were scattered over a 300 kb region. This distribution was reflected in sequence comparisons, which showed the tandemly arranged genes to have greatest identity between themselves, with the scattered copies showing more divergence (Alexandre *et al.*, 1996). At least several of the *GRESAG4.1* genes are transcribed in both bloodstream and procyclic forms. Control of

adenylate cyclase activity is cryptic; there is no evidence of G-protein involvement (Pays *et al.*, 1997) although the putative receptor-like structure of the trypanosome adenylate cyclases is quite different from that of mammalian cyclases, and may suggest activation is via homodimerisation following ligand binding. Activation by peptide binding to the extracellular domain has not been demonstrated, although tsetse fly proventricular extracts can induce activation (Van den Abbeele *et al.*, 1995). There is a correlation of activation with calcium mobilisation in the bloodstream form (Rolin *et al.*, 1990), although this may be coincident as the release of VSG is calcium-linked, and indeed all treatments that release VSG result in adenylate cyclase activity (Voorheis *et al.*, 1981, 1982). Experiments with protein kinase C inhibitors suggests a role in both VSG release and adenylate cyclase activation, as the processes occur concomitantly, although evidence shows that they are not obligatorily coupled (Rolin *et al.*, 1996).

2) Putative cell surface receptors: Aside from those discussed above a number of cell-surface proteins have been identified, by the combined techniques of surface labelling and cloning (Pays and Nolan, 1998). The suggested function of a number of these has been as structural proteins, possibly associated with VSG or PARP, depending on the life cycle stage. Other surface proteins have been identified but are as yet not sufficiently characterised for putative functions to be ascribed, particularly in view of the fact that those isolated by cloning have not been found to have significant homology with any known genes (Nolan *et al.*, 1997; Jackson *et al.*, 1993). Those identified by techniques like catalysed surface radioiodination (Gardiner *et al.*, 1983; Jackson and Voorheis, 1990) have not been characterised or the genes isolated.

3) G-proteins: To date, there has been no proof of the existence of functional heterotrimeric G-proteins in *T. brucei*. Putative α and β subunits have been identified (Coulter and Hide, 1995; Lips and Pays, unpublished) that are preferentially expressed in the bloodstream and procyclic form, respectively. There is evidence for both stimulatory and inhibitory G-proteins in *T. cruzi* (Coso *et al.*, 1992; Oz *et al.*, 1994; Eisenschlos *et al.*, 1986) but not *T. brucei*. Control of adenylate cyclase activity, described above, does not appear to be G-protein coupled in *T. cruzi* as agents known to perturb G-protein responsive cyclases do not affect trypanosomal adenylate cyclase activity (Torruella *et al.*, 1986; Eisenschlos *et al.*, 1986). Of the small guanine nucleotide-exchange factor proteins, a homologue of the nuclear Ran family has been described in *T. brucei* (Field *et al.*, 1995). Three homologues of the Rab family of small GTPases that are involved in endocytosis have been isolated in *T. brucei* (Field *et al.*, 1998).

4) Protein kinases: Protein kinases in eukaryotes form a large superfamily, defined as enzymes that use the γ-phosphate of ATP or GTP to form phosphate monoesters on serine/threonine residues (alcohol groups) or tyrosine residues (phenolic groups). They are related in possessing 11 homologous kinase catalysis domains which consist in total of approximately 250-300 amino acids (Hanks, 1991). Phylogenetic analysis based only on the homologous kinase catalysis domains broadly classified kinases into four groups, which can then be further subdivided into families, based on homology of non-kinase catalysis domains, substrate specificities and modes of regulation (Hanks and Hunter, 1995). The four groups are: 1) The ACG group, which includes the cyclic-nucleotide-dependent family, the protein kinase C family and the β-adrenergic family; 2) The CaMK group, which includes the families of protein kinases regulated by calcium and/or calmodulin; 3) The GMGC group, which includes the mitogen activated protein (MAP) kinase family, the cyclin dependent kinase (CDK) family and CDK-related kinases; 4) The 'conventional' protein tyrosine kinase (PTK) group, including the SRC family and the various families of peptide hormone receptor tyrosine kinases (RTKs).

Protein kinases that fall outwith these four groups can usually be classified as defined families in terms of being clearly related to each other by structure/function. It has proven difficult to define broader generalisations that could be used to group these kinases into a larger category (Hanks and Hunter, 1995). One notable type of kinase that falls outwith the four groups are those that display dual specificity i.e. are capable of phosphorylating both serine/threonine and tyrosine residues (Lindberg *et al.*, 1992), hence the designation of the 'conventional' tyrosine kinases above. Dual-specificity kinases are not closely related to the conventional PTKs and map throughout the protein kinase phylogenetic tree (Lindberg *et al.*, 1992). The completion of the *Saccharomyces cerevisiae* genome sequencing project has revealed that budding yeast has 113 kinase genes, corresponding to about 2% of the total number of genes. Although many of the genes discovered fall within the conventional four kinase groups - including cAMP-dependent kinases, Ca^{2+}-calmodulin protein kinases, MAPKs and CDKs - there are no conventional protein tyrosine kinases (Hunter and Plowman, 1997). However, there are dual-specificity kinases and tyrosine phosphatases, as well as some kinases unique to budding yeast (Hunter and Plowman, 1997). The ubiquity of phosphorylation networks in eukaryotes from yeast to mammals strongly suggests that such networks should exist in trypanosomes, with a comparable percentage of the genes of *T. brucei* encoding protein kinases. Furthermore, the kinases involved could be identified as phosphorylation differences and differential kinase activities between life cycle stages or at different points in the cell cycle.

Difficulties in achieving good synchronisation of cultured *T. brucei* (chapter 1.5) has led investigators to concentrate on observable changes between life cycle stages. Phosphorylation differences have been observed in proteins between the long slender and short stumpy bloodstream forms (Aboagye-Kwarteng *et al.*, 1991). Phosphorylation on tyrosine

residues has also been demonstrated (Parsons *et al.*, 1991). Multiple proteins phosphorylated on tyrosine were detected both in the bloodstream and procyclic forms, and the patterns of phosphotyrosine proteins varied during the life cycle. As regards analysis of the differential kinase activities between life cycle stages, multiple kinase species have been detected by use of renaturation kinase assays (Parsons *et al.*, 1993); phosphoamino acid analysis revealed that the phosphorylated residues were serine and threonine, with no phosphotyrosine. However, it should be noted that phosphotyrosine species were previously noted to be of very low cellular abundance (Parsons *et al.*, 1991) and difficult to detect by radiolabelling.

Monoclonal antibodies raised against purified phosphotyrosine proteins has allowed the characterisation of two phosphotyrosine proteins, of 44 kDa and 46 kDa, that are phosphorylated on serine and tyrosine (Parsons *et al.*, 1994). Transformation from the bloodstream to the procyclic form was accompanied by increased tyrosine phosphorylation of both proteins. Fractionation of trypanosome lysates by isoelectric focussing has also been used to characterise autophosphorylating kinases (Hide *et al.*, 1994). This approach detected a family of 60 kDa and 90 kDa serine/threonine protein kinases in bloodstream form lysates; the 60 kDa family was capable of utilising both ATP and GTP whereas the 90 kDa family could only use GTP. Despite the uniformity of molecular weight the isoelectric points within the families varied, suggesting that they represented two sets of isoenzymes (Hide *et al.*, 1994). A protein kinase C activity has also been identified in bloodstream form *T. brucei* (Keith *et al.*, 1990).

Due to the difficulties of obtaining synchronous populations of any particular proliferative stage of *T. brucei*, the focus has necessarily been of the differential kinase activities and phosphorylation profiles between life cycle stages. However, a cell cycle regulated serine/threonine protein kinase activity has been reported (Gale *et al.*, 1994). An 89 kDa protein, named SPK89, previously identified to have activity only in proliferative forms (Parsons *et al.*, 1993) was demonstrated to be specifically induced when stationary-phase procyclics were simulated to re-enter the cell cycle. SPK89 activity mapped predominantly to S phase with some residual activity in G2 (Gale *et al.*, 1994).

Complementary to the biochemical methods is the approach of using degenerate primers to fish for genes. Using degenerate primers to conserved kinase catalysis domains, seven distinct kinase fragments were amplified from cDNA, two of which showed differential accumulation of transcript between life cycle stages (Hua and Wang, 1994). The full-length gene of fragment *TbPK-A1* was isolated and renamed *KFR1*, as it showed significant homology to the *S. cerevisiae* kinases *KSS1* and *FUS3*, and the KFR protein was present at twice the abundance in the long slender bloodstream form as compared to the procyclic form (Hua and Wang, 1994). By analogy with KSS1/FUS3 function in yeast, KFR1 may be implicated in a MAP kinase-like pathway in *T. brucei*. Another of the fragments isolated was used to isolate a full-length gene with significant homology to the *Drosophila* kinase *POLO* (Graham *et al.*, 1997), named *tbPLK*,

which is expressed in both bloodstream and procyclic forms. This appears to be a member of the POLO-like kinase (PLK) family, hence the name; the *PLK*s have been implicated in mitotic control. A slightly different PCR-based approach is to use degenerate oligonucleotides with genomic DNA, isolate the whole gene and then investigate the relative transcript/protein presence between life cycle stages. By this approach, two protein kinase genes named *NRKA* and *NRKB* have been isolated, both of which are constitutively expressed through the life cycle (Gale and Parsons, 1993). *NRKA* and *NRKB* show homology to the *NEK1/NIMA* subfamily over the catalysis domains but are probably not functional homologues as there is no homology over the non-catalytic regions.

Other kinase genes cloned from *T. brucei* include three genes related to protein kinase A (PKA); the genes appear to code for PKA catalytic subunits, and have been named *tbPKAα*, *tbPKAβ* and *tbPKAγ* (Boshart and Mottram, 1997). *TbPKAγ* is expressed constitutively, whereas the other two are expressed to higher levels in the bloodstream forms. A protein kinase C (PKC) activity has been defined biochemically in *T. brucei* (Keith *et al.*, 1990) and a putative PKC gene cloned. The gene encodes a protein predicted to be 97 kDa and to have greatest homology to the unconventional PKCι (Selbie *et al.*, 1993) and PKCλ (Akimoto *et al.*, 1994) isoforms. In addition, a number of genes related to the *CDK* gene family have been isolated and partially characterised (see chapter 1.5). Although tyrosine protein kinase activity has been well documented in *T. brucei* (see above) as yet no tyrosine kinase gene has been isolated.

5) Protein phosphatases: Reversible phosphorylation of proteins is an important mechanism in the cellular control of phosphorylated proteins, as at any time the balance of activity of any particular phosphorylated protein will depend on the balance between the relative activities of kinases and phosphatases that act upon that protein. There are four types of serine/threonine phosphatase identified from a number of species (Cohen, 1989). Based on sequence comparisons, types 1, 2A and 2B are members of the same gene family, whereas type 2C is unrelated (Cohen, 1989). Tyrosine phosphatases are of receptor/transmembrane form (types I-V) with some non-membrane associated forms. Protein phosphatase activity has been documented in *T. brucei* (Walter and Opperdoes, 1982) and three phosphatase catalytic subunits have been isolated by PCR from a *T. b. rhodesiense* cDNA library (Erondu and Donelson, 1991). Two of the genes coded for closely related type 1 subunits, named *TPP1A* and *TPP2A*, the third coded for a type 2A subunit, named *TPP2* (Erondu and Donelson, 1991). *TPP1A* and *TPP2A* have 66% and 40% amino acid identity, respectively, to mammalian phosphatase 1 and 2A, while *TPP2* has 39% and 55% amino acid identity to mammalian phosphatase 1 and 2A. The *in vivo* functions of these phosphatases was indicated by treatment of procyclic-stage cells with okadaic acid, a specific inhibitor of type 1 and 2A phosphatases (Das *et al.*, 1994), in which co-ordination of nuclear and kinetoplast cytokinesis was disrupted. A life cycle stage-

regulated protein tyrosine phosphatase activity has been partially characterised (Bakalara *et al.*, 1995a) and there is evidence for multiple protein tyrosine phosphatase activities in *T. brucei* (Bakalara *et al.*, 1995b).

Although the evidence is fragmentary, it is clear that the basic mechanisms underlying signal transduction in higher eukaryotes are present in trypanosomes. The cryptic nature of some of these mechanisms may in time be revealed to be specific biological adaptions evolved to deal with the parasite's unique biology and life cycle. The greatest success in the field of trypanosome signal transduction has been in the characterisation and cloning of kinase genes, while information on processes such as cAMP and calcium signalling remain elusive, as does the isolation of specific cell-surface receptors. The significant aberrations to date are the failure to clone a tyrosine kinase gene, despite the fact that tyrosine phosphorylation and dephosphorylation have been demonstrated to occur in *T. brucei* (Parsons *et al.*, 1991, 1994; Bakalara *et al.*, 1995a,b) and the puzzling lack of functional G-proteins. While good progress is being made in identifying components of *the T. brucei* signalling machinery, to date not one single path has been mapped.

1.5: Cell cycle and differentiation in trypanosomes

The complete life cycle of *T. brucei* has been described (chapter 1.1). One of the most notable features is the alternation of proliferative and non-proliferative cell cycle forms, with much interest directed to the degree of coupling between progression through the cell cycle and differentiation processes (Matthews and Gull, 1994a,b, 1997). The study of the cell cycle in the proliferative life cycle stages, and of the interplay between the cell cycle and differentiation, has been hampered by the lack of a suitable method of inducing cell cycle arrest to achieve synchronous populations of life cycle stages tractable to *in vitro* culture. Hydroxyurea treatment is a suitable method for many cell types, but has been demonstrated to be unsuitable for *T. brucei* (Brun, 1980; Mutomba and Wang, 1996). Treatment with aphidicolin, an inhibitor of DNA polymerase, is often a useful method of introducing a G1/S block in mammalian cells, but it has also been demonstrated to be unsuitable for synchronisation for *T. brucei* (Mutomba and Wang, 1996). Other chemicals have been reported to interfere with normal cell cycle progression in *T. brucei*: Okadaic acid, a highly specific, reversible inhibitor of the protein phosphatases PP1 and PP2A in mammalian cells (Cohen *et al*, 1990) was tested on procyclic form *T. brucei* (Das *et al.*, 1994). Treatment with okadaic acid produced cells defective in organellar genome segregation and cytokinesis, but not M phase, resulting in multinucleate cells. Neither nuclear or kDNA synthesis appeared to inhibited, suggesting that okadaic acid

acts to inhibit cytokinesis and kinetoplast division without inhibiting DNA replication (Das *et al.*, 1994). The adenosine analogue (S)-9-(-3-hydroxy-2-phosphonylmethoxypropyl)adenine ((S)-HPMPA) was tested on the *T. b. brucei* and *T. b. rhodesiense* bloodstream form (Kaminsky *et al.*, 1998). As trypanosomes are incapable of purine synthesis *de novo*, the expected effect would be an inhibition of DNA synthesis and a resulting block of cells at G1/S and in S phase. This is what was observed, although the effect was specific to the nucleus and kDNA synthesis was not affected. However, synchronisation was not possible as the lag phase to recovery was too long and when cells undergoing cytokinesis were detected the populations were desynchronised (Kaminsky *et al.*, 1998). The CDK inhibitor flavopiridol introduces both a G1/S block and a G2/M block in procyclic *T. b. brucei* (Hassan, 1999). Partial synchrony can be achieved by arresting procyclic-stage cells by growing to stationary phase before inoculating to fresh medium (Gale *et al.*, 1994) or by removal of fatty acid-free BSA or low-density lipoprotein from the culture medium (Morgan *et al.*, 1993, 1996), but it is difficult to achieve acceptable reproducibility by these methods. However, by analysis of single cells and by methods involving mathematical analysis, it has been possible to develop a detailed picture of cell cycle progression in *T. brucei* procyclics and consequently a good range of cytological and morphological markers (Sherwin and Gull, 1989; Woodward and Gull, 1990; Robinson and Gull; 1991; Matthews and Gull, 1994b). Ten distinct morphological stages within the complete cell cycle of the procyclic have been described (Sherwin and Gull, 1989), allowing the exact positioning of any particular cell within the cell cycle. The temporal ordering of major events is as follows and is shown (Fig. 1.5).

1) Probasal body elongation
2) Initiation of a new flagellum by the posterior basal body
3) Generation of two new probasal bodies
4) Extension of the new flagellum, emergence from the flagellar pocket and consequent initiation and extension of the new paraflagellar rod
5) Kinetoplast division
6) Separation of kinetoplasts by basal body segregation
7) Nuclear mitosis
8) Cytokinesis

The exact timing of stages of the events above (occurring within a cell cycle typically lasting 8.5 hours) have been described as points within an idealised unit cell cycle, a useful concept (Sherwin and Gull, 1989). Flagellar growth, DNA replication and basal body replication appear to be highly co-ordinated, with a very high level of spatial ordering. This would be expected in a highly elongate cell with a strict requirement for correct segregation of a number

of single organelles, the nucleus, kinetoplast, mitochondrion and flagellum (Sherwin and Gull, 1989; Turner, 1992).

Determination of how the classical picture of four cell cycle phases - G1, M, G2, S - maps around the observance of these morphological events has been investigated (Woodward and Gull, 1990), in particular with reference to kinetoplast DNA replication ('S_K') and nuclear DNA replication ('S_N'). By monitoring 5-bromo-2-deoxyuridine (BRDU) incorporation to DNA with use of a specific monoclonal antibody, and use of a monoclonal antibody that recognised the basal bodies, three periods were defined: Period C, defined as cytokinesis for the nuclear cycle including post-M phase and cleavage; period D, defined as the period when kinetoplast division is occurring (demonstrated to be achieved by basal body movement: Robinson and Gull, 1991); and period A, defined as the period of time when the predivision cell contains two kinetoplasts. Determination of these periods was important as the period between completion of DNA replication and cell division is more complex in *T. brucei* than in most cells (Woodward and Gull, 1990). The timings of these periods are shown expressed as points in an idealised unit cell cycle (Fig. 1.6).

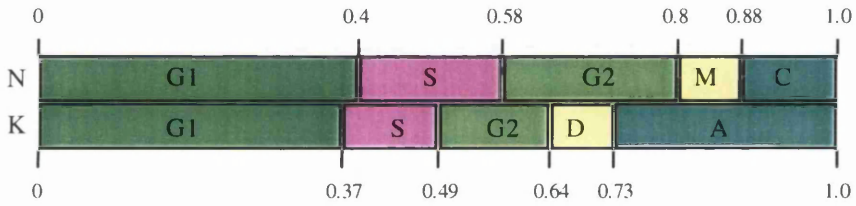


Fig. 1.6: Relative temporal order of events in the DNA cycles of *T. brucei*
The timing of events of the nuclear and kinetoplast cycles, expressed as points within an idealised unit cell cycle. N - nuclear cycle; K - kinetoplast cycle; S - period of DNA replication; M - nuclear mitosis; D - kinetoplast division; C - cytokinesis period for the nuclear cycle (including post-M phase and cleavage); A - period of time in which the predivision cell contains two kinetoplasts.

S_N is longer than S_K by approximately 30 mins, although this may simply reflect the proportionally lesser quantity of kDNA to replicate. However, $G2_K$ is entered about 45 mins before $G2_N$, therefore the replicative phases overlap, showing a non-co-ordinate start and finish to the two S phases. The termination of S_K before S_N argues for intra-organelle control of S phase i.e. S_K can terminate in a cellular environment where S_N continues, leading to a 'licensing factor' theory (Blow and Laskey, 1988). In this, the hypothetical factor is associated with DNA until it is displaced by replication, marking DNA as replicated and thus preventing more than one S phase per cell cycle. Period D represents the portion of the overall cell cycle in which the kinetoplast has an elongated, bilobed appearance. Periods C and A are lengthy, representing the

time required for the post-M phase nuclei and kinetoplasts to migrate to the appropriate positions for division. Interestingly, the relative positions of D and M vary according to species. For example, for *Crithidia luciliae* M occurs before D (Van Assel and Steinert, 1971). This may be a function of morphology as unlike the highly elongate *T. brucei*, *Crithidia* cells are more rounded and the nuclei, kinetoplasts and basal bodies assume a bilateral symmetry soon after organelle duplication. By contrast, in *T. brucei* the post-duplicative arrangement is more tandem, requiring extensive reorganisation to produce bilateral symmetry (Woodward and Gull, 1990). The close temporal association of DNA replication and basal body replication as well as precise organelle spatial positioning suggests a point where synchrony of organelle cycling might be achieved (Woodward and Gull, 1990; Matthews and Gull, 1994a,b; Robinson *et al.*, 1995).

Given the nature of the parasite's periodic cell cycle arrest and proliferation, and the fact that two of these arrested forms - the short stumpy and the metacyclic - are the infective forms that facilitate transmission between vectors and which begin differentiation to the next life cycle stage immediately after transmission, the next logical question is: To what extent, if any, is there interplay between the cell cycle and differentiation? Most studies have dealt with the differentiation most tractable to study, that from the short stumpy form to the procyclic form. This differentiation - often referred to as transformation - is coincident with re-entry to the cell cycle and progression through one full turn of the cell cycle, although the two processes can be uncoupled (Matthews and Gull, 1994b). The two most obvious markers for the transformation is the loss of the VSG coat which characterises the mammalian bloodstream forms and the appearance of the procyclic-specific procyclin protein, or PARP (procyclic acidic repetitive protein), as the dominant coat protein (Roditi *et al.*, 1989; Ziegelbauer *et al.*, 1993). Analysis of VSG loss and PARP appearance in the transformation of bloodstream to procyclic forms supported the hypothesis that there is a window in the cell cycle in G1 (or possibly G0) where the parasite is responsive to the differential trigger(s) that induce differentiation to the procyclic form, and that stumpy form cells are arrested in that window whereas slender form cells are not (Matthews and Gull, 1994a,b). Therefore the general picture is of the appearance of a mechanism that tends to arrest cells in the G1/0 window as the parasitaemia progresses, leading to cell cycle arrested cells that undergo morphological and gene expression changes. Transfer to the tsetse fly and exposure to induction factor(s) releases the cell cycle arrest and cells re-enter the cell cycle and concomitantly complete the differentiation to the procyclic form. The nature of the induction factor(s) *in vivo* are not understood. The commitment to differentiation has also been examined (Matthews and Gull, 1994b). This provided evidence that between 0-4 hours post-induction an irreversible commitment to differentiation is made. Drug treatment showed that DNA synthesis could be halted but procyclic-specific marker proteins still appeared; therefore cell cycle progression and differentiation are coincident processes that

can be uncoupled, at least after the induction of differentiation (Matthew and Gull, 1994b).

Analysis of the relative importance of the two factors in the *in vitro* transformation from stumpy to procyclic form - temperature shift and the presence of *cis*-aconitate - has been addressed (Matthews and Gull, 1997). The presence of *cis*-aconitate was essential for the triggering of all the differentiation changes - both antigenic markers like VSG loss and PARP appearance, as well as morphological changes - and was independent of the temperature. By contrast, re-entry to a proliferative cell cycle could only be triggered by the combined action of *cis*-aconitate and shift from 37°Cto 27°C (Matthews and Gull, 1997). The mechanism of temperature sensitivity is unknown, but this data provided evidence that in the absence of artificial drug treatment differentiation and cell cycle re-entry and progression are separable pathways, which are nevertheless *in vivo* precisely coincident processes controlled by cellular commitment in parallel, by unknown factor(s) and temperature shift (Matthews and Gull, 1997; Matthews, 1999).

The mechanism of induction of long slender bloodstream to stumpy form transformation has been studied in detail (Vassella and Boshart, 1996; Vassella *et al.*, 1997). The pleomorphic long slender bloodstream form, above a certain density, exhibits cell cycle arrest and transforms to the stumpy form (Vassella and Boshart, 1996). It was demonstrated that this was not due to the depletion of essential culture medium factor(s) or to cell-cell contact signalling, but to a stumpy inducing factor (SIF) that was accumulating in the medium, that was heat-stable and time-stable (Vassella *et al.*, 1997). Slender cells arrested by SIF have 1 nucleus and 1 kinetoplast, by cytometry were demonstrated to be in G1 (or G0), and differentiated synchronously to the procyclic form with the same kinetics as stumpy cells purified from whole blood. The cell cycle arrest appears to precede the phenotypic changes. The monomorphic strain MITat 1.2 did not respond to SIF, although conditioned medium from MITat 1.2 culture could transform pleomorphic strains, therefore suggesting the existence of a specific signalling pathway that couples density sensing to differentiation. Use of cAMP agonists and phosphodiesterase inhibitors demonstrated the involvement of the cAMP signalling pathway (Vassella *et al.*, 1997). At present it is not known whether SIF is a single or multifactor signal, a specifically synthesised pheromone-like factor or a catabolite. *T. brucei* is known to contain a family of unconventional adenylate cyclases (Alexandre *et al.*, 1990; Painavoine *et al.*, 1992; Alexandre *et al.*, 1996) therefore the scenario of SIF as a specific ligand for an adenylate cyclase, activating a cAMP pulse that initiates cell cycle arrest, is plausible. Unfortunately, little is known about the downstream effectors of cAMP signalling in trypanosomatids, although in the development of the stumpy form immediate effects downstream of the cAMP pulse are probably mediated by protein kinase A. Analysis of stage-regulated and differentiation-enriched transcripts in the transformation from the stumpy to the procyclic form has isolated two previously identified early transcripts - PARP and histone H2B - and two novel transcripts

(Matthews and Gull, 1998). One encoded a protein with similarity to the *Leishmania major* protein LACK (Leishmania activated protein kinase C receptor; Mougneau *et al.*, 1995; Julia *et al.*, 1996), the other encoded a protein with significant homology to the polypeptide sequence derived from purification of the *Crithidia fasciculata* cytochrome c oxidase complex (Matthews and Gull, 1998). Rapid downregulation of VSG and upregulation of PARP has previously been characterised (Pays *et al.*, 1993) as has the temporal expression of histone H2B (Ersfield *et al.*, 1996).

The detailed definition of progression of *T. brucei* through the cell cycle, at least in the procyclic form, and identification of good morphological and cytological markers, means that study of how cell cycle and life cycle interplay will be greatly facilitated. The definition of nuclear and kinetoplast cycles with at least some degree of co-ordination and some evidence of intra-organelle control raises interesting questions about the molecular mechanisms underpinning cell cycle progression. The existence of the cell cycle arrested stages and the demonstration that in the stumpy to procyclic transformation, cell cycle and differentiation can be uncoupled, raises questions about the tsetse to mammal transition. How are the nuclear and kinetoplast cycles co-ordinated and how does the cell achieve intra-organelle control? Does the proventricular to metacyclic differentiation also depend on a specific induction factor, and can cell cycle be decoupled from differentiation in this case? What are the signalling pathway(s) involved in a cAMP-mediated cell cycle arrest, and what are the effectors that change patterns of gene expression and produce phenotypic changes? The unique biology of the African trypanosome suggests that the answers to this question may be as unexpected as they are interesting.

While little is known about the signalling pathways involved in cell cycle control and differentiation in *T. brucei*, an increasing number of genes homologous to the genes that underpin cell cycle control in higher eukaryotes are being isolated from trypanosomatids. The first *CDK* gene isolated from a kinetoplastid was the *cfCRK* gene from *Crithidia fasciculata* (Brown *et al.*, 1992). This encoded a CDC2-like kinase that contained insertions of 66 and 79 amino acids between kinase catalysis domains VIb and VII, and X and XI, respectively; these insertions had not been found in *CDC2*-related genes from any other organism. A number of these *CDC2*-related kinase (*CRK*) genes have been isolated by PCR from *T. brucei* and *Leishmania mexicana*. The *lmmCRK1* gene was isolated during an unrelated experiment in screening for proteinases (Mottram *et al.*, 1993). The *lmmCRK1* gene encoded a 34 kDa protein containing residues at homologous positions to those residues known to be critical in the regulation of CDKs, such as the Tyr15 and Thr161 residues (Pines, 1995). In addition, it contained all eleven required kinase catalysis domains and also possessed a modified PSTAIR box, containing only one amino acid substitution compared to human CDK1 (Mottram *et al.*, 1993; Mottram, 1994). LmmCRK1 was demonstrated to possess histone H1 kinase activity,

showing stage-specific regulation: The kinase was active in promastigotes, whether in logarithmic growth or stationary phase, but not active in the proliferative amastigote stage, suggesting a role in differentiation rather than cell cycle progression (Mottram *et al.*, 1993). It was unable to complement an *Schizosaccharomyces pombe CDC2* temperature-sensitive mutant. As the protein was detected to equivalent levels in all three life cycle stages, the evidence pointed to post-translational regulation of activity, consistent with the control of CDK activity observed in higher eukaryotes (Solomon *et al.*, 1993; Mottram *et al.*, 1993; Pines, 1993, 1995; Morgan, 1995). *LmmCRK1* is essential, as judged from chromosome plasticity resulting from targeted gene disruption to generate null mutants (Mottram *et al.*, 1996). The *tbCRK1* gene was isolated by library screening, using a fragment produced by PCR from a *T. brucei* cDNA library using degenerate primers. TbCRK1 was found to have 73% identity to lmmCRK1 (Mottram and Glasssmith, 1995). The tbCRK1 protein possess the required kinase catalysis domains and residues involved in phosphorylation, and also has a degenerate PSTAIR box, with 2 amino acid substitutions. Ectopic expression of *tbCRK1* from an episomal vector can at least to some degree fill in the function of *lmmCRK1*, as the leishmanial gene can be deleted in this case (Mottram *et al.*, 1996). Promastigotes null for *lmmCRK1* with *tbCRK1* expressed episomally grow with normal kinetics as compared to wild-type, although when induced to transform to amastigotes growth kinetics do not match that of wild-type promastigotes transformed to amastigotes (Mottram *et al.*, 1996). The inference is that *tbCRK1* will perform a homologous function in trypanosomes and that it will also be an essential gene.

The *lmmCRK3* gene was also isolated by the approach of library screening following amplification of a gene fragment by PCR from cDNA. The lmmCRK3 protein has six amino acid substitutions in the PSTAIR box as compared to human CDK1, and is expressed throughout the life cycle (Grant *et al.*, 1998). Relative transcript abundance between life cycle stages for promastigote:metacyclic:amastigote is 5:1:1, although this does not accurately reflect protein abundance and kinase activity; lmmCRK3 histone HI kinase activity is highest in promastigotes, high in amastigotes and very low in the cell cycle arrested metacyclics (Grant *et al.*, 1998). This corresponded well to the previously documented histone HI activity profile of SUC-binding *CDC2*-related kinase (SBCRK; Mottram *et al.*, 1993). SBCRK was defined as a kinase activity that bound the *S. pombe* protein $p13^{SUC1}$. A leishmanial homologue of SUC1, $p12^{CKS1}$, was isolated (Mottram and Grant, 1996) that showed 70% identity to human CKS1 and 44% identity to *S. pombe* SUC1. The lmmCKS1 protein was shown to bind a strong histone HI kinase activity from promastigotes and amastigotes but not metacyclics, mirroring the profile of SBCRK and lmmCRK3 (Mottram and Grant, 1996; Grant *et al.*, 1998). Depletion of lmmCRK1 did not change this lmmCKS1-associated activity, indicating that it was not responsible for SBCRK (Mottram and Grant, 1996). Depletion of lmmCRK3 lowered the lmmCKS1-associated SBCRK activity by over 50%, indicating that lmmCRK3 is the major kinase subunit of SBCRK (Grant *et*

al., 1998). As SBCRK is active in the proliferative stages of *Leishmania* but not the nonproliferative, the data suggests that *lmmCRK3* is a good candidate for a functional *CDK1/CDC2* homologue. Deletion of *lmmCRK3* gave similar results to attempted deletion of *lmmCRK1*, showing it to be an essential gene, as would be expected (Hassan, 1999). The *tbCRK3* gene was isolated in the same manner as *lmmCRK3*, by PCR and library screening (Mottram and Glasssmith, 1995). It encodes a kinase with 78% identity to lmmCRK3, and like lmmCRK3 it contains 6 amino acid substitutions in the PSTAIR box and was unable to complement a yeast mutant (Mottram, 1994; Grant *et al*, 1998). Targeted deletion of *tbCRK3* generated a very small population of cells that appeared to be severely compromised in their ability to faithfully segregate organelles during cytokinesis (Glasssmith, 1997). Spindle formation was badly affected, with uneven distribution of nuclei and kinetoplasts leading to a variety of daughter cell types i.e. 2k 1n, 1k 2n, 2k 2n, in addition to anuclear cells 1k 0n (termed 'zoids').

Library screening at low stringency with the *lmmCRK1* gene isolated the *tbCRK2* gene (Mottram and Glasssmith, 1995). In common with the expanding family of *CRK* genes, *tbCRK2* also encoded a kinase with the conserved features mentioned above, although the PSTAIR box contained four substitutions. It also possessed an N-terminal extension of 41 amino acids as compared to tbCRK1, human CDK1 and CDK2. The *tbCRK2* gene was found to be unable to complement an *S. pombe CDC2* mutant (Mottram, 1994). A fifth member of the family, *tbCRK5*, was isolated as part of the EST sequencing project and shows highest identity to a mammalian meiosis-specific kinase (Van Hellemond and Mottram, unpublished). The function of tbCRK5 is unknown, although it is not localised to the nucleus but apparently associated with elements of the cytoskeleton.

In addition to *CDK* homologues, three cyclin homologues have been isolated from *T. brucei*. The *CYC1* gene was isolated by PCR using degenerate primers (Affranchino *et al.*, 1993) and described as a mitotic-type cyclin. However, other data has suggested a role in the transformation from the short stumpy bloodstream to procyclic form (Hua *et al.*, 1997). Two further cyclins, *CYC2* and *CYC3*, were isolated through complementation of a *Saccharomyces cerevisiae* G1 cyclin mutant by a *T. brucei* cDNA library (Van Hellemond *et al.*, 1999). Putative cyclins have not been isolated from any other trypanosomatid to date, although the *T. cruzi* CRK1 (which is homologous to lmmCRK1 and tbCRK1) has been demonstrated to bind mammalian cyclins A, D3 and E *in vivo* when expressed in COS-7 cells (Gomez *et al.*, 1998). CYC2 has been demonstrated to have a functional association with tbCRK3 (Van Hellemond *et al.*, 1999), the first evidence of a functional CRK-cyclin pair in trypanosomatids.

As the trypanosomatid equivalent of CDKs, the CRKs are likely to represent the endpoint of signalling pathways, at least as far as cell cycle control is concerned. In addition, they may represent integration points for cross-talk between the cell cycle and differentiation

events concerned with life cycle progression. In this study molecular genetics and biochemistry were used to attempt to analyse the function of *CRK2* and to further analyse the function of *CYC1*. In addition, in this study the isolation and partial characterisation of a novel trypanosome *CRK*, *CRK4*, is reported; *tbCRK4* has a high degree of sequence identity to the *C. fasciculata CRK* gene.

1.6: Molecular genetic analysis

The life cycle stages of *T. brucei* that are amenable to molecular genetic analysis are diploid. African trypanosomes are the only kinetoplastid in which sexual genetic exchange has been demonstrated to take place (Tait, 1990), and that appears to occur at low frequency during passage through the tsetse fly. The molecular karyotyping of different strains is in progress; the characterisation of chromosomes through increasing abundance of syntenic molecular markers (Melville *et al.*, 1998) is providing information on the genomic arrangement of *T. brucei*. Characterisation of kinetoplastid chromosomes recently took a step forward with the complete sequencing of the *Leishmania major* Friedlin chromosome 1 (Myler *et al.*, 1999), which showed an unusual and highly distinctive distribution of protein coding genes; the first 29 genes were all located on one DNA strand, and the following 50 genes were located on the other strand. This is consistent with the polycistronic arrangement seen in smaller regions (Swindle and Tait, 1996) but is the first known case of a head-to-head polycistron arrangement that spans an entire chromosome with absolute strand polarity. It will be interesting to see if this is a general feature of *Leishmania*, and whether other kinetoplastids - *T. brucei* included - also display this type of gene organisation.

As the life cycle stages of *T. brucei* that can be cultured are diploid, the well-established reverse genetic approach of studying gene function by targeted gene disruption is a feasible tool. Stable transfection of trypanosomes with the neomycin phosphotransferase (*NEO*) gene into the tandem tubulin locus was demonstrated to yield G418 resistant clones (ten Asbroek *et al.*, 1990). Transfection was more efficient with linearised DNA than with circular constructs, with longer regions of the tubulin locus in the targeting construct, and appeared to occur by homologous recombination (ten Asbroek *et al.*, 1990). Other data supported this (Lee and Van der Ploeg, 1990; ten Asbroek *et al.*, 1993). In trypanosomes episomal constructs are not well maintained and there appears to be a strong preference for integration (ten Asbroek *et al.*, 1993) although later studies generated more stable autonomously replicating episomes (Patnaik *et al.*, 1993; Patnaik, 1997) in which it was shown that episomes were only maintained if they contained transcriptional promoters. Analysis of targeted gene disruption (Eid and Sollner-Webb, 1991a,b) confirmed that gene replacement occurs by homologous recombination, and

that recombination was initiated by a crossover event between the linearised end of the transfecting DNA and a homologous region in the host genome. There are, however, problems with this straightforward approach resulting from the general plasticity of parasite genomes. It has been demonstrated in *Leishmania* that preservation of essential genes often occurs by the formation of genomic tetraploids, or aneuploids that are trisomic for the locus in question (Cruz *et al.*, 1993). In addition, relatively large chunks of chromosomes can apparently be translocated into other chromosomes (Dumas *et al.*, 1997). There is relatively little data on similar processes in trypanosomes. A number of genes suitable for use as selectable markers have been described, most of which were first used as selectable markers in *Leishmania.* Those used in these studies are the hygromycin phosphotransferase (*HYG*) gene, which confers resistance to the antibiotic hygromycin B (Lee and Van der Ploeg, 1991), the bleomycin (*BLE*) gene that confers resistance to the DNA-binding aminoglycoside phleomycin (Souza *et al.*, 1994), the puromycin acetyltransferase (*PAC*) gene which confers resistance to the glycopeptide puromycin (Freedman and Beverley, 1993), and the streptathricin acetyltransferase (*SAT*) gene that confers resistance to the antibiotic nourseothricin (Joshi *et al.*, 1995).

There are no easily manipulated and maintained episomal constructs for use of ectopic gene expression in *T. brucei*, as there is in *Leishmania* (LeBowitz *et al.*, 1990; Kelly *et al.*, 1992). However, systems utilising stable transfection via homologous recombination, with inducible expression of the subject gene and resistance marker allowing ectopic expression and selection, have began to address this issue in *T. brucei* (Wirtz and Clayton, 1995). The pHD vector series initially consisted of a modified PARP promoter that contained the operator region of the *E. coli* tetracycline resistance operon along with the appropriate 5' and 3' processing elements, linked to a copy of the *BLE* gene. The second component of the system was a *T. brucei* cell line that contained a constitutively expressed copy of the tetracycline repressor protein gene integrated into the α/β-tubulin gene array (Wirtz and Clayton, 1995). Following cloning of the subject gene into the pHD vector, transfection was followed by selection in the presence of phleomycin and tetracycline; the tetracycline bound the repressor protein, releasing repression of the modified promoter and allowing transcription of the subject gene and the *BLE* gene, with the phleomycin then selecting for stable recombinants. However, this system does not allow for the expression of toxic gene products. A later series of vectors (Biebinger *et al.*, 1997) corrected this fault by decoupling the control of expression of the resistance marker (rendering expression constitutive) from the inducible expression of the subject gene. In addition, the selectable marker was changed from *BLE*, which operates in a stochiometric binding fashion, to *HYG*, which operates enzymatically, thus helping to reduce background (Biebinger *et al.*, 1997). This system therefore provides an ideal tool for the expression of modified dominant negative alleles, and for overexpression of native alleles. In addition, expression of epitope-tagged dominant negative or wild type alleles may aid in elucidation of

gene function (example: Mottram *et al.*, 1996; Grant *et al.*, 1998). In parallel with a detailed description of the trypanosome cell cycle, an increase in the cytological and morphological markers available, and progress in the understanding of differentiation processes in trypanosomes and how these are linked to cell cycle progression (section 1.5), such systems will allow for greater dissection of the precise roles of trypanosome *CRK* and cyclin genes, particularly in cases where targeted deletion cannot produce a null genotype/phenotype.

The aims of this project were an investigation of the mechanisms of cell cycle control in *T. brucei*, specifically:

1) Analysis of *tbCRK2* gene function in procyclic-stage cells by targeted gene disruption.
2) Biochemical characterisation of tbCRK2 through the use of specific antibodies and the tetracycline inducible expression system.
3) Analysis of the *tbCYC1* gene function in procyclic-stage cells by targeted gene disruption.
4) Biochemical characterisation of tbCYC1 through the use of specific antibodies.
5) Isolation and characterisation of the *tbCRK4* gene and tbCRK4 protein.

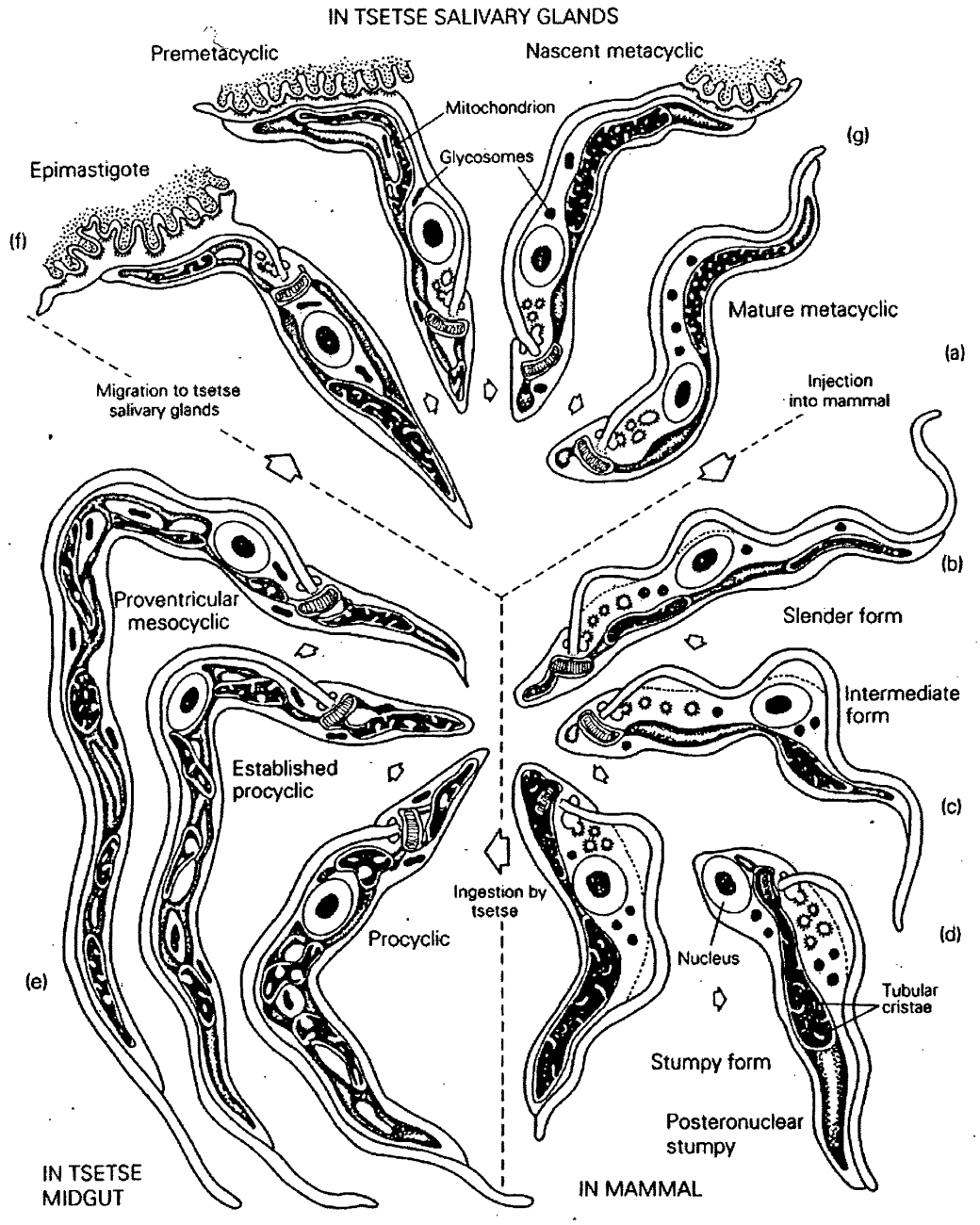


Fig. 1.1: Life cycle of *T. brucei*

The complete life cycle of *T. brucei*. Metacyclics are injected to the vascular system of a mammal (a), differentiate to the long slender form and divide by binary fission (b), then transform via the intermediate form (C) to the stumpy form (d) which are ingested by the tsetse during a bloodmeal. In the tsetse midgut the parasite divides as the procyclic form, transforms to the nonproliferative proventricular mesocyclic (e) and then migrates to the salivary glands. Here, the epimastigote form is assumed (f) and further division occurs before transformation to the mature, mammalian-infective metacyclic form (g). Taken from Cox, 1983.

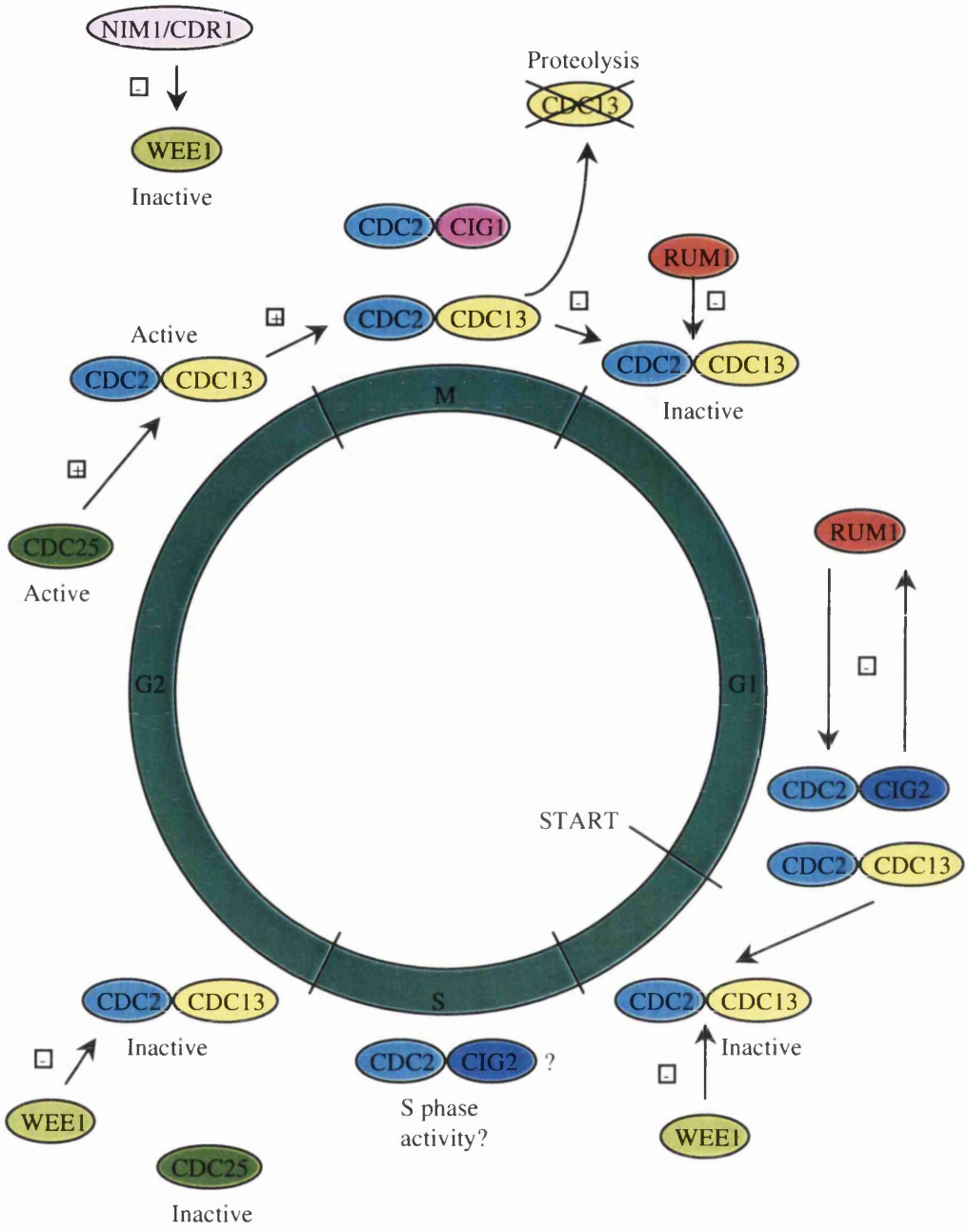


Fig. 1.2: Schematic of the *Schizosaccharomyces pombe* cell cycle

RUM1 inhibits any CDC2-CDC13 still present after cytokinesis, and continues to inhibit CDC2-CDC13 complexes as they form, as CDC13 levels rise throughout G1. CIG2 levels rise, and forming CDC2-CIG2 complexes are inhibited by RUM1. At a threshold point near START, CDC2-CIG2 acts to inhibit RUM1; the cell passes START and progresses through S phase, with CDC2-CDC13 held inactive by WEE1 Tyr15 phosphorylation. In late G2 the CDC25 phosphatase is activated, and in turn activates CDC2-CDC13, while WEE1 is inactivated. CDC2-CIG2 may have functions in S phase. CDC2-CIG1 has as yet undefined functions in M phase.

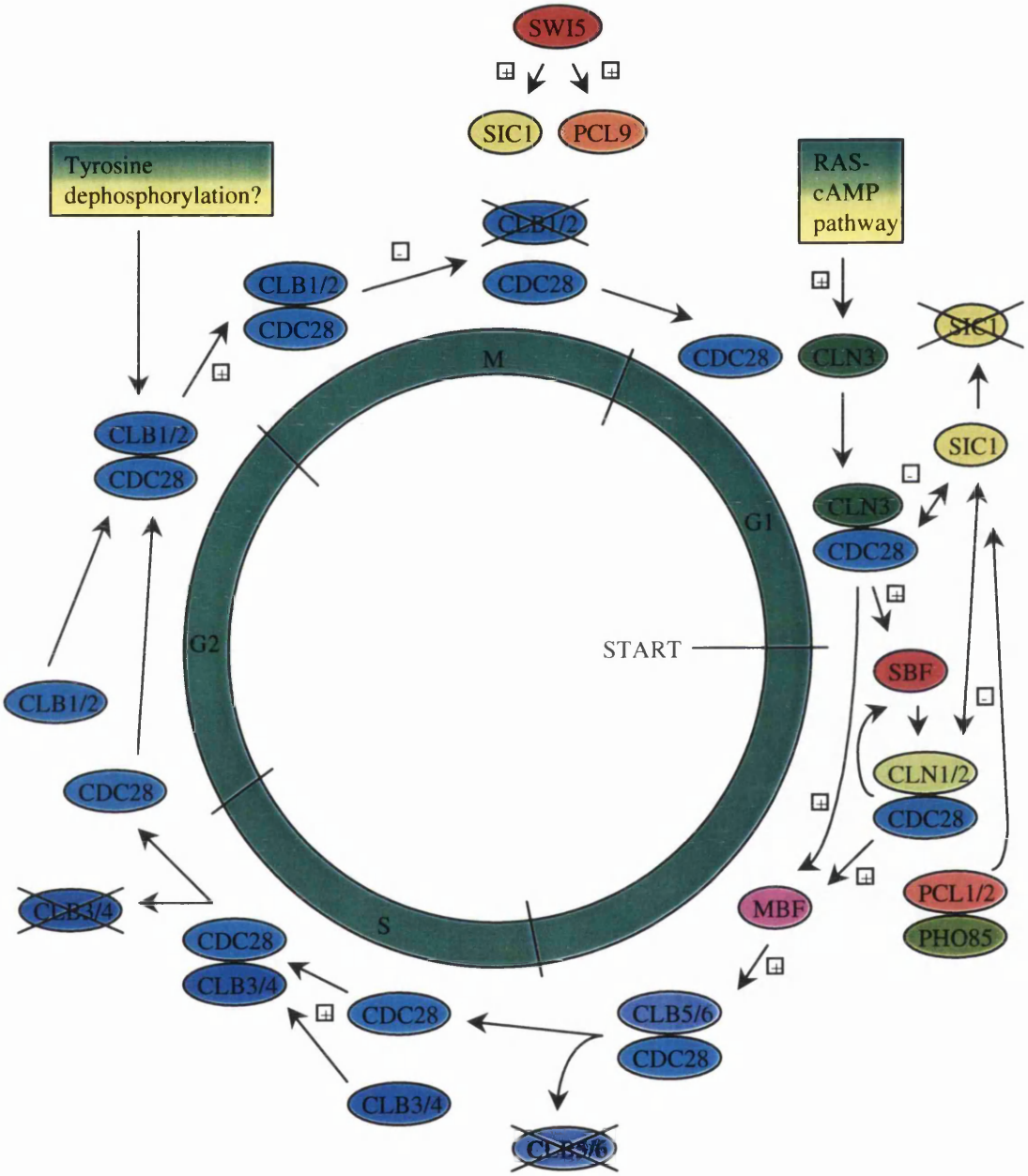

SWI5
SIC1
PCL9
Tyrosine dephosphorylation?
RAS-cAMP pathway
CLB1/2
CDC28
CLB1/2
CDC28
CDC28
CLN3
SIC1
M
SIC1
CLB1/2
CDC28
CLN3
CDC28
G1
START
G2
CLB1/2
SBF
CDC28
CLN1/2
CDC28
CLB3/4
MBF
PCL1/2
PHO85
S
CDC28
CLB3/4
CDC28
CLB5/6
CDC28
CLB3/4
CLB5/6

Fig. 1.3: Schematic of the *Saccharomyces cerevisiae* cell cycle

Near the end of M phase the CLB1/2 cyclins are destroyed and the inhibitory protein SIC1 and the cyclin PCL9 are induced by SWI5. CDC28-PCL9 is greatest at the M-G0 transition. Growth signals induce CLN3 expression but SIC1 inhibit the activity the CDC28-CLN3 complexes that form. After a threshold has been reached, CDC28-CLN3 induce CLN1/2 via SBF, and CDC28-CLN1-3 initiates SIC1 destruction. PHO85-PCL1/2 may also do this. Later in G1 CLN is downregulated and CDC28-CLN1/2 induce CLB5/6 via MBF. CDC28-CLB5/6 mediate entry to S phase, and as CLB5/6 expression wanes, CLB3/4 rises and CDC28-CLB3/4 moves the cell through S phase. CLB1/2 in induced in G2, and active CDC28-CLB1/2 drives the cell into M phase. CLB1/2 destruction at the end of M phase resets the cell cycle. CDC28 is phosphorylated on Thr167 by CAK1/CIV1 for activity regardless of cyclin partner.

Table 1: Mammalian CDKs and cyclins.

Cyclin	Associated CDK	Function
A	CDK1, CDK2	S-phase entry and progression, M phase
B1-3	CDK1	M phase entry and exit
C	CDK8	Transcriptional regulation, G1-S transition
D1-3	CDK4,CDK6	G1-S transition
E	CDK2	G1-S transition
F	?	Unknown
D2	CDK5	DNA damage response?
H	CDK7	CDK activation and transcriptional regulation
I	?	Unknown
K	CDK9	Transcriptional regulation?
T1-3	CDK9	Transcriptional regulation

Proteolytic machinery
Mitogenic signals
G1
S
G2
M
Restriction point (R)
Active
Inactive
CDK4/6
CDK2
CDK1
CDK3
CDK7
pRB
E2F
DP1
CDC25
PP1
WEE1
MYT1
A
B
D
E
H

Fig. 1.4: Schematic of the metazoan cell cycle

Cells in early G1 have no mitotic-type cyclins, hence CDK1 is inactive. Upon mitogenic stimulation, cyclin D is upregulated, binds CDK4/6 and drives the cell past the restriction point via pRB/E2F-mediated induction of late G1 and S phase genes, including cyclin E and cyclin A. As the restriction point is passed cyclin D is destroyed, providing rapid feedback to limit the influence of CDK4/6-cyclin D. CDK2-cyclin E complexes form and participate in a positive feedback loop to maintain pRB inactivity and hence E2F function. CDK3 is active at the G1-S transition. Cyclin E is destroyed as the cell enters S phase, and CDK2 instead complexes with cyclin A, which is synthesised in late G1. CDK2-cyclin A negatively regulates the activity of E2F. Rising levels of cyclin B mean the CDK1-cyclin B complexes begin to form at the end of S phase and in G2, and CDK7-cyclin H phosphorylates Thr161, but the complex is held inactive by Thr14 phosphorylation by MYT1 and Tyr15 phosphorylation by WEE1. The CDC25 phosphatase is held inactive during much of G2, possibly by the action of phosphatase PP1. In late G2 both MYT1 and WEE1 become hyperphosphorylated and inactivated. PP1 activity falls in parallel with a sharp increase of phosphorylation of CDC25, and the active CDC25 dephosphorylates CDK1-cyclin A/B complexes on Thr14 and Tyr15, activating the complexes and driving the cell into M phase. Cyclin A and, later, cyclin B are destroyed during M phase, and both newly formed daughter cells re-enter G1 with inactive CDK1.

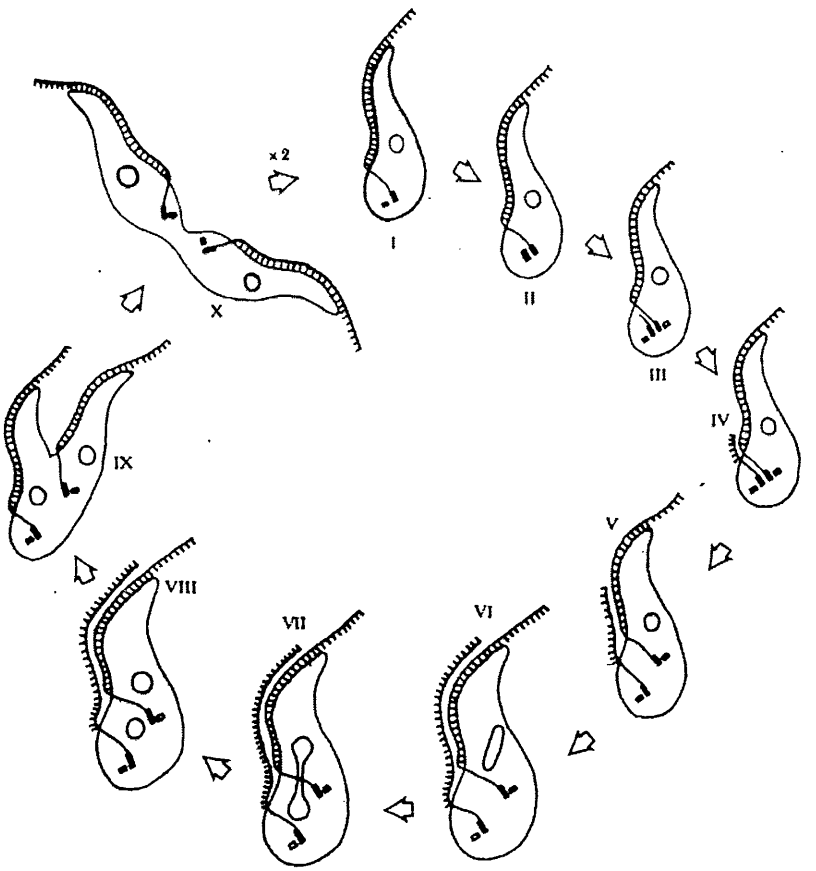


Fig. 1.5: Schematic representation of the *T. brucei* cell cycle

Diagram of the morphological changes occurring in the cytoskeleton of procyclic form *T. brucei*. The stage I cell is newly formed. In the stage II cell probasal body extension has occurred, and in the stage III cell two new probasal bodies have formed and elongation of the daughter flagellum has begun. In the stage IV cell the daughter flagellum has emerged from the flagellar pocket, and in the next stage the basal bodies begin to move apart before the daughter flagellum is full-length. Mitosis has initiated in the stage VI cell, and by stage VII M phase has progressed to the point where the nucleus is elongate. M phase is complete in the stage VIII cell; one nucleus lies between the two basal bodies, the other anterior to both basal bodies. In the stage IX cell the cleavage furrow has initiated and by the final defined stage, X, fission is almost complete, with the two daughter cells lying opposed to each other and joined only by their extreme posterior regions. Taken from Sherwin and Gull, 1989.

CHAPTER 2: MATERIALS AND METHODS

2.1 General materials and methods

2.1.1: Growth and handling of parasites

T. b. brucei Swiss Tropical Institute of Biology strain 247 (STIB 247) was used throughout, except where specifically stated otherwise. For the *CRK2* expression experiments the strains used were transgenic clonal isolates derived from Antat 1.1 ('360' cell line) and STIB 427 Mitat 1 ('449' cell line), for the pHD430 and pHD675 vector work, respectively.

Procyclic form STIB 247 were grown in SDM-79 medium (GibcoBRL) supplemented with 10% foetal bovine serum (GibcoBRL) at 27°C (Brun and Schonenberger, 1979). Subpassage was typically done by inoculating fresh medium to a cell density of 5x 10^5 cells ml^{-1}. Antibiotics used for the selection of transgenic procyclics were as follows, used at the final concentrations in culture: Hygromycin (Sigma) at 50 $\mu g\ ml^{-1}$, phleomycin at 20 $\mu g\ ml^{-1}$, nourseothricin (W. Werner, HKI Jena) at 50 $\mu g\ ml^{-1}$, puromycin (Sigma) at 10 $\mu g\ ml^{-1}$. Tetracycline HCl (Sigma) used for the induction of expression of cloned genes from the pHD constructs (Wirtz and Clayton, 1995; Biebinger *et al.*, 1997) was used at various concentrations ranging from 0.2 ng ml^{-1} to 10 ng ml^{-1}. Antibiotics were kept at 4°C once made in solution and were discarded after one month.

Harvesting of the procyclic form was done by centrifugation at 600*g* for 10 min at room temperature. Culture medium was aspirated and the cell pellet resuspended in an equivalent volume of 1x phosphate buffered saline (PBS) pH 7.2. Subsequently, the cells were spun as before and the PBS decanted. For some applications the PBS wash was repeated to ensure lower contamination of protein from the culture medium.

Purification of bloodstream form parasites was performed according to the protocol of Lanham, 1968, 1970. Following test bleeds and cell counts, to calculate the parasitaemia and thus determine whether cell type was predominantly of the long slender bloodstream form or enriched in the short stumpy bloodstream form, the mice were killed and exsanguinated. Whole blood was spun at 300*g* for 12 min at 4°C, and the serum was discarded. The trypanosome pellet, sitting on top of the erythrocytes, was then carefully resuspended in 2-3 ml of phosphate saline glucose buffer (PSG) pH 8.05, so as to disturb the erythrocytes as little as possible. The resuspended trypanosomes were transferred to a fresh tube and applied to a pre-prepared column of 10x bed volume of the sample of DE52 anion exchange resin, pre-equilibrated with PSG pH 8.05. The column was washed with PSG pH 8.05 until the eluate ran cloudy with eluted trypanosomes, which was collected. Samples were removed for cell counts and slides,

and the eluted cells were spun at 500*g* for 12 min at 4°C. Following aspiration of the buffer, cell pellets were resuspended in 2-3 ml of PSG and aliquoted to eppendorfs; these were spun briefly at maximum speed in a bench microfuge to pellet the cells, and following aspiration of supernatant the pellets were kept at -80°C until use.

Cell counts were performed in an improved Neubauer chamber.

Cryopreservation of procyclics was as follows: A small culture, typically 10 ml, was grown to midlog phase, approximately 6x 10^6 cells ml^{-1}. Cells were spun at 600*g* for 10 min at room temperature and the supernatant aspirated. The cell pellet was resuspended to 6x 10^7 cells ml^{-1} in complete SDM-79 with 5% DMSO. 500 µl aliquots were transferred to 1 ml cryotubes (Nunc) and the tubes were wrapped in cotton wool, placed in a polystyrene box and kept at -80°C overnight. The next day the tubes were transferred to liquid nitrogen long-term storage. On removal from liquid nitrogen tubes were allowed to thaw at room temperature before the contents were transferred to 4.5 ml complete SDM-79 medium with appropriate antibiotics, prewarmed to 27°C.

2.1.2: Transfection and cloning of procyclics

Transfection of procyclic form *T. b. brucei* STIB 247 was carried out as described by Sherman, 1991. Cells were grown to 6x 10^6 cells ml^{-1} and harvesting by centrifugation at 300*g* for 20 min at room temperature. Cell pellets were resuspended to 6x 10^7 cells ml^{-1} in Zimmermans post fusion medium (ZPFM) and spun as before. Supernatant was aspirated and the pellets again resuspended to 6x 10^7 cells ml^{-1} in ZPFM. 500 µl aliquots of resuspended cells were dispensed to 0.4 cm cuvettes (BioRad) and 5-50 µg plasmid DNA added and mixed in by gentle swirling. Each cuvette was then pulsed twice at 1.5 kV, 25 µF, in a Gene Pulser (BioRad) before being immediately removed to 4.5 ml complete SDM-79, prewarmed to 27°C. Transfection cultures were then incubated at 27°C for 6-12 hours to allow the cells to recover from electroporation before appropriate antibiotics were added and the cultures diluted to 10 ml, and cloned if so required.

Cloning of procyclic form cells was done by the semisolid plate method (Carruthers and Cross, 1992). Conditioned medium was made by growing wild-type STIB 247 procyclics to approximately 1x 10^7 cells ml^{-1} before centrifugation at 900*g* for 10 min to pellet the cells. The supernatant was decanted and centrifuged twice more as previously. The final supernatant was decanted and frozen at -70°C for several weeks, or kept at 4°C for 1-3 days, until required. A 10x suspension of low melting point type VII agarose (Sigma) was prepared and autoclaved, and then remelted by incubation in a 50°C waterbath. SDM-79 medium containing 20% FBS, 20% conditioned medium and 7.5 µg ml^{-1} haemin was prepared, filter sterilised, temperature

equilibrated in the same waterbath and added to the agarose. Final agarose concentration was 0.65%. After addition of appropriate antibiotics and thorough mixing 20 ml aliquots were poured into 9 cm Petri dishes. Plates were left to solidify for 1 hour and were dried in the TCR flow hood for exactly 1 hour before being spread with 100 µl of a suspension of transfected cells or control cells. Plates were incubated standing upright at 27°C until colonies started to appear (approximately 2 weeks). Colonies were lifted to 1 ml prewarmed complete SDM-79 medium, with appropriate antibiotics, using a sterile cell scraper.

2.2 Molecular materials and methods

2.2.1: *E. coli* strains

E. coli strain XL1-Blue (Stratagene) was used throughout for cloning, growth and general maintenance of plasmid DNA. For λFIXII work *E. coli* strain LE392 (Stratagene) was used. For the expression of recombinant histidine-tagged proteins *E. coli* strain M15[pREP4] (Qiagen) was used. For the expression of recombinant GST fusion protein the protease-deficient *E. coli* strain BL21 (Pharmacia Biotech) was used.

For short-term maintenance bacteria containing plasmid DNA were streaked to LB plates with appropriate antibiotics. Following overnight incubation at 37°C plates were kept at 4°C for up to one month, with individual colonies being picked with sterile picks to liquid LB medium as required.

For long-term maintenance individual colonies of bacteria containing plasmid DNA were inoculated to 3 ml LB with appropriate antibiotics and incubated at 37°C overnight with shaking at 200-250 rpm in a rotary shaker incubator. Approximately 16 hours later 1 ml of each culture was mixed with 1 ml of sterile LB containing 40% glycerol and 2% peptone, and this stabilate kept at -80°C. To safeguard plasmid maintenance, small-scale plasmid 'minipreps' were also made and kept at -20°C.

Bacteria were cultivated in LB broth with appropriate antibiotics in a rotary 37°C incubator; good aeration was ensured by not exceeding 20% of the flask volume with culture. Except in the case of cultures for plasmid minipreps, generally a small saturated overnight culture was used to inoculate a larger culture the next day, typically at a ratio of 1:100. This was the method used for midiprep cultures, and cultures for recombinant protein expression.

2.2.2: Plasmid isolation

Isolation of plasmid DNA from bacteria was effected using a variety of commercially available plasmid preparation kits. Small scale plasmid preparations - 'minipreps' - were done

using the WIZARD miniprep kit (Promega) or the Tip20 kit (Qiagen). Briefly, cell pellets were resuspended in TRIS buffer containing EDTA and RNaseA. Alkaline lysis of the bacteria was effected by addition of an NaOH solution containing SDS, followed by neutralisation with an acidic solution of potassium acetate to precipitate chromosomal DNA and cellular debris. The lysate was cleared by centrifugation and plasmid DNA specifically bound to proprietary resins. The bound DNA was washed in a medium salt-alcohol buffer that elutes contaminating RNA and traces of proteins, and the plasmid DNA was then eluted in a small volume of 10 mM TRIS pH 8.0 (WIZARD kit) or with a high salt-alcohol buffer (Tip20 kit). Plasmid DNA eluted from the Qiagen tips was then desalted and precipitated by isopropanol precipitation, washed with 70% ethanol, air-dried and resuspended in a small volume of 10 mM TRIS pH 8.0. Spectroscopic analysis was used in conjunction with electrophoresis of small aliquots of purified plasmid DNA on 1% TAE agarose gels to estimate DNA concentration.

Medium-scale plasmid preparations - 'midipreps' - were done using the Tip100 kit (Qiagen). The methodology is exactly the same as for minipreps, only scaled up.

2.2.3: Restriction digestion and ligation of plasmid

Restriction endonuclease digestion of plasmid DNA was typically carried out using 0.2-0.6 μg DNA in a total reaction volume of 20 μl at 37°C, except in specific cases where there was a requirement for incubation at a different temperature for maximal activity of the specific restriction endonuclease. Restriction enzymes and buffers used were obtained from either GibcoBRL or New England Biolabs. Where possible, combination digests using two enzymes were carried out simultaneously using a buffer compatible to both enzymes. In cases where this was not possible digestion was carried out using one enzyme and the buffer conditions then changed to suit the second enzyme, for example by raising the salt concentration. When buffer incompatibility precluded this, digestion with the first enzyme was followed by electrophoresing the cut DNA on an agarose gel and extraction of the DNA with the QIAquick Gel Extraction kit (Qiagen), then digestion with the second enzyme.

In restriction digests of large (100 μg) amounts of DNA intended for transfection, the extent of digestion was followed over time by electrophoresing small aliquots on TAE agarose gels. Once digestion was complete, the DNA was purified by two extractions with phenol:chloroform equilibrated by saturation with TRIS EDTA buffer (TE) pH 8.0 and one extraction with chloroform, followed by ethanol precipitation. The pellet was washed twice with large volumes of 70% ethanol. After brief drying under vacuum (approximately 5 min) DNA pellets were resuspended to 2 mg ml^{-1} in sterile distilled water preheated to 65°C, aliquoted and stored at -20°C until use.

Ligation reactions were performed using either the Amersham ligation kit or New England Biolabs T4 ligase and buffer. In both cases ligation reactions and conditions were according to the manufacturer's protocol, typically using a molar ratio of cut plasmid:insert fragment of 1:3, where possible. Reaction volume was typically 20 µl. The New England Biolabs system was far more frequently used, in which case 400 units of T4 ligase (for cohesive-end ligations) or 2 000 units of T4 ligase (for blunt-end ligations) were added to a 20 µl ligation reaction containing various concentrations of cut plasmid and insert fragment, and reaction buffer at 1x final concentration. Reactions were left to proceed overnight at 16°C in an insulated waterbucket.

2.2.4: Transformation of *E. coli*

Transformation of plasmid DNA or ligation reactions was done using either competent XL1-Blue cells prepared by the calcium chloride method (Sambrook *et al*, 1989), electrocompetent XLI-Blue cells, or commercially available competent XL1-Blue cells (Stratagene). Calcium chloride competent cells were prepared by growing a fresh culture of XL1-Blue *E. coli* to an optical density at 600 nm of 0.6 before harvesting by centrifugation at 6 000*g* for 10 min at 4°C. Cell pellets were resuspended in ice-cold 50 mM $CaCl_2$ to one-tenth the volume of the original culture and incubated on ice for 20 min. The cells were collected by centrifugation as previously, the resuspension and incubation step repeated, the cells collected again by centrifugation as previously and resuspended in 2 ml ice-cold 50 mM $CaCl_2$ per 50 ml original culture. An equivalent volume of LB with 40% glycerol and 2% peptone was thoroughly mixed in, 50 µl aliquots dispensed to eppendorfs and cells frozen quickly at -80°C. For use, aliquots were thawed on ice, transferred to sterile 15 ml Falcon tubes, and appropriate amounts of plasmid DNA or ligation reactions mixed in by gentle swirling, followed by a 25 min incubation on ice. Transformation reactions were then placed in a 42°C waterbath with a 40 second heat-shock window and a further 3 min incubation on ice before the addition of 0.5-3.0 ml of LB and incubation at 37°C in a rotary shaker for 1 hour to recover. Transformed cells were plated on predried LB plates at 100 µl per plate, with appropriate antibiotics added to the LB when pouring the plates and IPTG/X-gal where appropriate [2% X-gal with 40 mM IPTG] prespread on the plates once predried, and incubated, inverted, at 37°C overnight.

2.2.5: DNA gel electrophoresis

Agarose gel electrophoresis was carried out using agarose from GibcoBRL or SeaKem with TRIS acetate buffer (TAE), except in the case of gels to be used for Southern blotting, in which case TRIS borate buffer (TBE) was used. Stock solutions of 50X TAE and 50x TBE were prepared, autoclaved and dilutions of these used in casting agarose gels and preparing

electrophoresis buffers. Typical agarose concentration gels was 1%, although for resolution of smaller DNA fragments (500 bp and less) 2% agarose gels were used and for Southern blotting 0.8% TBE agarose gels were used. The molecular weight marker used was the 1 kb ladder (GibcoBRL) at a concentration of 0.5 μg per lane.

DNA in agarose gels was visualised by the addition of ethidium bromide to the gel mix and the electrophoresis buffer at a concentration of 0.5-2.0 μg ml^{-1}. Visualisation was by exposure to UV light using a transilluminator (Appligene) and images were recorded using a gel documentation system (Appligene).

DNA was extracted from agarose gels by using either SpinX minicolumns (Costar) or a Gel Extraction kit (Qiagen). In the former, the gel chip was added to the column, frozen briefly (30 min) at -70°C, thawed for 15 min at 37°C and then centrifuged in a bench microfuge to elute the DNA. Problems with this method include recovered DNA being relatively dilute and thus requiring subsequent precipitation, and agarose carry-through, which can inhibit subsequent enzymatic reactions. For these reasons the Qiagen kit was adopted for most purposes. DNA fragments were excised from the agarose gel with a scalpel, transferred to eppendorfs and weighed, then 3 volumes of dissolution buffer added and eppendorfs incubated at 50°C for 10 min to dissolve the agarose chip. DNA was bound to a resin and debris cleared by centrifugation, contaminating RNA and protein traces removed by washing with a salt-ethanol buffer and the DNA eluted by addition of a small volume of 10 mM TRIS pH 8.5. For downstream applications of ligation or probe labelling, the DNA was eluted with sterile distilled water preheated to 65°C, and kept at -20°C until use.

2.2.6: DNA sequencing

Manual DNA sequencing was carried out using the USB Sequenase 2.0 kit according to the manufacturer's instructions (Amersham). Reactions were run on wedge gels using glycerol tolerant buffer to prevent distortions. Typically, the gel was cast the night before use and then pre-run at 65 watts (approximately 2500 V) for 30 min prior to sample loading. Samples were heated to 80°C immediately before loading at 3 μl per lane, and once loaded were electrophoresed at the same voltage until the dye-front had come off the end of the gel. Once the run was complete gels were soaked in several changes of fixing solution [5% acetic acid, 15% methanol] to remove the urea, backed onto 3 mm Whatmann filter paper and dried in a slab dryer (BioRad) at 80°C for approximately an hour and a half. They were then exposed to autoradiography film (DuPont) by direct exposure in a hypercassette (Amersham) for twelve hours to seven days, depending on signal strength. This method of sequencing was used initially, until automatic cycle sequencing became available. However, for obtaining sequence very close (10-50 bp) to the primer site, this was still the method of choice, as cycle sequencing

does not provide good resolution in this range. The use of a manganese modification to the Sequenase kit allows good results in this range.

Automatic DNA cycle sequencing was carried out using the Perkin-Elmer Amplitaq FS kit according to the manufacturer's instructions. 400 ng of template DNA and 3.2 pmol of primer were used per reaction, and reactions were performed in an Appligene GeneAmp 2400 machine, cycles programmed according to the manufacturer's instructions. DNA for this purpose was always prepared using the Qiagen Tip20 kit, as this was generally agreed by all operators to give the optimal results, with the DNA eluted from the resin with distilled water.

Data analysis was performed using 373 Sequence Analysis Software on an Apple Mac computer. Once the sequence trace had been examined and ambiguities removed manually the appropriate section of good data was selected, and exported to UNIX, to be manipulated by GCG packages.

2.2.7: Isolation of genomic DNA and Southern blotting

Isolation of genomic DNA from STIB 247 procyclics was performed using the Nucleon II kit (Scotlab) which utilises a silica suspension for purification of DNA. 1x 10^7 procyclic cells were lysed by resuspension in 2 ml Nucleon buffer B and treated with 0.75 μg ml^{-1} RNaseA at 37°C for 30 min, and then deproteinised by treatment with 1 M sodium perchlorate. Following extraction with chloroform precooled to -20°C the silica suspension was added and the mixture centrifuged at 1400*g* for 3 min at 4°C. The aqueous, DNA-containing upper phase was then removed and any residual silica suspension collected by centrifugation at 1300*g* for 1 minute. The supernatant was removed and DNA was precipitated by addition of 2x volumes of 100% ethanol at -20°C. DNA was removed by hooking it with a heat-sealed pasteur pipette, rinsed in a large excess of 70% ethanol and briefly air-dried. The DNA pellet was then redissolved in 50 μl of 10 mM TRIS pH 8.0 to give an approximate final concentration of 1 mg ml^{-1}, and stored at 4°C.

Southern blotting was performed by the capillary method as described (Sambrook *et al*, 1989). Briefly, digested samples were electrophoresed on 0.8% TBE agarose gels, either for several hours at 70-90 V or overnight at 25 V, in both cases until the dye-front had migrated approximately three-quarters of the way down the gel. Gels were stained by incubation in 1x TBE buffer with 1 μg ml^{-1} ethidium bromide at room temperature for 25 min, DNA visualised on a gel documentation system (Appligene) and a visual image taken, and gels rinsed in a large excess of distilled water for 10 min. Gels were then incubated in several changes of denaturing buffer for 45 min with gentle agitation and then several changes of neutralisation buffer for 45 min with gentle agitation. After neutralisation, DNA was transferred to Hybond-N supported nitrocellulose (Amersham) by capillary action using 10x SSC as the transfer buffer. Following

overnight transfer membranes were air-dried and fixed by exposure to UV light. Filters were pre-hybridised at 65°C in prehybridisation buffer for 4 hours. The prehybridisation solution was then decanted and the radioactive probe added in 5 ml of fresh prehybridisation solution and incubated at 65°C overnight. The probe was decanted the next morning and the membrane washed at 65°C with shaking for 2x 10 min in 2x SSC with 0.1% SDS, then for 3x 15 min in 0.1x SSC with 0.1% SDS. The membrane was then exposed to Renaissance autoradiography film (DuPont) in a hypercassette (Amersham).

Probes for Southern blotting were prepared by gel purifying the particular DNA fragment from TAE agarose gels using the Qiagen Gel Purification kit and eluting the DNA in distilled water. 25 ng of purified DNA was labelled to high specific activity using the Prime-It II kit (Stratagene) with 3000 Ci $mmol^{-1}$ ^{32P}-dCTP, according to the kit protocol. Briefly, purified DNA was mixed with random oligonucleotide primers, heated to 100°C for 5 min and allowed to cool to room temperature to promote annealing of the random primers to the DNA fragment. Reaction buffer, radiolabelled dCTP and 5 units of Exo(-) Klenow enzyme were added, the reaction incubated at 37°C for 10 min to allow extension, and the reaction stopped. Labelled probe was then purified from unincorporated nucleotide on a NucTrap column (Stratagene) using a beta shield device (Stratagene) according to the manufacturer's instructions. Purified probes were stored in eppendorfs, and sealed in lead pots at -20°C until use.

2.2.8: Isolation and manipulation of RNA

Standard precautions were taken when isolating and manipulating RNA: All tips were specially autoclaved for this purpose and used only for RNA work and all solutions, where possible, were treated with 0.1% DEPC.

Isolation of total RNA from procyclics was performed using the Total RNA Midi kit (Qiagen), according to kit instructions with one exception: Trypanosomes are smaller by volume than mammalian and other higher eukaryotic cells, and thus proportionally more cells per prep must be used to obtain the same quantity of RNA. Accordingly, ten times as many cells were used per midiprep as recommended by the manufacturers for mammalian cells. Briefly, cells were lysed and genomic DNA sheared in a highly denaturing buffer that generated an RNase-free environment and stabilised RNA; many proteins were then removed by salt precipitation. RNA was collected by rapid precipitation, resuspended, and RNA bound to a resin that only binds nucleic acids. The resin was then washed with a medium salt buffer that specifically elutes DNA but leaves RNA bound. RNA was eluted with a high salt wash, collected by isopropanol precipitation, washed with large volumes of 70% ethanol, and resuspended in a small volume of RNase-free water. RNA was stored at -80°C until use.

Isolation of poly[A]+ RNA was done using the Oligotex Direct mRNA kit (Qiagen); again, ten times as many cells per prep was used as recommended in the kit manual. The kit exploits the fact that eukaryotic mRNA species end in a homopolymer of 20-250 adenosine nucleotides. Cells were lysed and homogenised using a QIAshredder, diluted with Dilution buffer and centrifuged to remove cellular debris and protein. The supernatant was removed to a fresh eppendorf and Oligotex suspension added, allowing the poly[A] transcript tails to hybridise to oligo[dT] primers which were coupled to latex microparticles. Following washing to remove nonbound material poly[A]+ RNA was eluted in 5 mM RNase-free TRIS pH 7.5 preheated to 70°C. Poly[A]+ RNA was stored at -80°C until use.

First-strand cDNA synthesis was carried out as follows: Eppendorfs were labelled either plus (+) or minus (-). Each received 5 µg of total RNA or 1 µg of poly[A]+ RNA, 60 pmol oligo[dT] or oligo[dT]-Anchor primer, and DEPC-treated distilled water to 12.5 µl. Tubes were incubated at 65°C for 10 min to denature the RNA, then chilled on ice to allow annealing. Following a brief spin, each tube received First Strand Buffer (GibcoBRL) to 1x final concentration, 50 mM DTT, 1 mM dNTPs and 0.5 µl of RNasin (GibcoBRL). Incubation at 42°C for 2 min was followed by addition of 25 units of Molony Murine Leukaemia Virus (M-MLV) reverse transcriptase (GibcoBRL) to the (+) tube and 1 µl of DEPC-treated distilled water to the (-) tube. Incubation at 42°C for 30 min allowed first-strand synthesis to occur, and the reactions were stopped by the addition of 480 µl TE pH 8.0 to each tube and heating at 65°C for 10 min. First-strand cDNA was stored at 4°C.

RNA for Northern blots were run on 1x MOPS/0.7 M formaldehyde gels according to the protocol supplied with the instructions for Hybond-N membrane (Amersham). Agarose gel strength was 1.5%, and no ethidium bromide was used as it can inhibit RNA transfer to the membrane. Following electrophoresis, the gel was rinsed for 15 min in RNase-free water and then in 10x SSC. RNA was transferred to Hybond-N membrane by capillary blotting as for Southern blots, using 10x SSC as the transfer buffer. Once transferred, membranes were partially dried, UV fixed and stained with methylene blue to visualise the RNA. Following destaining the membranes were then treated as for Southern blots. The RNA molecular weight markers used was an RNA ladder (New England Biolabs).

2.2.9: PCR reactions

PCR reactions were carried out using two different enzymes: Taq (*Thermus aquaticus* DNA polymerase, Promega and Applied Biosystems) and the high-fidelity Pfu (*Pyrococcus furiosus* DNA polymerase, Stratagene), the latter only being used in a few cases where fidelity was of greater importance than the enzyme's processivity or the abundance of the end-product.

PCR using plasmid DNA was performed in 10 µl reactions with 1 µl of an appropriate dilution of plasmid (typically, approximately 200-300 pg DNA per reaction) in thin-walled 200 µl PCR tubes. The other components of the reactions were 50 pg of each primer, 0.4 units of Taq DNA polymerase and 0.9 µl of 11.1x PCR mix. Reactions were carried out in a Perkin-Elmer GeneAmp 2400 machine. Reactions involving amplification of very GC-rich regions of DNA were facilitated by addition of DMSO to a final concentration of 5% in the reaction mix; DMSO inhibits Taq to approximately 50%, therefore double the usual amount of Taq per reaction was used in these cases. Extension times were based on an assumption of 30 seconds per kilobase, and annealing temperatures based on the thermal Tm values of the primers involved. For reactions using the high-fidelity but slow-procession polymerase Pfu extension times were doubled. PCR reaction end-point products were electrophoresed on 1% TAE agarose gels.

Screening by PCR of bacterial colonies, principally used to screen bacteria transformed with ligated plasmid, was performed as follows: Single colonies were picked off plates using a sterile pick and scraped off into 200 µl thin-walled PCR tubes. The picks were then transferred to 3 ml LB with appropriate antibiotics and set to grow in a rotary 37°C incubator. Meanwhile, PCR reactions were set up as normal in the tubes containing the cell scrapes but with no added DNA. A prehold of 4 min at 94°C was sufficient to lyse the cells, with the annealing temperature based on the thermal Tm of the primers and an extension time based on the length of the target region to be amplified. Once complete, reactions were electrophoresed on TAE agarose gels and based on the results of this, LB cultures were either discarded or left to grow up overnight for plasmid DNA preps.

PCR reactions with genomic DNA were performed as described for plasmid PCR reactions; PCR reactions using first-strand cDNA were performed as described for plasmid PCR reactions, except that double the volume of template (2 µl) was used per 10 µl reaction.

2.2.10: Screening λFIXII library

A λFIXII genomic library was used to screen for the *CRK4* gene, to isolate a clone or clones representing the full open reading frame. The bacterial host used for the phage was LE392, which was prepared by picking a single colony from an LB plate and inoculating 20 ml LB medium containing 10 mM maltose and 10 mM $MgSO_4$. This was grown at 37°C with shaking (225 rpm) until the optical density of the culture at 600 nm was between 0.5-1.0. Cells were harvested at 2000*g* for 20 min at 4°C, then resuspended in sterile 10 mM $MgSO_4$ to an optical density at 600 nm of 0.5. Cells prepared this way were stored at 4°C and used in under 3 days. For λFIXII plating on petri dishes, 100 µl of prepared LE392 were mixed with 100 µl of

sterile phage buffer containing λFIXII and incubated for 20 min at 37°C to allow adsorption of phage to the cells. After this each cell-phage mix was inoculated to 3.5 ml BBL top agarose at 48°C, mixed by gentle inversion, and poured onto predried BBL bottom agarose plates. Plates were incubated at 37°C, inverted, overnight or until plaques appeared and lysis was complete.

Phage lifts for Southern blotting from plates or petri dishes was performed by laying a clean sheet of Hybond-N on the surface of the top agarose for 1-2 min, using needles to mark orientation on the membrane and plate. The membrane was then carefully lifted off, submersed DNA side up for 5 min in denaturing solution, DNA side up for 5 min in neutralising solution, and then rinsed briefly in 2x SSC before being air-dried and UV fixed. The rest of the procedure was as for a standard Southern blot. Agar plugs were removed to 1 ml phage buffer, incubated at 37°C for 2-4 hours, one drop of chloroform added and the phage solution stored at 4°C.

λFIXII clones were grown for the isolation of phage DNA by the liquid culture method as described in the Promega kit manual. Briefly, 500 µl of an overnight LE392 culture was mixed with 10-20 µl of agar plug eluate and incubated at 37°C for 20 min. This was then inoculated to 100 ml prewarmed LB with 10 mM $MgSO_4$ and grown at 37°C in a rotary incubator for 6 hours. Cell lysis was not clear at this point, so 500 µl chloroform was added and the culture left to shake for a further 15 min. The lysate was then centrifuged at 8 000*g* for 10 min at 4°C to pellet cellular debris. The supernatant was removed to a fresh tube and phage DNA purified using the Magic Lambda Preps DNA Purification System (Promega). Phage DNA was eluted in a small volume of 10 mM TRIS pH 8.0.

2.3 Biochemical materials and methods

2.3.1: Protein gels and Western blotting

SDS-PAGE gels were cast and run according to the method of Lammeli (Lammeli, 1970) as adapted for the BioRad Mini-Protean system. Typically, resolving gel acrylamide concentration was 12.5 %, but for better resolution of some closely migrating proteins a resolving gel acrylamide concentration 7.5 % strength was used. SDS-PAGE running buffer (SPRB) was made as a 10x stock and used at 1x concentration. Usually, gels were run at 130 V at room temperature until the dye-front had just migrated off the bottom of the resolving gel. Loading buffer for protein samples was either four times sample buffer (FSB) or three times sample buffer (TSB, New England Biolabs). Gels were stained with Coomassie blue stain and destained with either fast destain or slow destain. Gels were dried in a slab gel drier (BioRad)

or an EasyBreeze apparatus (Hoeffer). Silver staining was carried out using commercially available kits from either Sigma or BioRad, according to the manufacturer's instructions. Molecular weight markers used were Kaleidoscope coloured markers (Amersham), low-range Rainbow coloured markers (BioRad) and both prestained and non-prestained Low Range markers (New England Biolabs).

Western blotting of gels was carried out under a variety of conditions, as requirements for good results vary between different antisera. The typical protocol used was as follows: Gel blotting was done in a wet-blot tank (BioRad) at 250 mA for 2 hours at 4°C, or at 25 mA overnight at 4°C, in Towbin transfer buffer (Towbin *et al.*, 1979). Transfer was to either Hybond-N supported nitrocellulose membrane (Amersham) or to PVDF membrane (DuPont). Because PVDF has a higher specific capacity for protein binding and generally resulted in lower backgrounds, this was usually the membrane of choice. Following transfer, membranes were briefly stained with Ponceau S to check for the transfer of proteins, destained with 5% acetic acid, and lanes demarcated with a soft lead pencil; non-prestained markers were also marked at this point with a soft lead pencil. After staining and destaining, PVDF membranes were rewetted by briefly submerging in 100% methanol and then rinsing in TRIS-buffered saline with Tween-20 (TBST) buffer. Blocking of nonspecific binding sites for immunoglobulins on the filters was performed by incubation of the membranes for 1 hour at 4°C in blotto with 0.05 % sodium azide. Binding of primary antiserum to target proteins was carried out using varying titres of monoclonal antibody or polyclonal antiserum in blotto with 0.05 % sodium azide at 4°C for 2-4 hours. Following this, membranes were washed for 6x 5 min in large (200-300 ml) volumes of 1x TBST on a rotary shaking platform at room temperature, and then incubated with secondary HRP-conjugated antibody in blotto, either anti-rabbit antiserum or anti-mouse antiserum (both from Promega). Following incubation of membranes with secondary antibody, they were washed as previously in 1x TBST. Blots were developed by incubation of the membrane in ECL reagents (Amersham) or SuperSignal Substrate system (Pierce). As the Pierce system is very powerful optimal exposure was often achieved by diluting the active solution mix 1:10 in distilled water before incubation of membrane

2.3.2: Preparation of protein samples

Small-scale (mini) crude whole cell *E. coli* lysates were prepared by resuspending cell pellets in an appropriate volume of lysis solution (LS) and adding either one-third or one-quarter volume of TSB or FSB, respectively. The suspension was then needled 5-6 times through progressively smaller needles: 21G, 23G and 25G, the suspension transferred to a fresh eppendorf and boiled at 100°C for five min.

Soluble (S-14) and insoluble protein fractions were made by spinning whole cell lysates at maximum speed in a bench microfuge for 10-20 min at 4°C. The soluble protein fraction was aspirated and transferred to a fresh eppendorf, insoluble protein was resuspended in fresh LS. For larger-scale soluble protein lysate preparation cell pellets were resuspended in a phosphate- or TRIS-based saline buffer and sonicated using a microtip. Lysates were centrifuged at 10 000*g* for 20 min at 4°C to pellet cellular debris, the supernatant aspirated and processed as appropriate.

Crude whole cell lysates of *T. brucei* were prepared by taking harvested, washed cells (see chapter 2.1.1), resuspending the cell pellet to approximately 5x 10^8 cells ml^{-1} in LS plus protease inhibitors (LSI) and incubating on ice for 20 min. An appropriate volume of sample buffer was then added, either TSB or FSB, and the lysate was processed as for *E. coli* whole cell lysates. S-14 protein fractions were made as described for E. coli lysates. S-100 lysates were made by performing the lysis step and then centrifuging at 100 000*g* for 45 min at 4°C in a Beckman ultracentrifuge. The supernatant was aspirated and was used for immunoprecipitations, selections and kinase assays. Estimation of protein concentration was performed using the Biorad DC Protein Microassay kit, with a set of immunoglobulin standards as reference.

2.3.3: Expression of fusion proteins in *E. coli*

The systems used for the production of recombinant fusion protein in *E. coli* were the QiaExpress system (Qiagen) and the GST Gene Fusion system (Pharmacia Biotech).

The pQE32 vector was chosen for expression of the *CYC1* gene; this vector adds an N-terminal hexahistidine tag, and includes an ATG start as well as all-frame stop codons. The pQE60 vector was used in the generation of histidine-tagged *CRK2*, used to create the pHD675-CRK2His construct (pGL234), as the vector's polylinker allowed for easy tag insertion and subsequent subcloning, and for the expression of full-length CRK2.

The pQE32-CYC1 construct (pGL107) was transformed to M15[pREP4] *E. coli*, and a series of pilot expression experiments carried out. A small-scale native purification was performed as per the manufacturer's instructions: 2x 1.5 ml LB was inoculated with 500 µl of an overnight saturated culture of M15[pREP4]pGL107. The cultures were grown for 30 min at 37°C before addition of IPTG to a final concentration of 2 mM to one of the cultures, and grown for a further 2 hours before harvesting. The cell pellets were resuspended in SB, incubated with 1 mg ml^{-1} lysozyme for 30 min, put through 5 freeze/thaw cycles in dry ice/ethanol and a 100°C heating block. The supernatant from an S-14 spin was added to 50 µl of a 50% slurry of Ni-NTA resin (prewashed in SB) and incubated for 30 min at 4°C with gentle

agitation. Brief centrifugation collected the resin, the supernatant was aspirated and bound protein eluted by the addition of SB containing 200 mM imidazole.

A subcellular localisation check was performed for CYC1His by inoculating 100 ml LB at 1:100 with an overnight saturated culture of M15[pREP4]pGL107 and growing to an OD at 600 nm of 0.8. After harvesting, cell pellets were processed by the protocol described by the manufacturer (Qiagen) to give soluble, insoluble and periplasmic shock protein fractions.

Pilot denaturing purification of CYC1His was performed by inoculation of 2 L of LB with 10 ml of a saturated overnight culture of M15[pREP4]pGL107. When cell density had reached an OD at 600 nm of 0.8, IPTG was added to 1 mM and the culture grown for a further 2 hours before harvesting as two equivalent aliquots. One cell pellet was resuspended in 5 ml QIAexpress buffer A, the other resuspended in 5 ml QIAexpress buffer B. The resuspended cells were sonicated and the supernatants from a centrifugation at 4 000*g* for 10 min at 4°C were loaded to 2 ml bed volume columns of Ni-NTA resin that had been pre-equilibrated in buffer A or buffer B, as appropriate. The columns were washed and CYC1His eluted as shown:

Column 1. Washed with buffer A, buffer C and eluted by the protonation method (buffers D, E and F).

Column 2. Washed with buffer B, buffer C and eluted by the protonation method (buffers D, E and F).

Column 3. Washed with buffer A, buffer C and eluted with imidazole (buffer C with 300 mM imidazole).

Column 4. Washed with buffer B, buffer C and eluted with imidazole (buffer C with 300 mM imidazole).

From the pilot denaturing purification above, a denaturing protocol with elution of CYC1His by protonation was chosen to produce milligram quantities of recombinant protein. M15[pREP4]pGL107 were grown, induced with 2 mM IPTG for 3 hours and harvested as described.. The cell pellet was resuspended in 5 ml of QIAexpress buffer A, sonicated, centrifuged at 2 000*g* for 10 min at 4°C, and the supernatant applied to a 4 ml bed volume Ni-NTA column previously equilibrated in buffer A. The resin was washed with 15 column volumes of buffer B, then with 15 column volumes of buffer C. The elution program was 5x 3 ml buffer D, 5x 3 ml buffer E and 6x 3 ml buffer F. Samples of each eluate fraction were analysed on Lammeli gels and by spectroscopic measurement at 280 nm, to ensure that the bulk of the CYC1His had eluted in the first two F eluate fractions. The protein-containing fractions were pooled and dialysed against 2.5 L of dialysis buffer over 2 days with three changes of buffer. Following dialysis, precipitated CYC1His was collected by centrifugation at 4 000*g* for 10 min at 4°C, the supernatant aspirated and the protein pellet resuspended in 1 ml of dialysis

buffer, sonicated briefly to break up precipitated protein aggregates with a microtip. Purified CYC1His was then dispensed as 50 µl aliquots to eppendorfs and stored at -20°C until required.

The pQE60-CRK2 construct (pGL217) was transformed to M15[pREP4] *E. coli* and a pilot induction of CRK2His was performed as described by the manufacturer (Qiagen). Pilot time-course induction and pilot denaturing purification was carried out as described by the manufacturer (Qiagen).

The pGEX-5X-2 vector was chosen for expression of a fragment of the *CRK4* gene, as this vector preserved the reading frame of the GST-CRK4 fusion. pGL134 was transformed to competent BL21 cells and a small-scale pilot time-course expression was carried out by inoculating LB at 1:100 using a saturated overnight culture of BL21pGL134 and growing to an OD of 0.8 at 600 nm before induction with 2 mM IPTG. Cultures were then grown for 4 hours with samples withdrawn hourly, and crude whole cell lysates made as described in chapter 2.3.2. A solubility check was carried out on an induced cell pellet as described in chapter 3.2.6.

Small-scale pilot induction and purification was performed by induction of BL21pGL134 with 2 mM IPTG. Cells were harvested, resuspended in SB and sonicated, and centrifuged. The S-14 supernatant was added to a 50% slurry of Glutathione-Sepharose 4B (GS4B) resin and incubated for 10 min at room temperature with gentle agitation. Meanwhile the cellular debris (insoluble protein) was resuspended in QIAexpress buffer B/TSB, syringed to resuspend and heated to 100°C for 5 min. The resin was collected by brief centrifugation, the supernatant aspirated, added to TSB and heated to 100°C for 5 min (soluble nonbound protein). The resin was washed with 1x PBS pH 7.2 and bound protein eluted by the addition of Glutathione Elution Buffer (GEB) and incubation at room temperature for 5 min. The resin was collected by brief centrifugation, the elute aspirated and added to TSB, and heated to 100°C for 5 min (soluble bound protein).

Induction and purification of milligram quantities of GST-CRK4 was done by the column method as described in the manufacturer's kit instructions. 4x 500 ml LB were each inoculated at 1:100 with a saturated overnight culture of BL21pGL134, grown to an OD at 600 nm of 0.8, induced by the addition of 0.5 mM final concentration IPTG and grown for 4 hours. Cells were harvested by centrifugation at 7 500*g* for 10 min at 4°C. Cell pellets were resuspended in a total of 100 ml 1x PBS pH 7.2 and sonicated, TritonX-100 was added to 1% and the lysate was centrifuged at 12 000*g* for 10 min at 4°C. The supernatant was applied to a 2 ml bed volume column of GS4B. The resin was washed with 1x PBS pH 7.2 until the protein content of the flow-through was negligible, and bound GST-CRK4 was eluted with 3x 3 ml GEB, each time incubating for 10 min at room temperature with gentle agitation. Eluate fractions were dialysed overnight at 4°C against 2 L 1x PBS pH 7.2. Purified GST-CRK4 was dispensed as small aliquots to eppendorfs and stored at -20°C until required.

2.3.4: Generation of antisera and purification of antibodies

The general schedule of immunisation of New Zealand white rabbits with purified recombinant protein was as described in Harlow and Lane (1988). Initially, a preimmune test bleed of several millilitres was taken before a subcutaneous injection with 200 μg of protein homogenised in an equivalent volume of complete Freund's reagent. Between four and six week after primary immunisation, secondary immunisation was performed using 200 μg of protein in incomplete Freund's reagent, in the same range of volume. An immune test bleed was taken approximately one week after this secondary immunisation. Two weeks after secondary immunisation a programme of booster immunisations began, each administered two weeks apart and with a test bleed being taken prior to each boost. For booster immunisation only 20 μg protein in incomplete Freund's reagent was required. The animal was exsanguinated and the antiserum stored at -20°C until use directly in experiments, or until antibody purification.

Once a specific response was detected with a particular antiserum on Western blots, the antiserum was partially purified by selecting specifically for the IgG fraction, using affinity purification with the immobilised recombinant bacterial protein A/G kit (Pierce) as described by the manufacturers. Sodium azide added to 0.03% and gelatin to a final concentration of 0.1%, to stabilise the antibodies, and the purified antibodies stored at 4°C.

Antisera raised using GST fusion proteins will contain a strong anti-GST response, and this can cause problems in immunoprecipitations and Western blots due to cross-reactivity to endogenous GST-like proteins in the protein samples under study. Consequently, anti-GST antibodies were removed from the anti-CRK4 antiserum used in such applications by using prepacked immobilised GST columns (Pierce) as described by the manufacturers. An aminolink immobilised GST-CRK4 column was made to further increase the specificity of the anti-CRK4 antibodies by specifically purifying only those antibodies that bind the CRK4 portion of the fusion protein. An immobilised column was made as per the manufacturer's instructions (Pierce). Freshly purified GST-CRK4 was made and the concentration was calculated using the Biorad DC protein microassay system; subsequently 10 mg of GST-CRK4 was used to couple a 2 ml bed volume column of aminolink gel. The GST-CRK4 sample was diluted 1:3 in 0.1 M sodium phosphate pH 7.0. After equilibration of the column with 0.1 M sodium phosphate pH 7.0 the GST-CRK4 sample was added to the column with 200 μl of reductant solution, and the reaction incubated at room temperature for 6 hours with gentle agitation. Subsequently the column was drained and washed with 5 ml 0.1 M sodium phosphate pH 7.0, the flow-through being collected and dialysed overnight against an excess of 1x PBS pH 7.2 for protein concentration determination. Any remaining active sites on the gel were blocked by washing the column with 4 ml of 1.0 M TRIS pH 7.4 and then incubating the gel with 2 ml 1.0 M TRIS pH 7.4 and 40 μl of reductant solution for 30 min at room temperature with gentle agitation.

Finally, the column was drained, washed with 1.0 M NaCl and then 1x PBS pH 7.2 with 0.05% sodium azide, and stored under 1x PBS pH 7.2 with 0.05% sodium azide at 4°C. A BioRad DC protein assay of the dialysed column flow-through and wash fraction showed a coupling efficiency of approximately 98%.

Specific purification of anti-CRK4 antibodies was carried out using aliquots of antisera which had been processed to remove all anti-GST antibodies, as described above. Following pre-equilibration of the GST-CRK4 aminolink column with 1x PBS pH 7.2 the partially purified antiserum diluted 1:3 with 1x PBS pH 7.2, applied to the column and the column incubated for 90 min at room temperature with gentle agitation. The column was then drained and washed with 5x 5 ml 1x PBS pH 7.2. Bound antibodies were eluted by application of 8x 1 ml of Immunopure IgG Elution buffer (Pierce) and then 8x 1 ml of 100 mM glycine pH 2.8. Eluate was collected as 1 ml fractions in eppendorfs (containing 50 µl 1 M TRIS pH 9.5 for the glycine elutant fractions, to neutralise them immediately) and spectroscopic measurement at 280 nm used to identify the fractions containing the majority of the eluted anti-CRK4 antibodies. These fractions were pooled and dialysed overnight at 4°C against a large excess of 30 mM sodium borate pH 7.2. After dialysis, 0.1% gelatin and 0.05% sodium azide was added to stabilise the antibodies.

Polyclonal antisera raised against peptide-carrier protein antigens were affinity purified, after protein A/G purification, on immobilised peptide columns. The peptides used had been generated with C-terminal cysteine residues for sulfhydryl coupling to carrier proteins, therefore they were coupled to a solid support matrix for antibody purification using the Sulfolink Immobilisation kit (Pierce) as described by the manufacturer. Affinity purification was carried out by loading an aliquot of the antiserum, diluted 1:1 with Immunopure Gentle Ag/Ab Binding buffer on a peptide column pre-equilibrated with the same buffer and incubating at room temperature for 60 min on a rolling platform. The column was drained and washed with 20 bed volumes of the same buffer, and antibody was eluted with 8 ml of Elution buffer followed by 8 ml of 100 mM glycine pH 2.8. 0.5 ml fractions were collected in eppendorfs. In the case of the glycine wash fractions the eppendorfs contained 50 µl 500 mM TRIS pH 8.5, to immediately neutralise the eluate fractions. Spectroscopic analysis was carried out to determine the fraction containing the bulk of the eluted antibodies; these were pooled and dialysed at 4°C overnight against a large excess of 30 mM sodium borate buffer. Sodium azide added to 0.05% and the antibodies kept at 4°C.

2.3.5: Immunoprecipitation-linked kinase assays and Western blots

Immunoprecipitation was carried out using the same basic protocol: An S-100 soluble protein lysate was incubated of aliquots of the lysate for 2-4 hours at 4°C with either polyclonal

antiserum or monoclonal antibodies. For controls antiserum/antibodies that had been preblocked by preincubation with a saturating amount of peptide or purified recombinant protein overnight at 4°C were also performed. Minicolumns containing small aliquots of Sepharose-protein A beads (50-80 µl of a 50% slurry) were prepared by washing the bead slurry with LS and then preblocking the beads with 250 µl of 10 mg ml^{-1} BSA in LSI for 2 hours, to block nonspecific binding sites. The antibody-lysate mixes were then added to the columns and left to bind for 1-3 hours at 4°C with agitation. Subsequently, columns were drained and washed extensively with LS, high salt lysis solution (HSLS) and then either kinase assay buffer (KAB) or 50 mM MOPS pH 7.2 buffer with 200 mM NaCl, depending on whether the beads were to be used in kinase assays or Western blots.

Immunoprecipitation-linked kinase assays were performed by resuspending Sepharose-protein A beads, incubated with antibody-lysate samples, in KAB and transferring to eppendorfs. After a brief spin in a bench microfuge to pellet the beads the KAB was aspirated and 20 µl of kinase assay mix (KAM)was added, making sure the beads were well resuspended. After incubation at 30°C for 20 min the reactions were stopped by addition of 20 µl of TSB or FSB and heating to 100°C for 5 min. The beads were centrifuged briefly before 20 µl aliquots of samples were electrophoresed on 12.5% Lammeli gels. Gels were stained with Coomassie blue, dried in a slab dryer (BioRad) and exposed with phosphorescent markers to autoradiography film (DuPont) in hypercassettes (Amersham) for 12-72 hours.

Immunoprecipitation-linked Western blots were performed by preparing beads and purifying antibody-lysate mixes as described above, and resuspending the beads in 50 mM MOPS pH 7.2 with 200 mM NaCl. After aliquoting to eppendorfs and brief centrifugation to collect the beads and aspirating the supernatant, beads were resuspended in an appropriate volume of a 1:1 mix of 50 mM MOPS pH 7.2/200 mM NaCl and TSB or FSB, and heated to 100°C for 5 min before samples were electrophoresed on 12.5% Lammeli gels. This method was only useful when the primary antibody or antiserum in the Western blot was from a different species of animal from that used in the immunoprecipitation.

2.3.6: Nickel-agarose selection of protein

Selection of CRK proteins from trypanosome S-100 lysates with $p12^{CKS1}$ and $p13^{SUC1}$ aminolinked beads was performed as described by Mottram and Grant (1996). An S-100 lysate was prepared. Minicolumns containing 40 µl of a 50% slurry of $p12^{CKS1}$, $p13^{SUC1}$ and control beads were prepared by washing the bead aliquots with 5 ml LS, adding 250 µl of 10 mg ml^{-1} BSA in LSI and incubating for 120 min at 4°C with agitation. Subsequently, 100 µl of S-100 lysate per column was added and the columns incubated for 120 min at 4°C with agitation. The

columns were drained, washed with 5 ml lysis solution with glycerol (LSG), 5 ml HSLS and 5 ml LS. Bead aliquots were resuspended in 1 ml 50 mM MOPS pH 7.2, transferred to eppendorfs, collected by brief centrifugation, resuspended in 66 µl 50 mM MOPS pH 7.2/33 µl TSB and heated to 100°C for 5 min. 15 µl aliquots were electrophoresed on 12.5% Lammeli gels for Western blotting.

Ni-NTA selection of CRK2His was performed by growing 10 ml cultures of the nonclonal procyclic population WUMP 1163 to midlog in the presence of tetracycline hydrochloride. S-100 lysates were prepared. 80 µl aliquots of Ni-NTA resin were dispensed to minicolumns and washed with 5 ml LSG with inhibitors (LSG). 200 µl of S-100 lysate was added per column, and the columns incubated at 4°C for 60 min with agitation. Columns were drained, washed with 5 ml LSG, 5 ml HSLS and then either 5 ml LS without 1% TritonX-100 (LS-T), if the intended use was for stained Lammeli gels or Western blotting, or with 5 ml KAB, if the intended use was for kinase assays. For the former, bead aliquots were resuspended in 1 ml LS-T, transferred to eppendorfs, collected by brief centrifugation, resuspended in 100 ml of a 2:1 mix of LS-T and TSB, and heated to 100°C for 5 min. For kinase assays, bead aliquots were resuspended in 1 ml KAB, transferred to eppendorfs and collected by brief centrifugation. Bead aliquots were resuspended in 20 µl KAM, incubated at 30°C for 20 min and the reaction stopped by addition of 20 µl TSB. 20 µl aliquots of reactions were electrophoresed on 12.5% Lammeli gels, the gels were dried and exposed with phosphorescent markers to autoradiography film. In most cases, regardless of the final application, samples of S-100 (induced or not with tetracycline) lysates and columns flow-through (nonbound soluble protein) were taken to confirm induction of CRK2His and binding to the resin.

2.4 Buffers and solutions

PBS:
100 mM sodium phosphate pH 7.2
200 mM NaCl

ZPFM:
132 mM NaCl
8 mM KCl
8 mM Na_2HPO_4.anhydrous
1.5 mM KH_2PO_4.anhydrous
1.5 mM magnesium acetate
0.9 mM calcium acetate

50x TAE:	2 M TRIS
	2 M sodium acetate trihydrate
	50 mM EDTA pH 8.0
	Adjusted to pH 7.2 with HCl
50x TBE:	2 M TRIS pH 7.2
	2 M boric acid
	50 mM EDTA pH 8.0
	Adjusted to pH 7.2 with HCl
PSG:	60 mM Na_2HPO_4.anhydrous
	4 mM $NaH_2PO_4.2H_2O$
	45 mM NaCl
	55 mM D[+] glucose
	Adjusted to pH 8.05 with NaOH
TE:	10 mM TRIS
	1 mM EDTA pH 8.0
	Adjusted to pH 8.0 with HCl
Nucleon buffer B:	400 mM TRIS
	60 mM EDTA pH 8.0
	150 mM NaCl
	1% SDS
	Adjusted to pH 8.0 with HCl
Denaturing Buffer:	0.5 M NaOH
	1.5 M NaCl
Neutralising buffer:	1.0 M TRIS pH 7.5
	1.5 M NaCl
10x SSC:	1.5 M NaCl
	1.5 M trisodium citrate pH 7.0

Prehybridisation buffer:

6x SSC
10x Denhardt's solution
0.1% SDS
0.1% sodium pyrophosphate
100 μg ml^{-1} salmon sperm DNA

Denhardt's solution:

5% Ficoll 400
5% polyvinylpyrrolidine (PVP)
10 mg ml^{-1} BSA fraction V

11.1x PCR mix:

495 mM TRIS pH 8.8
123 mM $(NH_4)_2SO_4$
50 mM $MgCl_2$
0.5% 2-mercaptoethanol
50 μM EDTA pH 8.0
11.1 mM each dNTP
1.25 mg ml^{-1} BSA

Phage Buffer:

100 mM NaCl
50 mM TRIS pH 7.5
10 mM $MgSO_4$
0.01% gelatin

10x SPRB:

250 mM TRIS pH 8.3
1.92 M glycine
1% SDS

FSB:

200 mM TRIS pH 6.8
400 mM 2-mercaptoethanol
8% SDS
40% glycerol
A few crystals of bromophenol blue

Coomassie stain:

0.5% Coomassie Brilliant Blue R-250
30% methanol
10% acetic acid

Fast destain:	45% methanol 10% acetic acid
Slow destain:	10% methanol 5% acetic acid
Towbin transfer buffer:	25 mM TRIS pH 8.3 93 mM glycine 20% methanol
1x TBST:	20 mM TRIS pH 7.6 300 mM NaCl 0.00005% Tween-20
Blotto:	5% nonfat dried milk in 1x TBST 10% horse serum
LS:	50 mM MOPS pH 7.2 100 mM NaCl 1 mM EDTA pH 8.0 1 mM EGTA pH 8.0 1 mM sodium orthovanadate 10 mM NaF 1% TritonX-100
LSI:	Same as LS with the following protease inhibitors: 100 mg ml-1 leupeptin 5 mg ml-1 pepstatin A 500 mg ml-1 Pefabloc SC 1.25 mM 1,10-phenanthroline
Aminolink reductant:	1.0 M sodium cyanoborohydride in 10 mM NaOH
HSLS:	Same as LS but with 500 mM NaCl
LSG:	Same as LS but with 10% glycerol

LSGI:	Same as LSG but with protease inhibitors as in LSI
QIAexpress buffer A:	6 M guanidine hydrochloride 0.1 M NaH_2PO_4 0.01 M TRIS Adjusted to pH 8.0 with NaOH
QIAexpress buffer B:	8 M urea 0.1 M NaH_2PO_4 0.01 M TRIS Adjusted to pH 8.0 with NaOH
QIAexpress buffer C:	Same as buffer B but pH 6.3
QIAexpress buffer D:	Same as buffer B but pH 5.9
QIAexpress buffer E:	Same as buffer B but pH 4.5
QIAexpress buffer F:	6 M guanidine hydrochloride 0.1 M acetic acid
SB:	50 mM sodium phosphate pH 7.8 300 mM NaCl
GEB:	10 mM reduced glutathione in 50 mM TRIS pH 8.0
KAB:	50 mM MOPS pH 7.2 20 mM $MgSO_4$ 10 mM EGTA 2 mM DTT
KAM (10 reactions):	186 µl KAB 8 µl 100 mM dATP 5 µl 10 mg ml^{-1} protein substrate 1 µl $^{\gamma 32P}$-dATP

CHAPTER 3: ISOLATION AND ANALYSIS OF *tbCRK4*

3.1: Introduction

Kinetoplastid *CRK* genes fall into four distinct classes, based on sequence analysis and comparisons between examples from different organisms. *CRK1* has been isolated from both *Trypanosoma* and *Leishmania* species; *lmmCRK1* is expressed in the proliferative promastigote and amastigote forms, although it only displays histone HI kinase activity in the promastigote form (Mottram *et al.*, 1993). *CRK3* has also been isolated from both organisms, and is a strong contender for a functional *CDC2* homologue as it correlates with SUC-binding CRK kinase (SBCRK) activity (Grant *et al.*, 1998). *CRK2* has only been isolated to date from *T. b. brucei*, and has a distinct N-terminal extension sequence of 41 amino acids as compared to other CDKs (Mottram and Smith, 1995). Phylogenetic analysis clearly shows that these genes bear no greater homology between themselves than to CDK genes from other eukaryotes, thus forming three classes (Mottram and Smith, 1995).

The first CRK to be isolated was detected using a PCR-based approach with degenerate primers. Lisa Brown and Dan Ray (UCLA) amplified and isolated a gene fragment from *Crithidia fasciculata* (Brown *et al.*, 1992) which had significant homology to *CDK* genes from other eukaryotes. Isolation and analysis of the full-length *cfCRK* gene revealed it to be novel member of the *CDK*-like family. While *cfCRK* possessed the usual features of this kinase subfamily it had in addition two large insert sequences, of 66 and 79 amino acids, located between kinase catalysis domains VIb and VII and domains X and XI, respectively (Fig. 3.10). Multiple alignment of kinase catalysis domains I-VIb and VII-X with the same domains from *CDC2* homologues from human, zea mays, drosophilia melanogaster and budding yeast *CDC28* showed a high level of conservation. Kinase catalysis domain XI, in comparison with domain XI from *CDC2*/*CDC28* from the aforementioned organisms, showed less conservation, consistent with what has been observed for *CDC2* homologues. Overall, *cfCRK* was most similar to mouse *CDC2*, with 47% identity and 67% similarity. In the PSTAIRE motif cfCRK had four amino acid substitutions as compared to human CDK1. Human CLK1 is a serine/threonine protein kinase more distantly related to CDK1 than the rest of the CDK family; it contains an insert domain between kinase catalysis domains X and XI, therefore corresponding to the second insert domain of cfCRK. However, there was no significant similarity between the insert in CLK1 and the inserts in cfCRK. FASTA searches of the GenBank database using the sequences of the two insert domains of cfCRK did not find any alignments of significance. Southern blotting revealed that *cfCRK* was a single copy gene, and Northern blotting using total RNA and poly[A]+ RNA, probed with a fragment of the gene covering the majority of the kinase domains,

detected a 3.8 kb polyadenylated transcript. Northern blotting using total RNA and poly[A]+ RNA, using a fragment produced by PCR amplification of the first insert domain only, also detected a polyadenylated 3.8 kb transcript. The size of the transcript was larger than expected, as the coding sequence of the transcript was approximately 1.4 kb. Although no trypanosome gene has been found to contain introns, the presence of two inserts in *cfCRK* which otherwise has a high degree of sequence identity with *CDC2*-related genes, suggested that the inserts could represent introns. The Northern blot data repudiated this. A 60 amino acid cfCRK peptide was expressed as a fusion protein with glutathione-S-transferase and used to raise a polyclonal rabbit antiserum. Western blotting with this anti-cfCRK antiserum detected a protein of the predicted size (53 kDa) in *C. fasciculata* whole cell extracts. In addition, the anti-cfCRK antiserum detected a protein of 48 kDa in *C. fasciculata* whole cell extracts that cross-reacted to an antiserum raised to the 16-residue PSTAIR motif of CDC2, suggesting that *C. fasciculata* may contain a protein with a more highly conserved PSTAIR motif. Further PCR-based experiments resulted in the isolation of a further three *C. fasciculata* kinase gene fragments, all of which clearly belonged to this new category of *CRK* genes, based on sequence comparisons and the presence of insert domains (Lisa Brown, pers. comm.). Interestingly, there was no significant homology between the two insert domains of *cfCRK*, renamed *CRK4* as *CRK1-3* had been described previously (Mottram, 1994), and the two insert domains of the other three cloned gene fragments.

A trypanosome homologue of *cfCRK* was identified by Dr. Noel Murphy (ILRAD, Nairobi) as a partial gene fragment clone during a search for *CDC2*-related kinases using PCR with degenerate primers. This fragment encoded kinase catalysis domains I-VIb and a first insert domain, and by sequence analysis was clearly a *CRK*; the presence of the insert fitted it into the *CRK4* subclass. Dr. Murphy provided the sequence of the PCR fragment with the aim of isolating the full-length gene.

3.2 Results

3.2.1: Isolation of *CRK4* fragments by PCR

Initially, a PCR-based approach was undertaken to attempt to isolate clones representing full-length *CRK4*. Oligonucleotides OL355 and OL356 were designed as gene-specific primers to amplify a gene-internal fragment (Fig. 3.1). To amplify the 5' end of the gene the 39 base-pair splice leader (SL) sequence was exploited, with a primer corresponding to the SL sequence being used in conjunction with an internal antisense primer OL356; all trypanosome pre-mRNAs are *trans*-spliced to a small leader RNA (Agabian, 1990; Perry and Agabian, 1991). To

amplify the 3' end of the gene the poly[A] tail of the 3' end of the transcript was exploited. Total RNA was isolated from *T. b. brucei* STIB 247 procyclics using the Qiagen Total RNA kit as described (chapter 2.2.8) and first-strand cDNA synthesis was performed (chapter 2.2.8) using the poly[dT]-anchor primer.

To isolate the gene-internal fragment, predicted from known sequence to be 760 bp (N. Murphy, unpublished) a number of PCR reactions were performed at differing annealing temperatures using primers OL355 and OL356, with a dominant fragment of approximately 760 bp expected to persist at higher annealing temperatures (Fig. 3.2). At 37°C annealing four main fragments were seen, with the dominant one at approximately the correct size (lane 2); at 42°C annealing the three subsidiary fragments were much fainter (lane 4), and at 48°C they were no longer present, leaving only the 760 bp fragment (lane 6). This fragment was excised, gel purified and cloned to pTAG vector (pGL138). Automatic sequencing using the AmpliTaq kit (Perkin-Elmer) and the vector-specific primers pTAG5' and pTAG3' showed the fragment to have identical sequence to *CRK4*. Completion of full-strand sequence of this fragment with the pTAG vector primers and the *CRK4*-specific primers shown (Fig. 3.1) showed it to be the expected 760 bp fragment from *CRK4*, with no base substitutions as compared to Dr. Murphy's sequence.

To isolate the 5' end of *CRK4* an annealing temperature range set of reactions were done with OL356 and the SL primer on a Robocycler machine (Stratagene). 12 identical reactions across a range from 42-62°C were performed, thus providing an approximate annealing temperature difference of 1.7°C between reactions (reactions 1-7, Fig. 3.3). Because of the loss of specificity due to using the non-transcript-specific SL primer it was expected that more nonspecific fragments would be observed, particularly at the lower end of the temperature range, but that far fewer dominant fragments would persist at the high end of the temperature range. At the higher temperatures a number of fragments was detected, the major fragment of 0.5 kb. A PCR reaction at an annealing temperature of 60°C was performed to purify this fragment for cloning; however, after the reaction was electrophoresed on a 1% TAE gel and visualised, fragments that had not previously been detected were observed at approximately 0.85 kb and 1.0 kb (not shown). The predicted size of the amplification product was a minimum of 0.9 kb. The 0.5 kb, 0.85 kb and 1.0 kb fragments were cloned into pTAG vector. Preliminary sequencing of pTAG/1.0 kb (pGL139) using the pTAG vector primers showed CRK4 sequence, and consequently the 0.85 kb and 0.5 kb subcloned fragments were not sequenced. Completion of full-strand sequence of the pGL139 insert using the pTAG vector primers and the *CRK4*-specific primers shown (Fig. 3.1) covering that region, gave the 5' gene sequence and showed the splice site.

To isolate the 3' end of *CRK4* primer OL355 was used with the complementary anchor primer. A number of reactions were performed, differing only in the annealing temperature

used; only two fragments persisted above moderate annealing temperatures of 50-55°C, at approximately 1.0 and 1.4 kb (Fig. 3.4, lane 2). A calculation of the minimum amplification product size for the 3' end of the gene, required to encode the remaining kinase catalysis domains without a second insert domain, was approximately 1.3-1.4 kb. Consequently, only the 1.4 kb fragment was cloned to the pTAG vector. Sequencing using the vector-specific primers pTAG5' and pTAG3' showed a poly[dT] sequence and also gave an upstream sequence that was identified as the triosephosphate isomerase gene. The triosephosphate isomerase transcript is an abundant mRNA and is therefore one of the most highly represented components of an expression library. OL355 corresponds to the ATP-binding pocket of CRK4, which may explain the erroneous amplification of the triosephosphate isomerase cDNA. As the 3' end of the *CRK4* gene could not be isolated by PCR methods it was decided to isolate the gene from a genomic λphage library.

3.2.2: Isolation of *CRK4* by library screening

To attempt to isolate clones derived from genomic DNA for comparison to those isolated by PCR, and to obtain the 3' end of *CRK4*, a λFIXII *T. b. brucei* genomic library was screened. *Sau*3AI partial digestion of *T. brucei* DNA and digestion of the phage with *Xho*I to remove the stuffer fragment were used to construct the library. Both the digested phage and partially digested parasite DNA were subjected to a partial fill-in reaction to generate compatible cohesive ends, ligated and packaged. After establishing the titre a primary screen was carried out as described (chapter 2.2.10) with a low stringency hybridisation temperature of 50°C. The rationale for the low stringency screen was that given the evidence of a multiple *CRK4* gene family in *C. fasciculata*, it was extrapolated that such a family may also be present in *T. brucei*, hence a lower hybridisation temperature increased chances of detecting other homologues of *tbCRK4*. The insert from pGL138 was used as the probe in all rounds of phage screening; this fragment contains the first insert domain, and so was likely to confer added specificity to the probe for the *CRK4*.

The primary screen yielded only three clear but weakly hybridising plaques. The three primary plaques were titred and put through a second round of screening. Each gave approximately 15% positive plaques per plate. Six hybridising plaques were lifted from each of these secondary screens. Two secondary plaques representing each primary plaque were then chosen for tertiary screening. A single unique plaque from the tertiary screen was picked for further analysis

The three clones were labelled IA, IIA and IIIA; phage DNA was purified as described (chapter 2.2.10) for these clones and an initial analysis was carried out by performing restriction

endonuclease digestion. The enzymes used were *Not*I, *Sal*I, *Sac*I and *Xba*I, all chosen on the basis of site representation within the phage's polylinker. These sites sit outside the *Xho*I site in the phage's polylinker that was used in the generation of the library, and in addition none of them cut within the known sequence of *CRK4*. *Hin*cII was also used; this cuts within the phage and also several times within the known sequence of *CRK4*, however as a frequent cutter it was used to provide a multiple fragment pattern to compare the three clones. Therefore, restriction digestion (except in the case of *Hin*cII) was expected to produce the phage arms of approximately 23 and 9 kb, with a pattern of fragments representing the fragments generated from the genomic inserts. The fragment pattern produced from this initial restriction digestion (Fig. 3.5) clearly showed clones IA and IIA to have similar but not identical inserts; the notable differences between the two were in the *Sal*I digest, where IIA showed an additional two fragments at approximately 1.8 and 3.2 kb (IIA, lane 3) that were not seen with IA (IA, lane 3). In the *Xba*I digestion IA (lane 5) produced a single small fragment of 1 kb, whilst IIA (lane 5) produced a 1.2 kb fragment. Clone IIIA produced a completely different pattern of fragments. For clones IA and IIA, there were few fragments produced from *Xba*I (lanes 5) and *Sac*I (lanes 4), and none visible from *Not*I (lanes 2). The pattern of fragments in the *Hin*cII digestions seemed identical between clones IA and IIA (lanes 6). For clone IIIA, *Hin*cII sites were again frequent (lane 6) and there was a 1.1 kb fragment produced from *Not*I (lane 2). In contrast to IA and IIA, *Xba*I and *Sac*I each generated two large molecular weight fragments for IIIA (lanes 4 and 5), and *Sal*I generated only one or two non-phage-arm fragments, with a small fragment at about 1.1 kb (lane 3). Therefore clones IA and IIA represented similar genomic fragments, while clone IIIA was completely different.

Southern blotting was performed using the same restriction enzymes on purified DNA of the three phage clones; the probe used was again the insert from pGL138 (Fig. 3.6). For clones IA and IIA only very large molecular weight hybridising fragments were detected with *Not*I, *Sac*I and *Xba*I (lanes 1, lanes 3, lanes 4), indicating that CRK4 resides on a large DNA fragment. The *CRK4* probe hybridised to a 2.5 kb *Sal*I fragment for both clones (lanes 2), and a 0.5 kb *Hin*cII fragment (lanes 5). The pattern of hybridising fragments was identical for clones IA and IIA, supporting the view from the initial restrictions that these clones represent overlapping genomic fragments. Clone IIIA gave a totally different pattern of hybridising fragments; the *Not*I digest produced a single fragment of about 1.1 kb (lane 1). *Sal*I and *Sac*I produced large molecular weight single fragments (lanes 2 and 3), and *Xba*I gave a single fragment at around 5 kb (lane 4). Interestingly, the *Hin*cII restriction gave the same-sized hybridising fragment at approximately 0.4 kb (lane 5) as seen for clones IA and IIA, and also gave a very weakly hybridising fragment at an even smaller size.

3.2.3: Subcloning and sequencing of *CRK4*

Subclones representing the entire *CRK4* gene, with some flanking sequence, were cloned from the λ DNA. To obtain the 5' end of the gene plus some 5' flanking sequence, both clones IA and IIA were digested with *Sal*I and the hybridising 2.5 kb fragment cloned to the *Sal*I site in pBluescript. PCR using the T3 and T7 vector primers was performed to identify colonies that carried inserts of the correct size. For each λ clone, two independent 2.5 kb *Sal*I subclones were isolated and DNA prepared for sequencing. To obtain the 3' end of the gene and possibly some 3' flank sequence, clone IIIA was digested with *Not*I and the hybridising 1.1 kb fragment cloned to the *Not*I site of pBluescript. Screening and selection resulted in two positive 1.1 kb *Not*I subclones being isolated. DNA was prepared from each subclone for sequencing.

Prior to sequencing of the six subclones, a restriction digestion was performed to check that subcloned fragments would excise properly with the relevant enzymes; in addition, *Hinc*II and *Sau*3AI were also tested on all subclones, to provide a fragment pattern indicative of unique clones. The restriction patterns for subclones IIA/1 and IIA/2 were identical, but between the IA subclones there were clear differences in pattern with the *Hinc*II and *Sau*3AI digests; subclone IA/1 was clearly identical to both IIA subclones, while subclone IA/2 was quite different from these three (data not shown). Subclones IIIA/1 and IIIA/2 produced identical fragment patterns (not shown). A Southern blot of the six subclones digested with *Sal*I and *Not*I, as appropriate, and using the insert from pGL138 as a probe was carried out to determine which of the 2.5 kb subclones contained the predicted *CRK4* sequence; the 1.1 kb subclones were also tested for hybridisation. The autoradiograph showed hybridisation to both IIIA *Not*1 1.1 kb fragments but hybridisation only to subclone IA/2 *Sal*I 2.5 kb fragment, therefore subclones IA/1, IIA/1 and IIA/2 represented erroneous nonhybridising sequences, whilst subclones IA/2, IIIA/1 and IIIA/2 contained *CRK4* sequence (data not shown).

Sequencing using the vector's T3 and T7 primers was performed on the 3 positive *CRK4* subclone plasmids, IA/2, IIIA/1 and IIIA/2. The results of preliminary sequencing confirmed the restriction analysis: subclone IA/2 was identified as *CRK4* sequence with the T3 primer; this showed antisense sequence from the *Sau*3AI site at position 560 (relative to the A of ATG start) towards the 5' end of the gene. Therefore this subclone (pGL119) represented the 5' end of the gene and approximately 2 kb of upstream flank sequence. IIIA/1 and IIIA/2 subclones (pGL140 and pGL141,) both gave identical sequence with the T3 primer. The sequence obtained was again antisense, this time from kinase catalysis domain X towards the *Sau*3AI site involved in the generation of the library. At this point it could not be determined if the 2.5 kb *Sal*I fragment from IA/2 and the IIIA 1.1 kb *Not*I fragments were contiguous sequence.

Southern blotting of λFIXII clone IIIA DNA digested with a range and combinations of

enzymes was carried out, to identify a hybridising fragment larger than the 1.1 kb *Not*I subcloned fragment, and so to isolate a subclone that represented the complete 3' end of the gene plus some 3' flanking sequence (Fig. 3.7). Only three of the restrictions identified a hybridising fragment that was a suitable size for subcloning: The *Nco*I digest had a fragment at approximately 2.2 kb (lane 6). As *CRK4* has an *Nco*I site at position 578 (relative to A of ATG start) a 2.2 kb fragment would be predicted to contain the end of the gene plus approximately 1.4 kb of 3' flank sequence. The *Xba*I-*Xmn*I digest also showed a 2.2 kb fragment (lane 1), by contrast the *Sac*I-*Kpn*I digest showed a hybridising fragment at approximately 1.6 kb (lane 2), both of which would be predicted to contain less 3' flank than the 2.2 kb *Nco*I fragment. It was decided to subclone the *Sac*I-*Kpn*I fragment as although this fragment represented less 3' flank sequence, the restriction sites involved were judged more convenient for use in subcloning. Inspection of clone IIIA DNA digested with *Sac*I-*Kpn*I and electrophoresed on TAE midigels revealed that there were three fragments relatively close together; one fragment just over 1.5 kb, and a very close doublet at about 1.6 kb. Since it was impossible to tell from the Southern blot data which of these fragments was the one that hybridised with the *CRK4* probe, all three were cloned to pBluescript. PCR using primers OL348 and OL356 were used to screen for positives, with an expected amplification product of approximately 450 bp. The results showed that the slightly larger of the two fragments at 1.6 kb was the fragment that hybridised to the *CRK4* probe. This fragment was cloned into pBluescript, to give pGL218.

3.2.4: Analysis of *CRK4* sequence

Prior to complete sequencing of the three subclones, a Southern blot of STIB 247 genomic DNA digested with a variety of restriction enzymes was carried out, and probed with the insert from pGL138 (Fig. 3.8). Only high molecular weight hybridising fragments were observed except in the case of *Hin*dIII (lane 5) which showed a hybridising fragment at approximately 1.6 kb, and *Hin*cII (lane 4) which gave a hybridising fragment of approximately 0.4 kb. Following complete double-stranded sequencing of the gene and flanking sequences from subclones pGL119 (IA/2 *Sal*I 2.5 kb), pGL140 (IIIA/1 *Not*I 1.1 kb), pGL218 (IIIA *Sac*I-*Kpn*I 1.6 kb), the Southern blot data were corroborated. The sequence of the PCR-generated fragment in pGL138 overlapped pGL119 and pGL140, thus proving that they represented contiguous sequences. The complete sequence of *CRK4* with flanking sequences, primers and relevant restriction sites is shown (Fig. 3.9). From the region covered by the probe (insert from pGL138) the 0.4 kb *Hin*cII fragment detected on the Southern blot (Fig. 3.8) represents the fragment of 0.35 kb generated by cleavage at the *Hin*cII sites at positions 763 and 413 (relative to the A of ATG start). The 1.6 kb *Hin*dIII fragment detected represents the fragment of 1.59 kb

generated by the cleavage at the *Hin*dIII sites at positions 1623 and 783 (relative to the A of ATG start). As shown in the earlier schematic (Fig. 3.1), *tbCRK4* is the same as *cfCRK* in possessing two large insert domains, between kinase catalysis domains VIb and VII, and domains X and XI. Multiple alignment of tbCRK4 and cfCRK primary sequence with tbCRKs 1-3 and *S. cerevisiae* CDC28 shows that for both CRK4s the Insert Domains 1 and 2 are almost exactly the same size between cfCRK and tbCRK4 (Fig. 3.10). Multiple alignment was done using ClustalW at EBI (http://www2.ebi.ac.uk/clustalw/). All the kinase catalysis domains present in CRK1-3 are present in tbCRK4, including the conserved threonine and tyrosine residues (for tbCRK4, Thr17 and Tyr18) at the ATP binding pocket, the conserved threonine 161 equivalent (for tbCRK4 and cfCRK this residue is a serine, at positions Ser231 and Ser230, respectively). The PSTAIRE box equivalent is present, through degenerate, with five substitutions (GAPPSTAIREIALLKV) as compared to human CDK1 (EGVPSTAIREISLLKE). The DSEI box equivalent is also present although highly degenerate with four substitutions (GNTDVDQ) as compared to human CDK1 (GDSEIDQ). There is another conserved residue at position 425 - a serine residue - that is also a serine at the equivalent position in tbCRK3 and CDC28, and a threonine in tbCRK1, tbCRK2 and cfCRK. The striking feature is that there is no significant homology between the insert domains for the two genes. BLAST searches for the insert domains only of tbCRK4 failed to produce any significant matches for any of the databases used (data not shown). Further differences were predicted for the two proteins using tools local to the ExPasy server (http://www.expasy.ch/tools/). The cfCRK protein was predicted to have a molecular weight of 53.5 kDa and a p*I* of 6.4, whereas tbCRK4 was predicted to be slightly smaller -as expected - at about 51.4 kDa, with a p*I* of 8.5.

Similarities between tbCRK4 and cfCRK are revealed by other analysis, however. Hydropathy plotting using the ProtScale tool at the ExPasy server with the method of Hopp & Wood (Fig. 3.11) shows that Insert Domain 1 for both proteins is predicted to be largely hydrophilic (tbCRK4 upper panel, residues 140-210; cfCRK lower panel, residues 143-210). Insert Domain 2 for both is less uniformly hydrophilic (tbCRK4, residues 312-390; cfCRK, residues 311-390) but both contain a highly hydrophilic sequence of residues, for tbCRK4 from about amino acids 310-350 and for cfCRK from about amino acids 305-325.

3.2.5: Production of recombinant GST-CRK4

The GST Fusion system (Pharmacia Biotech) was used for the production of recombinant GST-CRK4 fusion protein, to raise a specific anti-CRK4 antiserum, for use in Western blotting and immunoprecipitations. This system involved the large immunogenic GST tag, which could be problematic to remove if removal of the tag was necessary. However, GST

is soluble and may help less soluble peptides to produce soluble protein as a fusion.

The segment of *CRK4* used was a *Hinc*II-*Hae*III fragment of 244 bp, representing almost exactly the first insert domain with virtually no kinase domain sequence (see Fig. 3.1). The rationale was that as this insert domain does not have significant homology to any protein sequence in the SwissProt database, it would be less likely to raise antibodies that could cross-react to other endogenous trypanosome proteins. The GST tag is 26 kDa, so the predicted size of the GST-CRK4 fusion protein was approximately 36 kDa.

Plasmid pGL138 was digested sequentially with *Hinc*II and *Hae*III, as the buffers for the enzymes are not compatible. The reaction was electrophoresed on a 1% TAE midigel, the 244 bp fragment excised, gel purified and cloned into the *Sma*I site of pGEX-5X-2. Transformants were screened by Southern blotting: Colony streaks were lifted to supported nitrocellulose filter and probed with the insert from pGL138. Eight positives were detected and further screened by PCR. Two clones containing the insert fragment in the correct orientation were detected; one of the clones was chosen, and DNA was prepared for restriction and sequencing. Restriction analysis and PCR reactions confirmed that this clone, pGL134 (Fig. 3.12) contained the insert fragment in the correct orientation. Sequencing of the 5' ligation join using primer OL350 was performed using a manual kit (USB) with the manganese modification, rather than automatic sequencing, as the ligation join was less than 50 bp from the primer. The sequence obtained showed that the insert fragment had integrated in the correct orientation and with the reading frame preserved. Consequently, pGL134 was transformed to the protease-deficient *E. coli* strain BL21 for pilot expression and purification of GST-CRK4.

A 20 ml culture in LB with 100 $\mu g\ ml^{-1}$ ampicillin was inoculated at 1:100 with a saturated overnight culture of BL21pGL134, and grown to an optical density of 0.8 at 600 nm. An uninduced control sample was taken, and IPTG added to a final concentration of 2 mM. Induced samples were then removed hourly, up to four hours. The remainder of the 4-hour induced culture (10 mls) was centrifuged at 10 000*g* for 10 minutes at 4°C, the medium decanted and the cell pellet frozen at -20°C. Crude whole cells lysates were prepared from the uninduced and induced samples (described in chapter 2.3.2), electrophoresed on a 12.5% Lammeli gel, stained with Coomassie blue and dried down (Fig. 3.13). The Lammeli gel clearly showed the induction of two proteins (lanes 3-6), although at 30 kDa and 44 kDa neither matched the predicted size of 36 kDa for recombinant GST-CRK4.

A solubility check was performed (Fig. 3.14) using the frozen cell pellet from the pilot expression experiment (above). The cell pellet was thawed, resuspended in sonication buffer and was incubated on ice for 30 minutes following addition of lysozyme to 1 $mg\ ml^{-1}$. Subsequently, they were put through 6 freeze/thaw cycles in dry ice/ethanol and a 100°C heating block and then spun at maximum speed in a bench microfuge for 20 minutes. The supernatant was aspirated and the cellular debris resuspended in 1 ml sonication buffer by syringing through

a G25 needle. 400 µl of FSB was added to the supernatant (soluble protein) and the cellular debris (insoluble protein), the samples heated to 100°C for 5 minutes and aliquots electrophoresed on a 12.5% Lammeli gel, the gel stained with Coomassie blue and dried. The same two proteins previously noted, of 30 kDa and 44 kDa, were observed in the soluble protein lysate (lane 3) but not in the insoluble protein lysate (lane 2). However, the insoluble protein sample was very dilute.

A small-scale native purification for GST-CRK4 was performed as described (chapter 2.3.3), and samples of insoluble protein, soluble nonbound protein and eluate protein electrophoresed on a 12.5 % Lammeli gel and visualised by Coomassie staining (data not shown). A 44 kDa protein but not a 30 kDa protein was detected in the insoluble protein sample; the 30 kDa protein showed in the soluble nonbound protein samples although the 44 kDa protein did not, but it was present with very little apparent contamination in the eluate protein sample. The 30 kDa protein did not show in the eluate samples. The presence of the 44 kDa protein in the insoluble (cellular debris) sample was likely a result of large over-expression of the recombinant protein, either allowing some soluble recombinant protein to be brought down in the pellet or representing some recombinant protein forming inclusion bodies. There was a disparity of 8 kDa between the predicted size of GST-CRK4 (36 kDa) and the 44 kDa protein selected by the GS4B resin. The isoelectric point (p*I*) of the CRK4 peptide was predicted to be 4.4, therefore it was possible that the molecular weight disparity was a result of retardation on the Lammeli gel due to the CRK4 peptide's acidity.

Based on the pilot assays, it was decided to use the GST system for a moderate-scale induction and purification of GST-CRK4. 4x 500 ml LB were each inoculated at 1:100 with a saturated overnight culture of BL21pGL134, grown to an OD at 600 nm of 0.8, induced by the addition of 0.5 mM final concentration IPTG and grown for a further 4 hours. Cells were harvested by centrifugation at 7 500*g* for 10 minutes at 4°C, cell pellets resuspended in a total of 100 ml 1x PBS pH 7.2 and lysed by sonication. The lysate was centrifuged at 12 000*g* for 10 minutes at 4°C to clear the cellular debris, the supernatant aspirated and applied to a GS4B resin column. The column was washed with 1x PBS pH 7.2 until the protein content of the flow-through was negligible as judged by absorbance at 280 nm, and bound GST-CRK4 was eluted with 3x 3 ml glutathione elution buffer. Eluate fractions were dialysed overnight against 2 000 ml 1x PBS pH 7.2. Samples of the dialysed eluate fractions were electrophoresed on a Lammeli gel to determine which fraction contained the bulk of the GST-CRK4 protein, along with samples of the crude soluble lysate, column flow-through and first wash fraction (Fig. 3.15). The column appeared to be overloaded as the induced 44 kDa protein was detected in the flow-through (lane 3) as well as the wash (lane 4). However, by far the majority of GST-CRK4 was in the first eluate fraction (lane 5), with a lesser amount in the second eluate fraction (lane 6). A number of other proteins were detected in each eluate fraction, all of which were below 40 kDa

and may have been degradation products. Samples of the first and second eluate fractions were bound separately to an excess of GS4B resin, washed thoroughly, eluted and electrophoresed on a Lammeli gel (not shown). The profile of proteins remained the same, indicating that all the smaller proteins were specifically binding the resin. The concentration of protein in the first eluate fraction was assessed using the Biorad DC protein assay kit, and found to have a GST-CRK4 concentration of 3 mg ml^{-1}. 200 μg of the first eluate fraction was used in the primary immunisation of two New Zealand rabbits, essentially as described in chapter 2.3.4. The antisera from the two rabbits were labelled anti-CRK4/A and anti-CRK4/B on receival. The purified GST-CRK4 protein in eluate fractions 1 and 2 were dispensed to eppendorfs as 100 μl aliquots and stored at -20°C.

3.2.6: Western blotting with anti-CRK4 antibodies

After exsanguination of both animals, test blots with both antisera showed that anti-CRK4/B gave more promising results (data not shown), therefore it was an aliquot of this antiserum that was purified by the two-step affinity purification. This involved one purification step on an immobilised GST column to remove anti-GST antibodies, and a second purification step on an aminolinked GST-CRK4 column to specifically purify only anti-CRK4 antibodies (for details see chapter 2.3.4). After the first step, removal of anti-GST antibodies, the flow-through fraction was tested against protein lysates of STIB 247 long slender bloodstream, short stumpy bloodstream and procyclic lysates, and also against TREU 869 long slender bloodstream lysates (Fig. 3.16). A doublet at 46-48 kDa was detected very weakly in the soluble protein lysate of all three STIB 247 life cycle stages (lanes 1, 3, and 5) that was competed with purified GST-CRK4 fusion protein (lanes 9, 11 and 13). Although intensity of signal varied both proteins were also present to lesser degrees in the insoluble protein lysates (lanes 2, 4, and 6). The larger of this doublet was present in both soluble and insoluble lysates of TREU 869 (lanes 7 and 8), although the smaller protein was not. Proteins of 55-60 kDa were detected at in the STIB 247 long slender bloodstream form soluble and insoluble lysates (lanes 1 and 2) that were competed with fusion protein. A protein of 32 kDa, not competed with purified GST-CRK4 fusion protein, was detected in the soluble lysate of all three STIB 247 life cycle stages (lanes 1, 3 and 5) and also in the short stumpy insoluble protein lysate (lane 4). This protein was also detected in both the soluble and insoluble lysates of the TREU 869 long slender bloodstream form.

Following the second purification step of specific purification of anti-CRK4 antibodies and dialysis of protein-containing pools, a Western was performed with these antibodies using fresh S-100 soluble protein lysates were made from all three STIB 247 life cycle stages; from freshly-harvested procyclics at midlog, and from freshly purified long slender bloodstream and

short stumpy bloodstream form. The lysates were electrophoresed on Lammeli gels, transblotted and probed with the purified anti-CRK4 antibodies (Fig. 3.17). Two proteins were observed in the procyclic lysate (lane 3), the 53 kDa protein being specifically competed by preblocking of the antibodies with a saturating quantity of fusion protein (lane 6), indicating the specificity of the reaction. The 53 kDa protein was also present in the short stumpy bloodstream form (lane 2), although the signal was far weaker, and was competed with fusion protein (lane 5). The 53 kDa protein was absent from the long slender bloodstream form (lane 1). There was also a slightly smaller protein, detected with weak intensity, present in the long slender bloodstream and short stumpy bloodstream form lysates (lanes 1 and 2) that appeared to be specifically blocked (lanes 4 and 5), and a slightly larger protein of weak intensity in the short stumpy bloodstream form lysate (lane 2) that was not competed with fusion protein (lane 5). A modified Western protocol was followed (Fig. 3.18) with the blocking step carried out at room temperature for 8 hours and the blot incubated with primary antibodies at a lower concentration and overnight at 4°C. This resulted in disappearance of all but the 53 kDa protein detected in procyclics (lanes 2 and 4) and more weakly in the short stumpy bloodstream form (lanes 1 and 3); for both lysates this protein was competed with fusion protein (lanes 5-8). The conclusion would therefore be that the purified anti-CRK4/B antibodies specifically detects CRK4 in soluble fraction protein lysates of STIB 247 short stumpy bloodstream and procyclic form, but not in soluble lysates of the long slender bloodstream form. Quantitatively, CRK4 abundance would appear to be low in the short stumpy bloodstream form, raised to a higher level in procyclics, and that the protein is not present in the long slender bloodstream form. The apparent size of the protein correlates well with the size of approximately 51-52 kDa as predicted from the gene sequence.

3.2.7: Selections using recombinant p12 and p13

To establish that CRK4 has one property of a CDK, that of binding to the small cofactor protein p13, from yeast, or p12, from *Leishmania mexicana*, procyclic S-100 soluble protein lysates were selected with beads cross-linked by amino linkage to either recombinant $p12^{CKS1}$, $p13^{SUC1}$ or to TRIS (control beads). Selection was carried out as described in chapter 2.3.6. Protein eluted from the beads was electrophoresed on 12.5% Lammeli gels, transblotted and probed with the purified anti-CRK4/B antibodies (Fig. 3.19, upper panel). Two proteins of approximately 60 and 66 kDa were found to bind p12 (lane 2) and p13 (lane 3) but also control beads (lane 1). The cross-reaction with the antibodies was strong, and appeared to be partially abrogated by preblocking the antibodies with fusion protein (lanes 4-6). This result could be repeated with several attempts at this experiment. In no case was a protein detected at the predicted size of 53 kDa, although in some of the blots the resolution was sufficiently low that

any signal from CRK4 could have been swamped by the signal from the two nonspecific proteins. To attempt to resolve this, lower percentage acrylamide Lammeli gels, with a resolving section of 7.5%, were run to achieve greater separation of proteins in this region. When probed with the purified anti-CRK4/B antibodies (Fig 3.19, lower panel) the same two nonspecific proteins were detected (lanes 7-9), although there appeared to be less blocking with fusion protein than previously observed (lanes 10-12). The 53 kDa CRK4 was never detected, suggesting that CRK4 does not bind either p12 or p13. Although internal positive controls were not performed, the same bead preparations were routinely being used by other researchers with the expected results, and it was therefore assumed that lack of CRK4 binding was not due to the beads.

3.2.8: Immunoprecipitation using purified anti-CRK4 antibodies

The purified anti-CRK4/B antibodies were used to immunoprecipitate CRK4 from S-100 soluble fraction protein lysates of STIB 247 procyclics, and linked to kinase assays, to try to demonstrate a CRK4-associated kinase activity. As Western blots had shown that the candidate polypeptide was present at greatest quantity in procyclic form soluble extracts, procyclic S-100 soluble lysates were incubated with either preimmune serum, purified anti-CRK4/B antibodies or purified anti-CRK4/B antibodies preblocked with a saturating amount of recombinant fusion protein, and kinase assays performed. Antibody-CRK4 complexes were bound to Sepharose-protein A beads, washed extensively to remove all other soluble proteins and resuspended in a kinase assay mix, containing $^{\gamma 32p}$-dATP and a substrate protein at a concentration of 5 $\mu g\ ml^{-1}$ (for full details see chapter 2.3.5). The pilot assay used phosvitin as a substrate (Fig. 3.20) as phosvitin has been demonstrated to be a substrate for CDK phosphorylation. A stronger kinase activity was detected with the preimmune serum (lane 1) than the anti-CRK4/B antibodies (lane 2) or the anti-CRK4/B antibodies preblocked with purified GST-CRK4 fusion protein (lane 3). It is always difficult to predict a good substrate for kinases, but as phosvitin had been used to assay the kinase activity of other CDC2-related kinases it was chosen as a possible substrate for CRK4; however, the result of the pilot kinase assay suggested that phosvitin may simply not be a substrate for CRK4. Further attempts were carried out with a range of substrates, including phosvitin, α-casein, β-casein, protamine and a histone mix. No specific kinase activity was obtained with any of these substrates (Fig. 3.21). As no significant difference between the anti-CRK4 antibodies and the antibodies preblocked with recombinant protein was found, with the exception of a small difference for α-casein (lanes 3 and 4), and in the case of protamine (lane 2), which showed a stronger signal with the preblocked antibodies (lane 9) it would appear that there is no kinase activity associated with CRK4 precipitated from STIB 247 procyclics.

3.2.9: Northern blotting and RT-PCR

To corroborate the findings from Western blotting that *CRK4* is expressed mainly in procyclics and to obtain a rough quantitation of transcript presence between life cycle stages, poly[A]+ RNA was prepared from STIB 247 long slender bloodstream, short stumpy bloodstream and procyclic forms, and a Northern blot performed (see chapter 2.2.8). Following RNA transfer from the gel to supported nitrocellulose membrane by capillary blotting, the membrane was stained with methylene blue and probed with the insert from pGL138. No transcript was detected in any of the life cycle stages, even after 14 days exposure of the filter to autoradiography film. However, as no mRNA was observed on methylene blue staining, and a re-probe of the filter with the insert from pGL130 did not detect the CRK2 transcript, it is possible that the RNA was degraded.

To check the integrity of the RNA preps, first-strand cDNA was made from the poly[A]+ RNA preps for all three life cycle stages using an oligo[dT] primer. In addition, control reactions were performed, identical to the first-strand synthesis reactions with the exception that the control reactions did not include reverse transcriptase (for details see chapter 2.2.8). PCR was then carried out with the first-strand cDNA reactions, labelled (+), and the control reactions, labelled (-). Initially, the primer combination OL355 and OL356 was used, to amplify a 760 bp fragment; these reactions worked but the end-point products were of very low abundance, therefore visualisation was poor. Consequently a nested approach was taken, using the first round primer combination OL355 and OL90 to amplify a 945 bp fragment, with the second round combination of OL348 and OL356, to amplify a 437 bp fragment, or OL348 and OL350, to amplify a 144 bp fragment. Aliquots of the control (-) samples were treated with RNaseA, labelled (-/-) as an additional negative control. A set of reactions using the aforementioned primers gave odd results; the expected amplification fragments were seen for the (+) reactions, but the exact same reactions were also seen for the (-) reactions (data not shown). Only the (-/-) reactions showed no amplification products. As it is not possible to amplify by PCR from mRNA, no conclusions could be drawn from these experiments.

3.3: Discussion

The PCR-based approach to isolation of the *CRK4* gene was only partially successful. The gene specific primers amplified a product of the predicted molecular weight (Fig. 3.2); cloning and sequencing of this 760 bp fragment showed it to be *CRK4*, with no base substitutions as compared to the partial sequence of Dr. Murphy (unpublished). The utilisation of the mini-exon 39 bp splice leader sequence along with one of the gene specific primers

(OL355) gave a less specific reaction (Fig 3.3), with several species persisting to high annealing temperatures. The three major fragments of 0.5 kb, 0.85 kb and 1.0 kb were cloned. Initial sequencing of the cloned 1.0 kb fragment showed *CRK4* sequence; further sequencing showed it to be the expected amplification product representing the 5' end of the gene, and indicated the acceptor site for *trans*-splicing to the 39 bp splice leader sequence (Fig. 3.1). Attempts to amplify the 3' end of *CRK4* using primer OL355 and the complimentary anchor primer gave two amplification products that persisted to reasonable annealing temperatures, of 1.0 kb and 1.4 kb. Based on calculations of the minimum amplification product size required to encode all the remaining required kinase catalysis domains without a second insert domain, only the 1.4 kb fragment was cloned and sequenced. Preliminary sequencing showed this to be a fragment of the triosephosphate isomerase gene; the triosephosphate isomerase mRNA is an abundant transcript and therefore forms one of the most highly represented species in an expression library. As previously mentioned primer OL355 corresponds to the ATP-binding pocket of CRK4, covering the sequence that encodes the amino acid sequence LGEGTYG. This is a consensus sequence that is conserved across proteins that bind ATP, therefore OL355 provided less specificity for *CRK4* than was initially believed and partnered with the complementary anchor primer that annealed all first-strand cDNAs, the erroneous amplification of a fragment of the triosephosphate isomerase gene is perhaps not as surprising as initially thought.

The library screen was carried out to obtain the 3' end of *CRK4*, also allowing a direct comparison of genomic and transcript-derived sequences. In addition this enabled isolation and sequencing of the *CRK4* flanking sequences, required for future work to develop targeted gene disruption constructs. The insert from pGL138 was used to screen the λFIXII *T. b. brucei* genomic library; this fragment covers kinase domains I to IX, and the first insert domain. A hybridising temperature of 50°C was employed, lower than would normally be used. Given the existence of a subfamily of *CRK4* genes in *C. fasciculata* it was hypothesised that there may be a similar subfamily in *T. brucei*, hence a lower hybridisation temperature would increase the likelihood of detection of *CRK4* homologues. The primary screen yielded only three hybridising plaques; secondary and tertiary screens isolated these three clones. The initial restriction analysis (Fig 3.5) and Southern blotting (Fig 3.6) showed that clones IA and IIA were similar whereas clone IIIA was quite different. There were differences notable in the restriction patterns of clones IA and IIA with *Sal*I, and a slight difference in the size of the one noticeable fragment produced with *Not*I, therefore leading to the supposition that clones IA and IIA represented overlapping sequences. The Southern blot (Fig 3.6) showed exactly the same pattern of hybridising fragments for clones IA and IIA, showing that these clones likely represented phage inserts covering a *Sau*3AI site internal to *CRK4* to sequence outwith the gene. Clone IIIA gave a different pattern of hybridising fragments, although IA, IIA and IIIA all gave a small molecular weight hybridising fragment of about 0.4 kb with *Hinc*II. There were two *Hinc*II sites within the

known sequence of *CRK4*, at positions 11 and 413 (relative to the A of ATG start) which would give the 0.4 kb fragment observed for clones IA and IIA (Fig. 3.1). As clone IIIA gave a different pattern of restriction fragments and hybridising fragments the assumption was that there was a *Hinc*II site approximately 200 bp into the phage arm sequence that gave a 0.4 kb fragment representing 200 bp of *CRK4* sequence and 200 bp of phage sequence that was detected by the probe. At this point, due to the presence of two *Sau*3AI sites in the region of the gene covered by the pGL138 insert probe at positions 266 and 561 it was not possible to tell if clone IIIA insert sequence was contiguous with the insert sequence of clones IA and IIA.

From the Southern blots the 2.5 kb *Sal*I hybridising fragment was subcloned from clones IA and IIA, and the 1.1 kb *Not*I fragment was subcloned from clone IIIA. Restriction analysis and Southern blotting of the chosen subclones showed that three of the four subclones of IA and IIA were identical but did not hybridise the probe, with the one distinct subclone (IA/2, pGL119) producing a different restriction pattern and hybridising the probe. Both IIIA/1 and IIIA/2 hybridised the probe and were identical by restriction analysis. Preliminary sequencing of pGL119 showed antisense *CRK4* 5' sequence from the *Sau*3AI site at position 560 toward the 5' end of the gene; therefore the pGL119 insert represented the 5' end of the gene and approximately 2 kb of 5' flanking sequence. Preliminary sequencing of pGL140 showed antisense *CRK4* 3' sequence from kinase catalysis domain X. As this subclone represented little if any 3' flanking sequence, further Southern blots were done to identify larger hybridising fragments from clone IIIA; the 1.6 kb *Sac*I-*Kpn*I fragment was identified (Fig. 3.7) and subcloned (pGL218).

Complete sequencing of the pGL140 insert and partial sequencing of the pGL218 and the pGL119 inserts gave the complete double-strand sequence of *CRK4*, showing that the pGL140 insert and both the pGL218 and pGL119 inserts represented contiguous *CRK4* sequence with the *Sau*3AI site at position 560 the site dividing the two types of subclone. The sequence from pGL138 (the PCR product obtained from OL355 and OL356) overlapped this *Sau*3AI site, proving that the two types of lambda subclones represented contiguous sequence. The complete ORF for CRK4, translated into peptide sequence, showed the presence of a second Insert Domain at an equivalent position to the cfCRK Insert Domain 2, and of approximately the same size. Multiple alignment of tbCRK4with tbCRK1-3, cfCRK and CDC28 showed that tbCRK4 possessed all the kinase catalysis domains as well as the CDK/CRK-specific features - a PSTAIRE and DSEI box (although both to some degree degenerate) as well as the conserved serine/threonine and tyrosine residues at equivalent positions as compared to CDK1 (not shown) and CDC28 (although the serine and threonine residues appear to be fairly interchangeable at these conserved points between CRKs). BLAST database matching searches using just the insert domain sequences from tbCRK4 did not find any significant matches, and there was little homology between tbCRK4 and cfCRK for Insert Domain 1 or Insert Domain 2.

However, hydropathy plotting of the two CRK4 proteins showed that both insert domains were hydrophilic, with a striking similarity in Insert Domain 2 showing as a highly hydrophilic spike on the plot. It is tempting to speculate that the insert domains - despite a lack of primary sequence homology - may adopt similar secondary structures and have relevant and homologous biological function.

The Pharmacia Biotech GST system was used to try and express a fragment of *CRK4* to produce recombinant protein. This system involves a large (26 kDa) highly immunogenic tag which could prove problematic to remove, should that have been necessary, but which would help to both stabilise the CRK4 portion of the recombinant protein and solubilise it. Purification of glutathione-S-transferase tagged protein is carried out under native conditions, hence GST-CRK4 would have to be soluble within the bacteria for this system to be feasible.

The section of *CRK4* chosen to be cloned to the expression vectors was a *Hinc*II-*Hae*III fragment of 244 bp that represented almost exactly the first insert domain with virtually no kinase domain sequence (Fig. 3.1). The rationale was that as this insert was not present in the other known *CRK* genes nor did database matching produce any significant matches, using it to produce recombinant protein and raise antisera would increase the chance of obtaining an antiserum specific for CRK4. The vector pGEX-5X-2 was chosen to engineer a fusion protein with the CRK4 portion C-terminal to the GST protein (pGL134; Fig. 3.12).

Pilot inductions with the expression construct pGL134 in *E. coli* strain BL21 (Fig. 3.13) showed two induced proteins of 30 kDa and 44 kDa. Comparing soluble and insoluble protein fractions of induced cultures showed that both the 30 kDa and the 44 kDa proteins in the pilot expression of GST-CRK4 were located in the soluble protein fraction (Fig. 3.14). A small-scale native purification was performed for GST-CRK4; the 44 kDa protein was seen in the insoluble protein fraction but not in the soluble nonbound protein fraction, although the 30 kDa was observed in the latter and therefore did not bind the immobilised glutathione matrix. The 44 kDa protein was also present with very little apparent contaminating protein in the eluate fraction. The assumption was therefore that the 44 kDa protein was GST-CRK4, and that the presence of some of this in the insoluble protein fraction was likely due to large over-expression of the recombinant protein thus either allowing some to be brought down with the cellular debris or representing some recombinant protein forming inclusion bodies. The predicted size of GST-CRK4 was 36 kDa, 26 kDa from the GST protein and 10 kDa from the CRK4 segment, hence there was a disparity of about 8 kDa between the predicted size and the observed size of the induced, purified protein. An examination of the isoelectric point of the CRK4 segment predicted a *p*I of 4.4. Highly acidic proteins, the exemplar being histones, are known to migrate anomalously on SDS-PAGE gels therefore it was possible that the low isoelectric point of the CRK4 segment of GST-CRK4 was causing the recombinant protein to be retarded on Lammeli gels, hence the molecular weight disparity of 8 kDa.

The GST system was then used to express and purify large quantities of GST-CRK4, which was used in the immunisation of two rabbits, by the schedule described in chapter 2.3.4. Preliminary Western blotting using the first bleed antisera showed both to be detecting many proteins, although most of them could not be competed with fusion protein (data not shown). The multiplicity of proteins observed could have been the result of a number of factors: The antisera were nonpurified, and as such contained large quantities of IgM and other serum proteins. Also, the GST system uses the glutathione-S-transferase gene from *Schistosoma japonicum*; anti-GST antibodies in the resulting polyclonal antisera can cause problems in immunological analysis due to cross-reactivity to endogenous GST-like proteins in the protein samples under study. Accordingly, following the exsanguination bleed of both rabbits and preliminary Western blots to determine which of the two gave lower background (data not shown) the anti-CRK4/B antiserum was purified in a two-step process. Anti-GST antibodies were removed from an aliquot of the antiserum by specific adsorption to an immobilised GST column. The column flow-through was tested in a Western blot using lysates of STIB 247 procyclic, long slender bloodstream and short stumpy bloodstream forms, and lysates of the TREU 869 long slender bloodstream form (Fig. 3.16). Many of the proteins previously observed that were not competed with fusion protein were absent. Specific proteins that could be competed with purified GST-CRK4 fusion protein were detected at 46 kDa and 48 kDa in both the soluble and insoluble protein lysates for all three STIB 247 life cycle stages, although these signals were very weak with the exception of the 48 kDa protein in the STIB 247 short stumpy bloodstream and long slender bloodstream form soluble lysates. In addition, there was a larger molecular weight protein detected in both STIB 247 long slender bloodstream form soluble and insoluble lysates that were competed with fusion protein. At this juncture the 3' end of *CRK4* had not been isolated, so there was no definite predicted size for the protein. A rough calculation of the size that would be added to the known peptide sequence in order to represent the remaining kinase catalysis domains suggested that in the absence of a very large second insert domain, CRK4 was most likely either the 46 kDa or the 48 kDa protein.

The GST column flow-through was then affinity purified on an immobilised Sepharose-GST-CRK4 aminolink column, and the purified anti-CRK4 antibodies tested in a Western blot with fresh soluble lysates made of STIB 247 procyclic, short stumpy bloodstream and long slender bloodstream forms, to attempt to clarify exactly what the antiserum was detecting (Fig. 3.17). This gave much clearer results, with one protein at approximately 54 kDa detected in the procyclic lysate that was competed with purified GST-CRK4 fusion protein. This protein was also present in the short stumpy bloodstream form lysate at lower intensity, but was absent from the long slender bloodstream form lysate. There were in addition faint proteins just above and below this dominant protein, but modification of the blotting protocol lead detection of only the 54 kDa protein (Fig. 3.18). The size of this protein corresponds well to the predicted size from

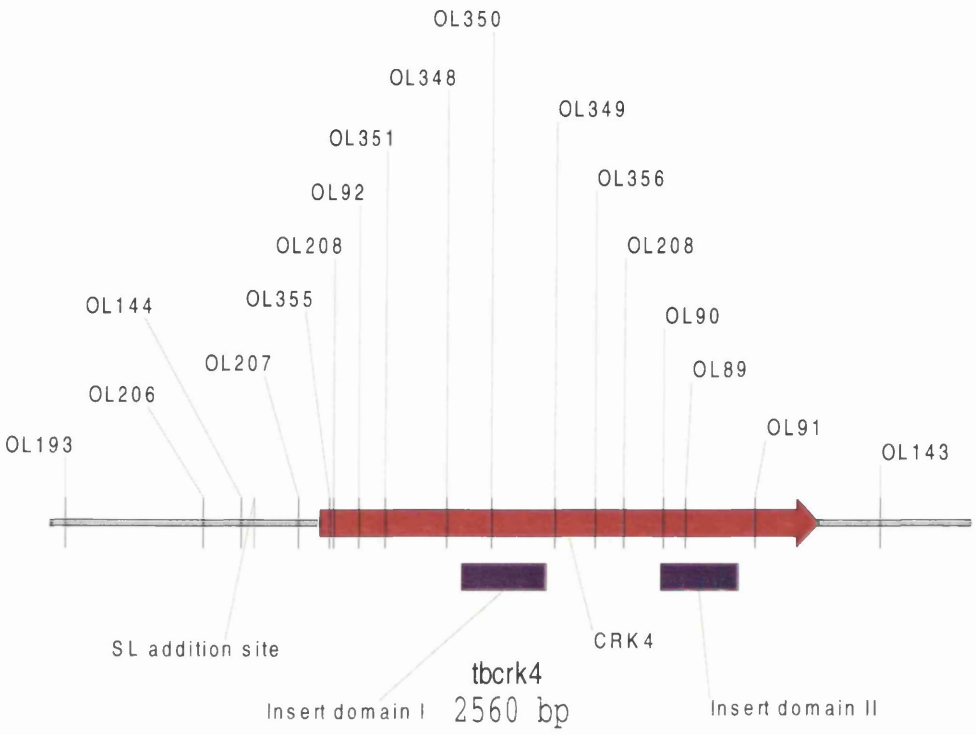
the complete open reading frame of the gene of approximately 51-52 kDa. The conclusion therefore is that *CRK4* is expressed to a low degree in the short stumpy bloodstream form and to a high degree in the procyclic form, but that it is not expressed in the long slender bloodstream form. Control of expression and activity could therefore be at the level of transcriptional control, mRNA translation (or mRNA stability) or post-translational modification.

The purified anti-CRK4/B antiserum was also used in immunoprecipitations linked to kinase assays, to attempt to demonstrate a functional kinase activity for CRK4 in procyclic form soluble extracts (Fig. 3.20). However, use of a range of substrates failed to show any significant activity that could be competed by fusion protein (Fig. 3.21). As the antiserum was raised to a soluble GST fusion protein containing a CRK4-specific section of 83 amino acids it would be expected that the CRK4 portion have secondary structure and possibly some tertiary structure. A hydropathy plot of the first insert domain showed it to be highly hydrophilic (Fig. 3.11) and therefore possibly forming a discrete domain structure exposed to the cytosol. The GST-CRK4 fusion could thus reasonably have been expected to have conformational epitopes, leading to a polyclonal antiserum that would perhaps be poorer at recognising denatured CRK4 in the context of a Western blot, but better at recognising native CRK4 in the context of an immunoprecipitation. The failure to demonstrate a CRK4-linked kinase activity from procyclic form extracts could have several explanations. Firstly, that although soluble the CRK4 portion of the GST-CRK4 fusion protein, outwith the context of the rest of CRK4, did not fold properly and that antibodies in the resulting antiserum were largely directed to linear peptide sequence and spurious conformational epitopes, giving specific detection on a Western blot but failing to recognise the native protein. Secondly, it is possible that the insert domains of CRK4 confer some novel functional interaction with other protein(s) or other cytosolic molecules that physically exclude the binding of antibodies to the domain, whether by stearically blocking antibodies or by causing conformational change of the insert domain. This possibility is intriguing, given previous speculation about the function of these domains, but beyond the fact that hydropathy plots show both to be highly hydrophilic, little could be extrapolated as database search matching found no homology to any other eukaryotic peptide sequences, nor any structural motif similarities. A third possibility is that CRK4 has been immunoprecipitated but that it is not active in the procyclic form of the parasite; this would appear to be counter-intuitive, as it is expressed at greatest abundance in this form, but it is possible that it is active at the transition between the short stumpy bloodstream form and the procyclic form, and is present though not active in cultured procyclics. Alternatively, it is possible that CRK4 activity is required in the proventricular mesocyclic or the epimastigote forms, and is held as an inactive complex until then. As it is difficult to predict protein substrates for kinases, the possibility that precipitated CRK4 from procyclic extracts is active but not against any of the substrates tested cannot be ruled out.

Selection with recombinant $p12^{CKS1}$ and $p13^{SUC1}$ from procyclic form soluble protein extracts were run on Lammeli gels, transblotted to PVDF membrane and probed with the purified anti-CRK4/B antiserum, to determine which of the recombinant proteins CRK4 bound. Previous work has shown that *T. b. brucei* CRK1, CRK2 and CRK3 all bind p12, and that CRK3 binds p13 weakly (Van Hellemond and Mottram, unpublished), therefore the same was expected to be observed for CRK4. However, repeated attempts at selection showed that CRK4 from procyclics does not bind either p12 or p13 (Fig. 3.19). There are several possible reasons for this. As speculated, CRK4 may be held inactive as part of a complex in the procyclic form, precluding binding to p12. Alternatively, CRK4 may simply not bind p12; the insert domains may confer novel biological function upon CRK4 such that it does not interact with p12 (or the as yet undiscovered *T. brucei* equivalent). Given the failure to demonstrate a kinase activity and the fact that CRK4 does not appear to bind p12, demonstration of binding a cyclin-like protein would seem to be necessary to classify this as a functional CDK homologue. The position of trypanosomes as one of the earliest divergents form the eukaryotic tree should be borne in mind, as CRK4 could represent divergence from a 'proto-CDK' to a CDK-like protein that has novel biological functions, possibly not related to control of the cell cycle.

To corroborate the Western blot results, that in STIB 247 CRK4 is not expressed in the long slender bloodstream form, expressed to low levels in the short stumpy bloodstream form and expressed to higher levels in the procyclic form, Northern blotting was performed on poly[A]+ RNA isolated from STIB 247 procyclic, long slender and short stumpy bloodstream forms. No transcript was detected when the membrane was probed with the insert from pGL138. As there was some doubt about transcript integrity, first-strand cDNA was synthesised; however, the results from this had to be discounted as the negative control gave positive results. Successful Northern blotting, or quantitative PCR, could delineate the question of differential transcription rates between the life cycle stages and whether control at the level of RNA stability or translation is a major point of control of expression. Control of CRK4 activity by post-translation modification has not been shown as functional kinase activity has not been demonstrated.

A


OL350
OL348
OL349
OL351
OL356
OL92
OL208
OL208
OL144
OL355
OL90
OL207
OL89
OL206
OL91
OL193
OL143
SL addition site
CRK4
tbcrk4
Insert domain I
2560 bp
Insert domain II

B

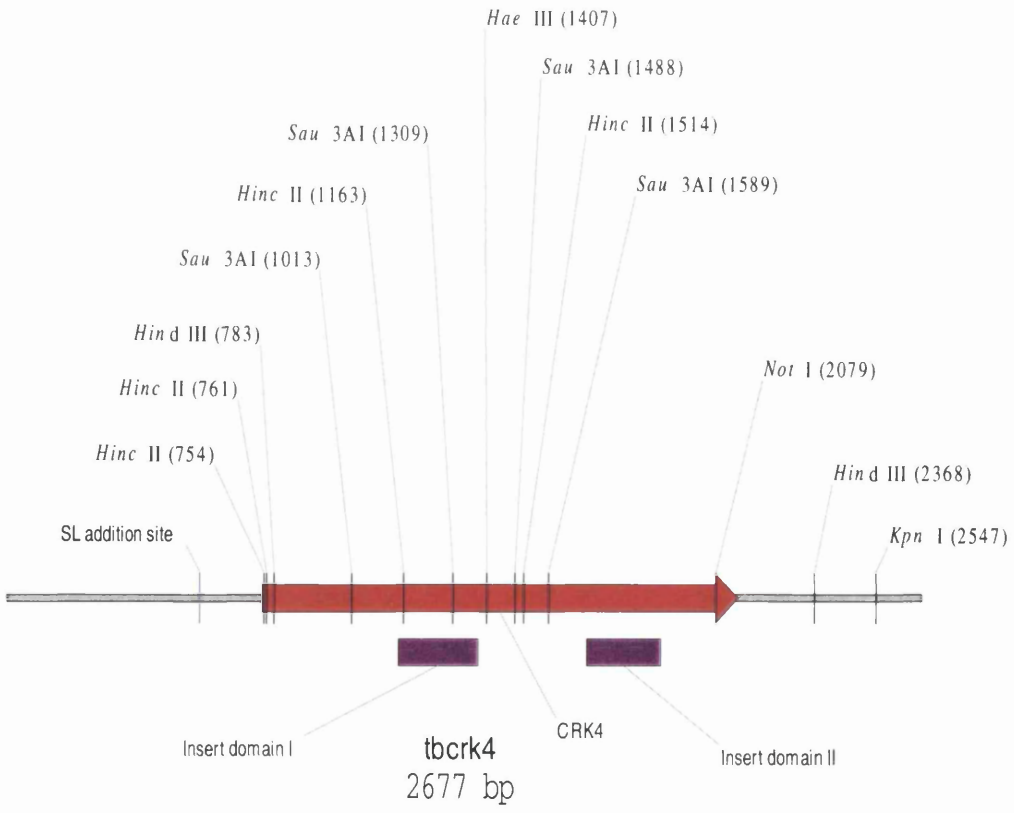
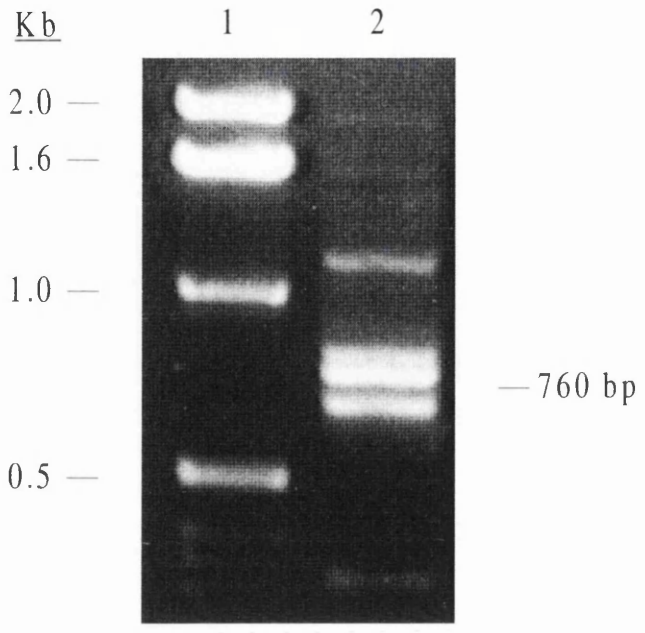

Hae III (1407)
Sau 3AI (1488)
Sau 3AI (1309)
Hinc II (1514)
Hinc II (1163)
Sau 3AI (1589)
Sau 3AI (1013)
Hind III (783)
Not I (2079)
Hinc II (761)
Hinc II (754)
Hind III (2368)
SL addition site
Kpn I (2547)
CRK4
Insert domain I
tbcrk4
Insert domain II
2677 bp

Fig. 3.1: The *T. brucei CRK4* gene

A) Schematic diagram of the *CRK4* gene showing the primers used to obtain full-strand sequence for the open reading frame (ORF) and the flanking sequences. The splice leader addition site is also shown (SL) and the two insert domains. B) Schematic diagram of the *CRK4* gene showing the relevant restriction sites. The *Sau*3AI site at position 1309 (560 relative to the A of ATG start) is the site relevant to the two contiguous lambda subclones. The *Hin*cII site at 1163 and the *Hae*III site at 1407 (414 and 658, respectively, relative to the A of ATG start) were the sites used for excision of Insert Domain 1 for the recombinant protein work.


Kb
1
2
2.0
1.6
1.0
0.5
760 bp

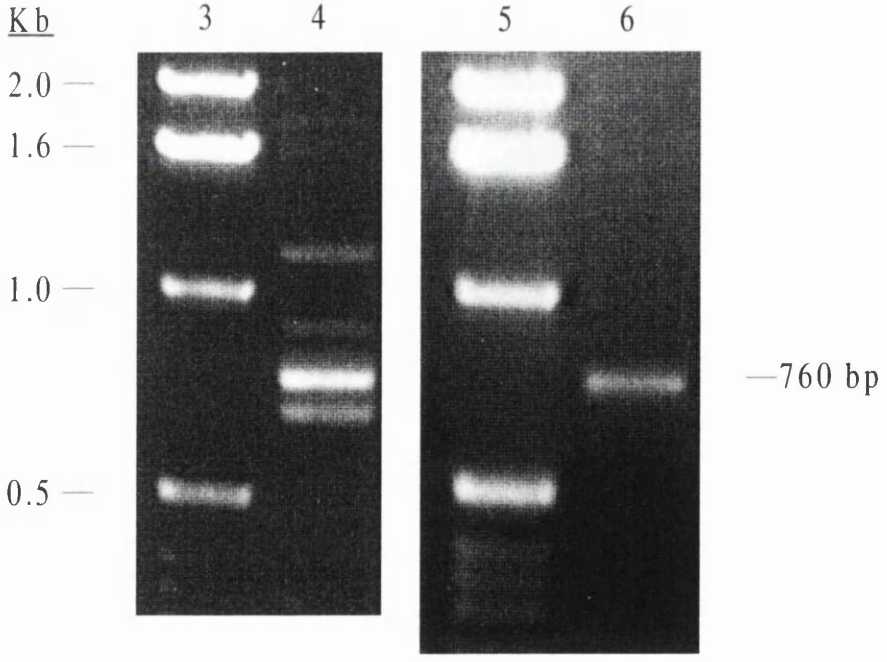

Kb
3
4
5
6
2.0
1.6
1.0
0.5
760 bp

Fig. 3.2: PCR from first-strand cDNA to amplify a *CRK4* gene fragment

PCR was performed on first-strand cDNA of *T. b. brucei* STIB 247, using 2 µl of cDNA in a reaction volume of 10 µl with 50 ng each of primers OL355 and OL356. Conditions were the same between reactions with only the annealing temperature altered; the end-point products were electrophoresed on 1.0% agarose gels and visualised by staining with 0.5 µg ml^{-1} ethidium bromide.

Lanes 1, 3, 5. 1 kb ladder markers (GibcoBRL)

Lane 2. 37°C annealing

Lane 4. 42°C annealing

Lane 6. 48°C annealing

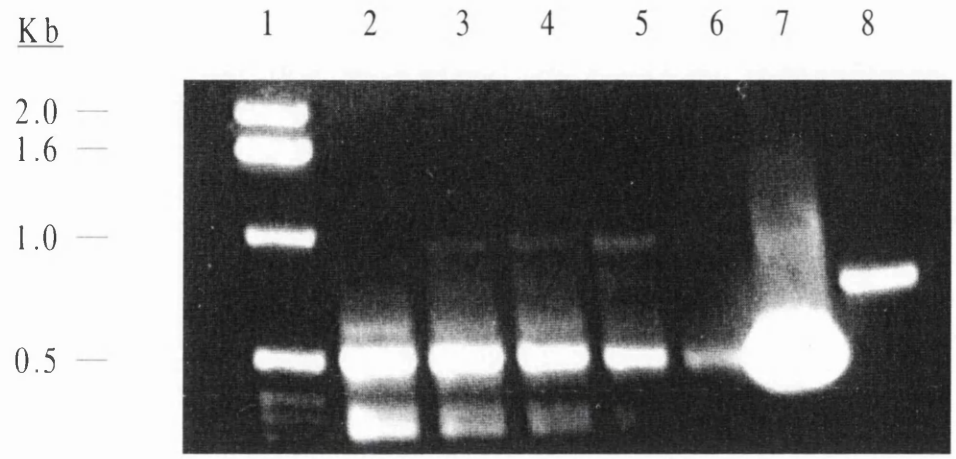


Fig. 3.3: PCR from first-strand cDNA to amplify the 5' end of *CRK4*

PCR was performed on first-strand cDNA of *T. b. brucei* STIB 247, using 2 µl of cDNA in a reaction volume of 10 µl with 50 ng each of primer OL356 and the splice leader (SL) primer. Identical reactions were run on a Robocycler machine (Stratagene) with an annealing temperature range of 42°C-62°C (12 reactions). The end-point products were electrophoresed on a 0.8% agarose gel and visualised by staining with 0.5 µg ml^{-1} of ethidium bromide. The range was run from left to right (lane 2 represents 56.4°C annealing, lane 6 represents 62.0°C annealing; the remaining reactions are not shown). Lane 7 is the 500 bp *SAT* gene amplified by SAT5' and SAT3' primers at 42°C; lane 8 is the 760 bp CRK4 product amplified with OL355 and OL356 primers at 48°C annealing. Markers (lane 1) were the 1 kb ladder markers (GibcoBRL).

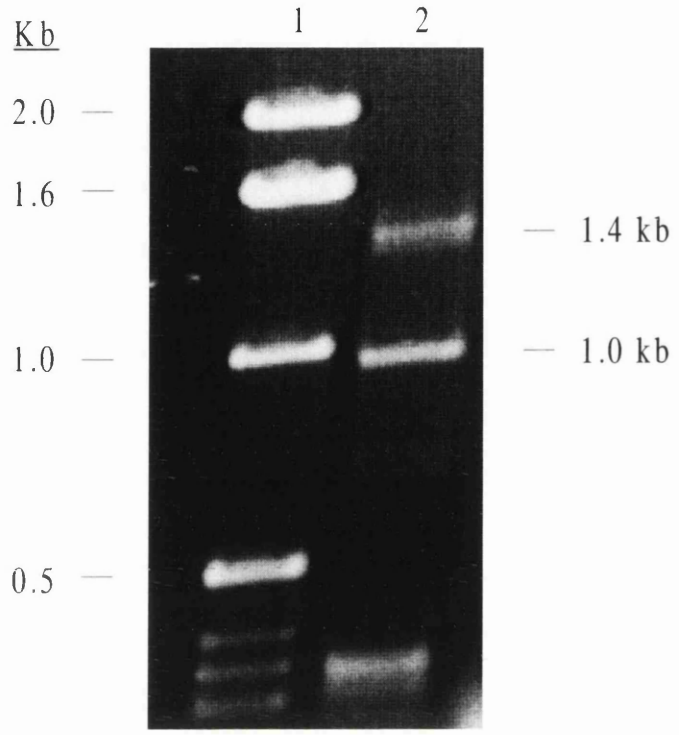


Fig. 3.4: PCR from first-strand cDNA to amplify the 3' end of *CRK4*

PCR was performed on first-strand cDNA of *T. b. brucei* STIB 247 that had been synthesised using an anchor-[dT] primer, using 2 µl of cDNA in a reaction volume of 10 µl with 50 ng each of primer OL355 and the complementary anchor sequence primer. The annealing temperature was 52°C. The amplified fragments of 1.0 kb and 1.4 kb are marked.

Lane 1. 1 kb ladder markers (GibcoBRL)

Lane 2. PCR reaction end-point products

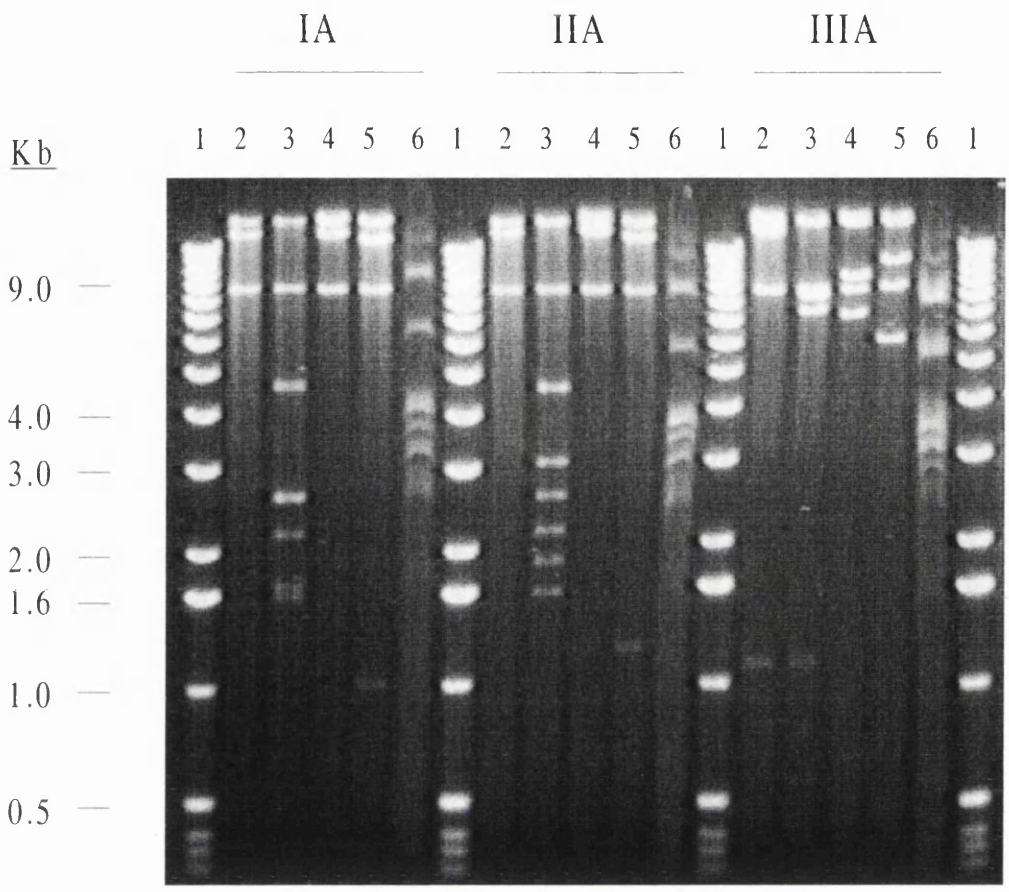


Fig. 3.5: Restriction digestion of purified λFIXII clones DNA

Approximately 1 µg of purified λFIXII clones IA, IIA and IIIA DNA was digested with restriction enzyme at 37°C for 2 hours and then electrophoresed on a 1.0% agarose gel. DNA was visualised by staining the gel with 0.5 µg ml^{-1} ethidium bromide.

Lanes 1. 1 kb ladder markers (GibcoBRL)

Lanes 2. *Not*I

Lanes 3. *Sal*I

Lanes 4. *Sac*I

Lanes 5. *Xba*I

Lanes 6. *Hinc*II

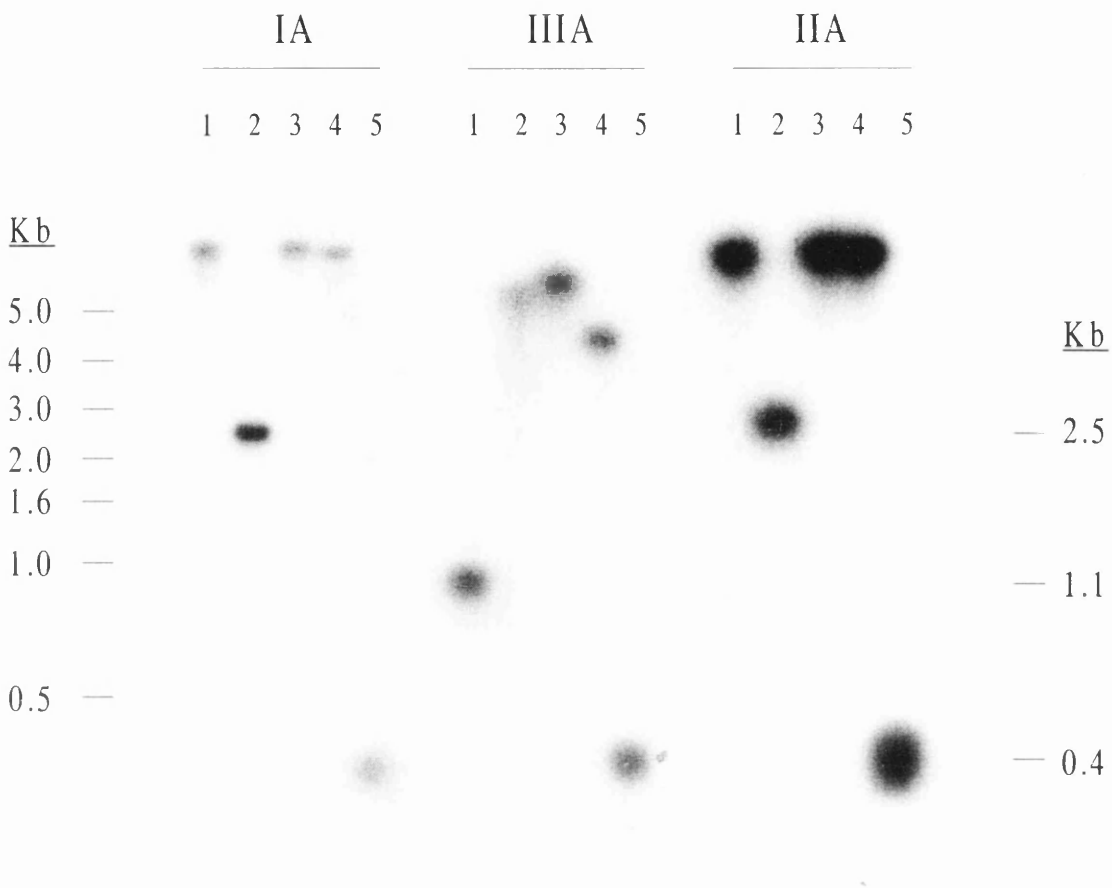


Fig. 3.6: Southern blot of λFIXII clones DNA

Approximately 1 μg of purified λFIXII clones IA, IIA and IIIA DNA was digested with restriction enzyme at 37°C for 2 hours and then electrophoresed on a 0.8% TBE agarose gel. DNA was transferred to a nitrocellulose membrane by the capillary method and fixed by UV illumination. The filter was prehybridised at 65°C for 4 hours and probed with the insert from pGL138. The sizes of the useful hybridising fragments are annotated. The markers were the 1 kb ladder (GibcoBRL).

Lanes 1. *Not*I

Lanes 2. *Sal*I

Lanes 3. *Sac*I

Lanes 4. *Xba*I

Lanes 5. *Hinc*II

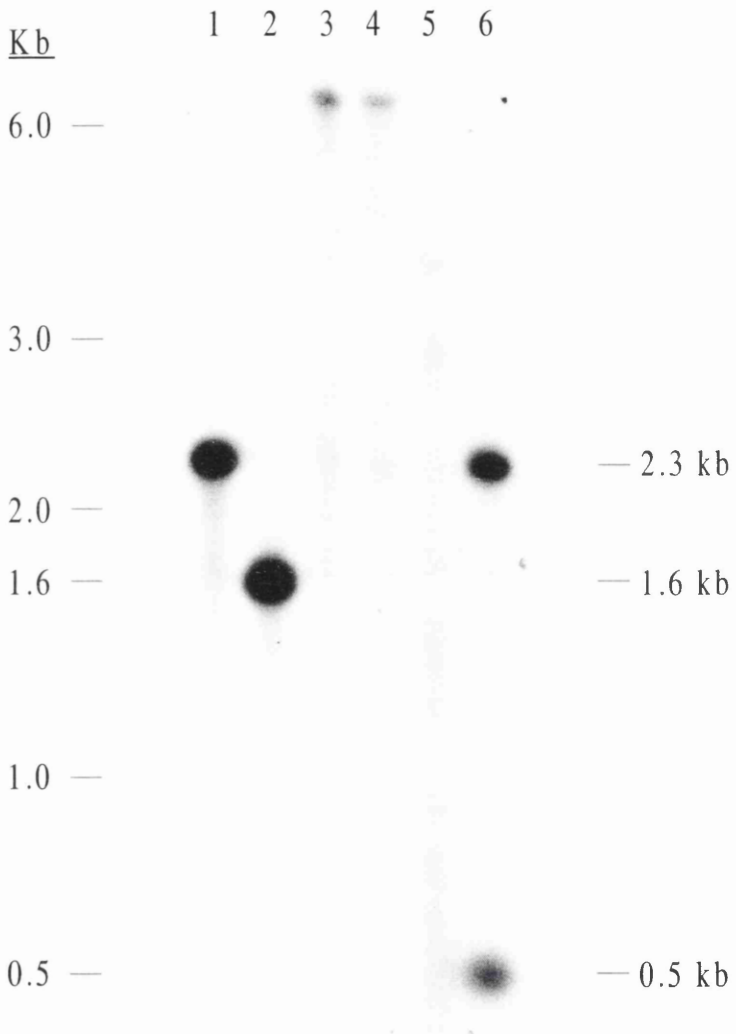

Kb
1 2 3 4 5 6
6.0
3.0
2.0
1.6
1.0
0.5
2.3 kb
1.6 kb
0.5 kb

Fig. 3.7: Southern blot of λFIXII clone IIIA

Purified λFIXII clone IIIA DNA (1 μg) was digested with restriction enzymes at 37°C for 2 hours and then electrophoresed on a 0.8% TBE agarose gel. DNA was transferred to a nitrocellulose membrane by the capillary method and fixed by UV illumination. The filter was prehybridised at 65°C for 4 hours and probed with the insert from pGL138. The sizes of the useful hybridising fragments are annotated. The markers were the 1 kb ladder (GibcoBRL).

Lane 1. *XbaI/XmnI*

Lane 2. *SacI/KpnI*

Lane 3. *XbaI/XhoI*

Lane 4. *XbaI/SpeI*

Lane 5. *PstI*

Lane 6. *NcoI*

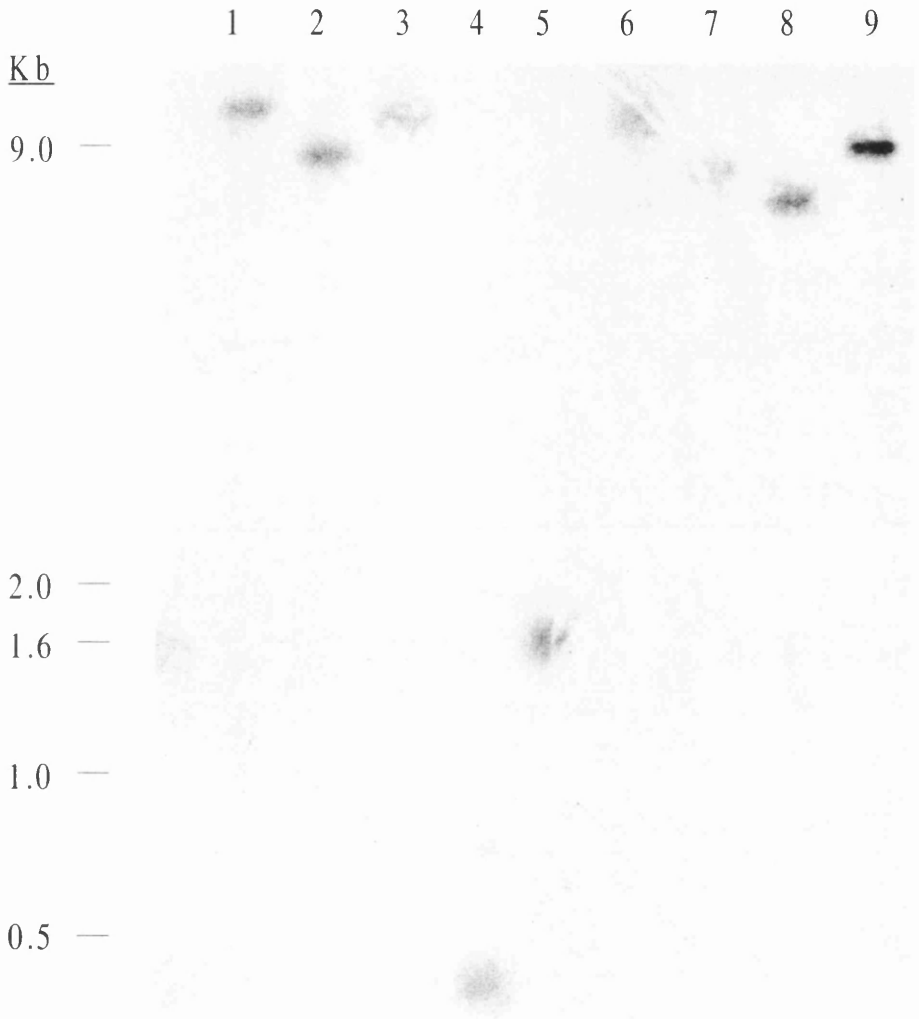

1 2 3 4 5 6 7 8 9
Kb
9.0
2.0
1.6
1.0
0.5

Fig. 3.8: Southern blot of STIB 247 genomic DNA with a *CRK4* PCR fragment

Genomic DNA was prepared from STIB 247 procyclics using the Nucleon kit. 2.5 μg of genomic DNA per lane was digested with enzymes for 4 hours at 37°C. Restriction digests were electrophoresed on a 0.8% TBE gel and blotted to Hybond-N supported nitrocellulose membrane (Amersham). The filter was fixed, prehybridised for 4 hours at 65°C, and probed with the 760 bp PCR fragment generated with OL355 and OL356 (insert from pGL138).

Lane 1. *Not*I

Lane 2. *Sal*I

Lane 3. *Xba*I

Lane 4. *Hinc*II

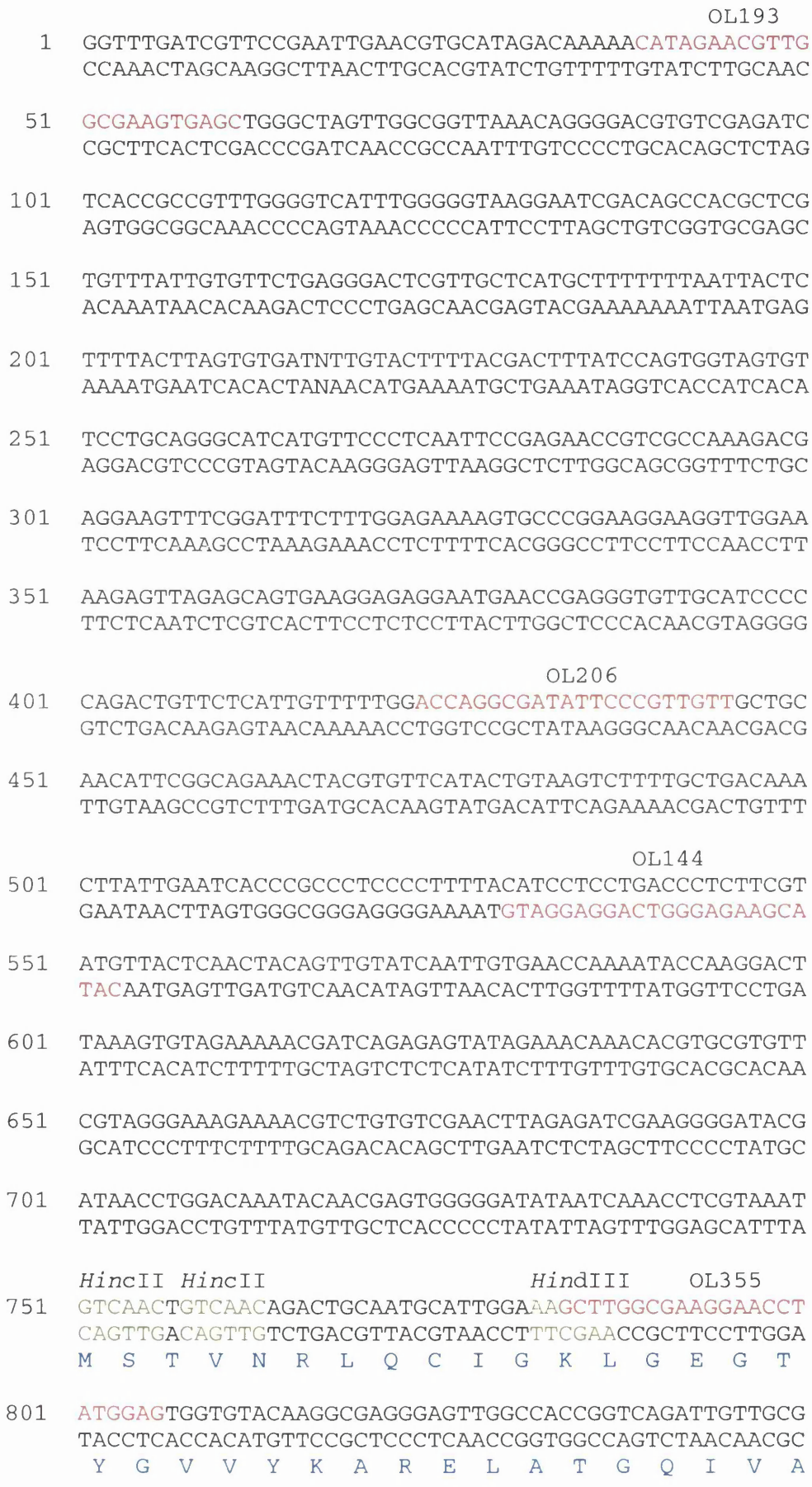
Lane 5. *Hin*dIII

Lane 6. *Bam*HI

Lane 7. *Sph*I

Lane 8. *Kpn*I

Lane 9. *Xho*I


OL193
1 GGTTTGATCGTTCCGAATTGAACGTGCATAGACAAAAACATAGAACGTTG
CCAAACTAGCAAGGCTTAACTTGCACGTATCTGTTTTTGTATCTTGCAAC
51 GCGAAGTGAGCTGGGCTAGTTGGCGGTTAAACAGGGGACGTGTCGAGATC
CGCTTCACTCGACCCGATCAACCGCCAATTTGTCCCCTGCACAGCTCTAG
101 TCACCGCCGTTTGGGGTCATTTGGGGGTAAGGAATCGACAGCCACGCTCG
AGTGGCGGCAAACCCCAGTAAACCCCCATTCCTTAGCTGTCGGTGCGAGC
151 TGTTTATTGTGTTCTGAGGGACTCGTTGCTCATGCTTTTTTTAATTACTC
ACAAATAACACAAGACTCCCTGAGCAACGAGTACGAAAAAAATTAATGAG
201 TTTTACTTAGTGTGATNTTGTACTTTTACGACTTTATCCAGTGGTAGTGT
AAAATGAATCACACTANAACATGAAAATGCTGAAATAGGTCACCATCACA
251 TCCTGCAGGGCATCATGTTCCCTCAATTCCGAGAACCGTCGCCAAAGACG
AGGACGTCCCGTAGTACAAGGGAGTTAAGGCTCTTGGCAGCGGTTTCTGC
301 AGGAAGTTTCGGATTTCTTTGGAGAAAAGTGCCCGGAAGGAAGGTTGGAA
TCCTTCAAAGCCTAAAGAAACCTCTTTTCACGGGCCTTCCTTCCAACCTT
351 AAGAGTTAGAGCAGTGAAGGAGAGGAATGAACCGAGGGTGTTGCATCCCC
TTCTCAATCTCGTCACTTCCTCTCCTTACTTGGCTCCCACAACGTAGGGG
OL206
401 CAGACTGTTCTCATTGTTTTTGGACCAGGCGATATTCCCGTTGTTGCTGC
GTCTGACAAGAGTAACAAAAACCTGGTCCGCTATAAGGGCAACAACGACG
451 AACATTCGGCAGAAACTACGTGTTCATACTGTAAGTCTTTTGCTGACAAA
TTGTAAGCCGTCTTTGATGCACAAGTATGACATTCAGAAAACGACTGTTT
OL144
501 CTTATTGAATCACCCGCCCTCCCCTTTTACATCCTCCTGACCCTCTTCGT
GAATAACTTAGTGGGCGGGAGGGGAAAATGTAGGAGGACTGGGAGAAGCA
551 ATGTTACTCAACTACAGTTGTATCAATTGTGAACCAAAATACCAAGGACT
TACAATGAGTTGATGTCAACATAGTTAACACTTGGTTTTATGGTTCCTGA
601 TAAAGTGTAGAAAAACGATCAGAGAGTATAGAAACAAACACGTGCGTGTT
ATTTCACATCTTTTTGCTAGTCTCTCATATCTTTGTTTGTGCACGCACAA
651 CGTAGGGAAAGAAAACGTCTGTGTCGAACTTAGAGATCGAAGGGGATACG
GCATCCCTTTCTTTTGCAGACACAGCTTGAATCTCTAGCTTCCCCTATGC
701 ATAACCTGGACAAATACAACGAGTGGGGGATATAATCAAACCTCGTAAAT
TATTGGACCTGTTTATGTTGCTCACCCCCTATATTAGTTTGGAGCATTTA
HincII HincII HindIII OL355
751 GTCAACTGTCAACAGACTGCAATGCATTGGAAAGCTTGGCGAAGGAACCT
CAGTTGACAGTTGTCTGACGTTACGTAACCTTTCGAACCGCTTCCTTGGA
M S T V N R L Q C I G K L G E G T
801 ATGGAGTGGTGTACAAGGCGAGGGAGTTGGCCACCGGTCAGATTGTTGCG
TACCTCACCACATGTTCCGCTCCCTCAACCGGTGGCCAGTCTAACAACGC
Y G V V Y K A R E L A T G Q I V A


OL92
851 TATAAACGCATTATCGTCGGCAGTGAGGATGATGGTGCACCATCGACCGC
ATATTTGCGTAATAGCAGCCGTCACTCCTACTACCACGTGGTAGCTGGCG
Y K R I I V G S E D D G A P S T
OL351
901 CATACGGGAGATTGCGCTTCTTAAGGTTCTAAAGCATAATAATATCGTGC
GTATGCCCTCTAACGCGAAGAATTCCAAGATTTCGTATTATTATAGCACG
A I R E I A L L K V L K H N N I V
951 GACTCTATGACGTGCTTTTCGAGCCGCCAAAACTTACTCTTATTTTTGAG
CTGAGATACTGCACGAAAAGCTCGGCGGTTTTGAATGAGAATAAAAACTC
R L Y D V L F E P P K L T L I F E
Sau3AI
1001 TACTGCGAATATGATCTTAGGAGATTCATGCACAAACACGCCTCCAGAGT
ATGACGCTTATACTAGAATCCTCTAAGTACGTGTTTGTGCGGAGGTCTCA
Y C E Y D L R R F M H K H A S R
1051 GCGCGCACAGGTTAAGGAAATTCTCAAACAAGTCCTCCTGGGCCTTCGGT
CGCGCGTGTCCAATTCCTTTAAGAGTTTGTTCAGGAGGACCCGGAAGCCA
V R A Q V K E I L K Q V L L G L R
OL348
1101 ACATGCATCAGCGGAGTGTCGTGCATCGAGACTTGAAACCAGACAACATT
TGTACGTAGTCGCCTCACAGCACGTAGCTCTGAACTTTGGTCTGTTGTAA
Y M H Q R S V V H R D L K P D N I
HincII
1151 TTTGTTAGGGTTGACGGAATATCAGGGTCTGAGGGCAGTGAAATGTCAGT
AAACAATCCCAACTGCCTTATAGTCCCAGACTCCCGTCACTTTACAGTCA
F V R V D G I S G S E G S E M S
OL350
1201 CTCTAGTAGGGAAACGGAAGTAGTTAATACGCGAGGGCACTGGGACGACA
GAGATCATCCCTTTGCCTTCATCAATTATGCGCTCCCGTGACCCTGCTGT
V S S R E T E V V N T R G H W D D
1251 GTGATGGGGCAGCCAACTGCCCCAAGAATGCGAACCCCACCGAAAACCCT
CACTACCCCGTCGGTTGACGGGGTTCTTACGCTTGGGGTGGCTTTTGGGA
S D G A A N C P K N A N P T E N P
Sau3AI
1301 AACGCCAGGATCGCTTCGGAGGACCATGGAGCCGATGGTACTGTCCCCTC
TTGCGGTCCTAGCGAAGCCTCCTGGTACCTCGGCTACCATGACAGGGGAG
N A R I A S E D H G A D G T V P
1351 TGGCTCAACCACTCCGACTTCACAGTTGGTAATTAAGTTGGGTGATTATG
ACCGAGTTGGTGAGGCTGAAGTGTCAACCATTAATTCAACCCACTAATAC
S G S T T P T S Q L V I K L G D Y
HaeIII OL349
1401 GGATGGCCCGTATAGAGAGTATCCCAGTAAAAAAGTACTCTCACGACGTG
CCTACCGGGCATATCTCTCATAGGGTCATTTTTTCATGAGAGTGCTGCAC
G M A R I E S I P V K K Y S H D V

*Sau*3AI

1451 GCATCCCTCTGGTATCGTAGCCCTGATGCGCTGCTGGGATCTGCAATGTA
CGTAGGGAGACCATAGCATCGGGACTACGCGACGACCCTAGACGTTACAT
A S L W Y R S P D A L L G S A M

*Hinc*II OL356

1501 CGGTTTCGCCGTTGACCTGTGGTCGGTTGGCTGTATCTTCGCAGAAATGG
GCCAAAGCGGCAACTGGACACCAGCCAACCGACATAGAAGCGTCTTTACC
Y G F A V D L W S V G C I F A E M

*Sau*3AI

1551 TTACGGGCGCCCCATTAATACGCGGAAATACCGATGTAGATCAGTTGCTC
AATGCCCGCGGGGTAATTATGCGCCTTTATGGCTACATCTAGTCAACGAG
V T G A P L I R G N T D V D Q L L

1601 AAGACGTTCAAGCTGCTTGGCACGCCCACACCTGAGACGTGGCCATCTAT
TTCTGCAAGTTCGACGAACCGTGCGGGTGTGGACTCTGCACCGGTAGATA
K T F K L L G T P T P E T W P S

1651 GAAAAATTGCCCCAAGGCCGTACAACTTCTCAAGGCTGCAGCCGAATTGG
CTTTTTAACGGGGTTCCGGCATGTTGAAGAGTTCCGACGTCGGCTTAACC
M K N C P K A V Q L L K A A A E L

OL90

1701 CCAGGGGGGAACCTGCAGAAAATGGAAAAAAAACAAAGAATTCACGGGGT
GGTCCCCCCTTGGACGTCTTTTACCTTTTTTTTGTTTCTTAAGTGCCCCA
A R G E P A E N G K K T K N S R G

OL89

1751 GTCCAAAAACAGCACGAGAGGAATCTCCCTAATGTGCAAAAAACTGCAAC
CAGGTTTTTGTCGTGCTCTCCTTAGAGGGATTACACGTTTTTTGACGTTG
V Q K Q H E R L N P N V Q K T A

1801 AGGCACCTCACACCTTTCCGCCGTAATGCGTGTAGGTATTGATGGAGAGA
TCCGTGGAGTGTGGAAAGGCGGCATTACGCACATCCATAACTACCTCTCT
T G T S H L S A V M R V G I D G E

1851 GTGGTTGTGTTGCGGACCCCAGGGACGAGCTGTATCATTTCCCGCCGGAG
CACCAACACAACGCCTGGGGTCCCTGCTCGACATAGTAAAGGGCGGCCTC
S G C V A D P R D E L Y H F P P E

1901 TTGTGTTTTCCATCCGCATTCGATGAGTATGTGAAAGCATCCGGCTTCCA
AACACAAAAGGTAGGCGTAAGCTACTCATACACTTTCGTAGGCCGAAGGT
L C F P S A F D E Y V K A S G F

OL91

1951 GGGACGCGTGGGTGTTGAAGGTACTGATTTTCTACGGAGGCTGATCCGTT
CCCTGCGCACCCACAACTTCCATGACTAAAAGATGCCTCCGACTAGGCAA
Q G R V G V E G T D F L R R L I R

2001 ATGAGCCGTCACAGAGGATGTCGGTTCAGGAGGCGCTGTCCCATCCTTTT
TACTCGGCAGTGTCTCCTACAGCCAAGTCCTCCGCGACAGGGTAGGAAAA
Y E P S Q R M S V Q E A L S H P F

*Not*I

2051 GTCAGTAACACAGTGGCACTAATGCAGCGGCCGCTTGATGTCATGGCTTC
CAGTCATTGTGTCACCGTGATTACGTCGCCGGCGAACTACAGTACCGAAG
V S N T V A L M Q R P L D V M A

```
2101 CATGTTACGCCAATCACTTCGAGAACACAGTGTATATCCGTAATATATAT
     GTACAATGCGGTTAGTGAAGCTCTTGTGTCACATATAGGCATTATATATA
     S  M  L  R  Q  S  L  R  E  H  S  V  Y  P  *

2151 ATATATATATATATATATCTACGATTATTTGTTTTGGAAGTGCAGGGGTG
     TATATATATATATATATAGATGCTAATAAACAAAACCTTCACGTCCCCAC

2201 TGGCTTCCTCATTATTGGAAAATAGGCAGTTCTTACAATGTAGTTGTTAC
     ACCGAAGGAGTAATAACCTTTTATCCGTCAAGAATGTTACATCAACAATG

2251 ACTTACAGGGATGTCTTCAACATAGTATAATTACGGATGATCCACAGCGG
     TGAATGTCCCTACAGAAGTTGTATCATATTAATGCCTACTAGGTGTCGCC

                          OL143
2301 TGGTCCCGGCAGCATTCATCGCATTTAGCGTTCTCTATCTTGCTACTTTC
     ACCAGGGCCGTCGTAAGTAGCGTAAATCGCAAGAGATAGAACGATGAAAG

                       HindIII
2351 CTTCTTTCAAGTGCATAAGCTTTCTTTACTCCCTTCTATATTTTCAAAGC
     GAAGAAAGTTCACGTATTCGAAAGAAATGAGGGAAGATATAAAAGTTTCG

2401 AGACAGAACGTCACCAATCAAATATCGTAGTCCCCCCCTCTATTGTTTTT
     TCTGTCTTGCAGTGGTTAGTTTATAGCATCAGGGGGGGAGATAACAAAAA

2451 GAATCTTCTCACTACTGTTTTTATTAGCGTTTTCCTCCATGACAGTTTCT
     CTTAGAAGAGTGATGACAAAAATAATCGCAAAAGGAGGTACTGTCAAAGA

                                                    KpnI
2501 TTCGCATCAATTGGAGTTACGTACCCCTGCCGTTGGATGTTGGTACCCAA
     AAGCGTAGTTAACCTCAATGCATGGGGACGGCAACCTACAACCATGGGTT

2551 TTCSCCCTATAGTGAGTCGTATTACGCGCGCTCACTGGCCGTCGTTTTAC
     AAGSGGGATATCACTCAGCATAATGCGCGCGAGTGACCGGCAGCAAAATG

2601 AACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCA
     TTGCAGCACTGACCCTTTTGGGACCGCAATGGGTTGAATTAGCGGAACGT

2651 GCACATCCCCCTTTCGCCAGCTGGCGT
     CGTGTAGGGGGAAAGCGGTCGACCGCA
```

Fig. 3.9: The *CRK4* gene and flanking sequences

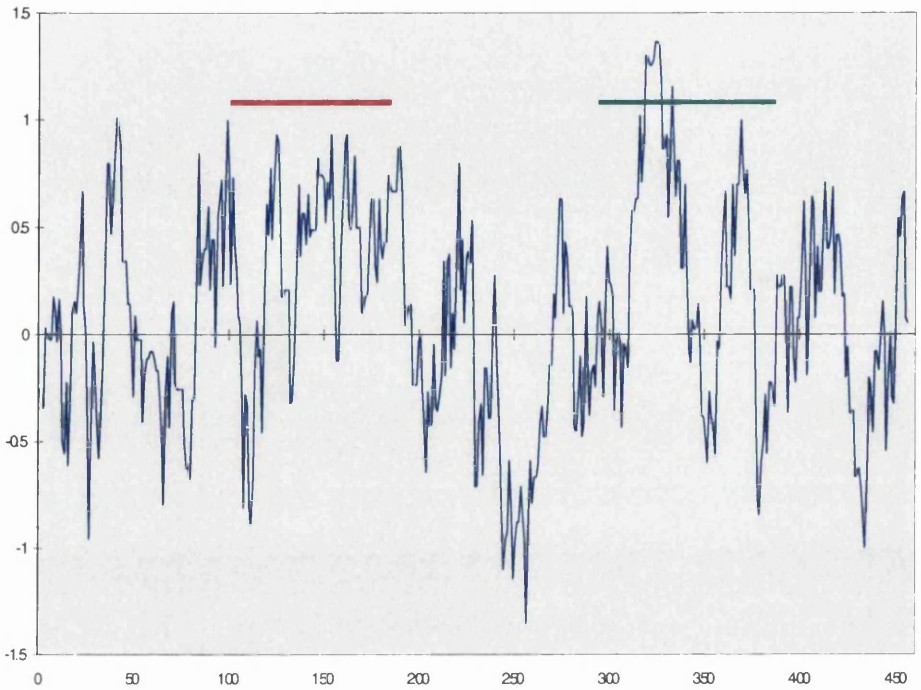
The complete ORF and flanking sequences with relevant restriction sites, primers and translation of the ORF is shown. Primer sequences are highlighted in red, on the upper strand (sense primers) and the lower strand (antisense primers). Only the restriction sites relevant to the formation of the lambda library, Southern blot data and subcloning of gene fragments have been shown. The protein translation (blue) shows the presence of the two insert domains (purple).

```
                                                                      I
TBCRK4          ---------------------------------------MSTVNRLQCIGKLGEGTYGVV 21
CFCRK4          ---------------------------------------MSTLGRYRHVVKLGEGTYGMV 21
TBCRK1          -------------------------------------------MGSRYERLQKIGEGSYGVV 19
TBCRK2          MQVQVQEGQTACDGSLRPLPSAGPASFVPRSLRPAPLRGTSTPDRYSRIEKVGEGSYGIV 60
TBCRK3          ----------------------MTMLGALTGRQLSSGLKDQFDRYNRMDILGEGTYGVV 37
CDC28           ---------------------------------------MSGELANYKRLEKVGEGTYGVV 22
                                                         .   :  :***:**:*

                          II                 III          IV
TBCRK4          YKARELATG---QIVAYKRIIVGSEDDGAPSTAIREIALLKVLKHNNIVRLYDVLFEPP- 77
CFCRK4          YKGTEIQTG---RVVAFKRMVVTSDDEGIPGAAIREICLLKELRHNNVVELFEVLFDPP- 77
TBCRK1          FRARDVTTG---TIVAVKRIRLEKEEEGVPCTAIREISILKELRHENIVRLLDVCHSEK- 75
TBCRK2          YKCHDNFTG---RTVAMKRIPLIVNDGGVPSTAVREVSLLRELNHPYVVRLLDVVLHEA- 116
TBCRK3          YRAVDRATG---QIVALKKVRLDRTDEGIPQTALREVSILQEIHHPNIVNLLDVICADG- 93
CDC28           YKALDLRPGQGQRVVALKKIRLESEDEGVPSTAIREISLLKELKDDNIVRLYDIVHSDAH 82
                ::  :  .*     ** *:: :   : * * :*:**:.:*: :..  :*.* ::

                       V                              VIa            VIb
TBCRK4          KLTLIFEYCEYDLRRFMHKHASRVR---AQVKEILKQVLLGLRYMHQRSVVHRDLKPDNI 134
CFCRK4          KITMIFELCDCDLKRYMESRPQRLLDANTEMRPILKQIFLGLEYLHGRCVVHRDMKPQNI 137
TBCRK1          RLTLVFECMEMDLKKYMDHVGGDLDAG--TIQEFMRSLLSGVRFCHERNVLHRDLKPPNL 133
TBCRK2          KLLLIFEYMEQDLQGMLK--QRNTAFVGGKLRRIMFQLLLGLHECHSRRFVHRDIKPSNI 174
TBCRK3          KLYLIFEYVDHDLKKALE--KRGGAFTGTTLKKIIYQLLEGLSFCHRHRIVHRDLKPANI 151
CDC28           KLYLVFEFLDLDLKRYMEGIPKDQPLGADIVKKFMMQLCKGIAYCHSHRILHRDLKPQNL 142
                :: ::**  : **:  :.             :: :: .:  *:   * : .:***:** *:

TBCRK4          FVRVDGISGSFGSEMSVSSRETEVVNTRGHWDDSDGAANCPKNANPTENPNARIASEDHG 194
CFCRK4          FVNVRGPDFAAMTALPSSPQQSMRVPHAGGTNGEAGRA---S-ANGNEHAPRPTAAEGSV 193
TBCRK1          LISREK---------------------------------------------------- 140
TBCRK2          LIDRKES--------------------------------------------------- 181
TBCRK3          LVTTDN---------------------------------------------------- 157
CDC28           LINKDG---------------------------------------------------- 148
                ::

                                VII                                VIII
TBCRK4          ADGTVPSGSTTPTSQLVIKLGDYGMARIESIPVKKYSHDVASLWYRSPDALLGSAMYGFA 254
CFCRK4          SPWEEAANTKDAPNQLIIKIGDFGLARVEEIPVKKYSHEVVTLWYRSPDVLMSSALYSYP 253
TBCRK1          ----------------ELKLADFGLGRAFGIPVKKFTQEVVTLWYRSPDVLLGSTQYGTP 183
TBCRK2          ----------------VVKLADFGLGRAFRVPLQTYTTEVMTLWYRAPEVLLGDKQYLPA 225
TBCRK3          ----------------SVKIADFGLARAFQIPMHTYTHEVVTLWYRAPEILLGEKHYTPA 201
CDC28           ----------------NLKLGDFGLARAFGVPLRAYTHEIVTLWYRAPEVLLGGKQYSTG 192
                                :*:.*:*:.*   :*:: :: :: :****:*: *:.   *

                     IX                                X
TBCRK4          VDLWSVGCIFAEMVTGAPLIRGNTDVDQLLKTFKLLGTPTPET--WPSMKNCPKAVQLLK 312
CFCRK4          VDIWSMGAIFFEMATSKVLFSGRNEDEQLLRMFWLLGSPTKET--WPSMMTYTGTMERLE 311
TBCRK1          VDIWSVGCIFAEMAIGAPLFTGKNDADQLLRIFQFLGTPNRQV--WPSMDTYPNSSNMLS 241
TBCRK2          VDVWSMGCVFAELARRRSLFAGDTAINQLFSIFQLLGTPTEAT--WRGVTSLPHHNVNFP 283
TBCRK3          VDMWSIGCIFAELARGKVLFRGDSEIGQLFEIFQVLGTPMDAEGSWLGVSSLPDYRDVFP 261
CDC28           VDTWSIGCIFAEMCNRKPIFSGDSEIDQIFKIFRVLGTPNEAI--WPDIVYLPDFKPSFP 250
                ** **:*.:* *:     :: * .   *::  * .**:*      * .:   .     :

TBCRK4          --AAAELARGEPAENGKKTKNSRGVQKQHER---------NLPNVQKTATGTSHLSAVMR 361
CFCRK4          RSSRAAAERQDLTVNGDVYVQQQQLQAQQQQPQQGSSPSHSSSRAPDLLTQLAHKRFYHS 371
TBCRK1          ---------------------------------------------------------- 243
TBCRK2          ---------------------------------------------------------- 284
TBCRK3          ---------------------------------------------------------- 262
CDC28           ---------------------------------------------------------- 251

                                                                  XI
TBCRK4          VGIDGESGCVADPRDELYHFPPELCFP-SAFDEYVKASGFQGRVGVEGTDFLRRLIRYEP 420
CFCRK4          ESAMQQRRESASSAANSYRLPVELWFDRPLFKEYMAATRCDVSVSPEGIDLLRRCLMYEP 431
TBCRK1          -------------------RPEFQQTLAATCEEQFQTNPAYAKLGPQGIDLLRWLLRYEP 282
TBCRK2          ------------------------------RWTAKPLRTAVPALDDDGVDLLRRMLCYNP 313
TBCRK3          ------------------------------KWSGKPLTQVLPTLDGDAVDLLSQMLRYNP 291
CDC28           ------------------------------QWRRKDLSQVVPSLDPRGIDLLDKLLAYDP 280
                                              .          :.  . *:*   : *:*

TBCRK4          SQRMSVQEALSHPFVSNTVALMQRPLDVMASMLRQSLREHSVYP 464
CFCRK4          NQRITAAEAVHHPYLERVPVPTAGSLDVLISSLMQTMETIHLL- 474
TBCRK1          SERLTAAQALEHPYFSVEF------------------------- 301
TBCRK2          RERITAYEALQHSYFDEVREEEVEKLMRFNGA------------ 345
TBCRK3          AERISAKAALQHPWFSDAMF------------------------ 311
CDC28           INRISARRAAIHPYFQES------------------------- 298
                 :*::.  *  *.:..
```

Fig. 3.10: Multiple alignment of CRK4 with CRKs

Multiple alignment of the sequences shown was done with the ClustalW application at the EBI website. As there were several spurious alignments of several amino acids for tbCRK1-3 and CDC28 within the insert domains of tbCRK4 and cfCRK, these have been manually corrected. Two underlying dots represents similar residues (conservative substitutions); an underlying star represents invariantly conserved residues. Dashes indicate spaces introduced to maximise alignment. Residues highlighted in red correspond to the PSTAIRE sequence motif of hsCDK1; residues highlighted in blue correspond to the DSEI sequence motif of hsCDK1; residues highlighted in green represent residues in conserved positions relative to residues in hsCDK1 that are known to be phosphorylated.

A


1.5
1
0.5
0
-0.5
-1
-1.5
0
50
100
150
200
250
300
350
400
450

B

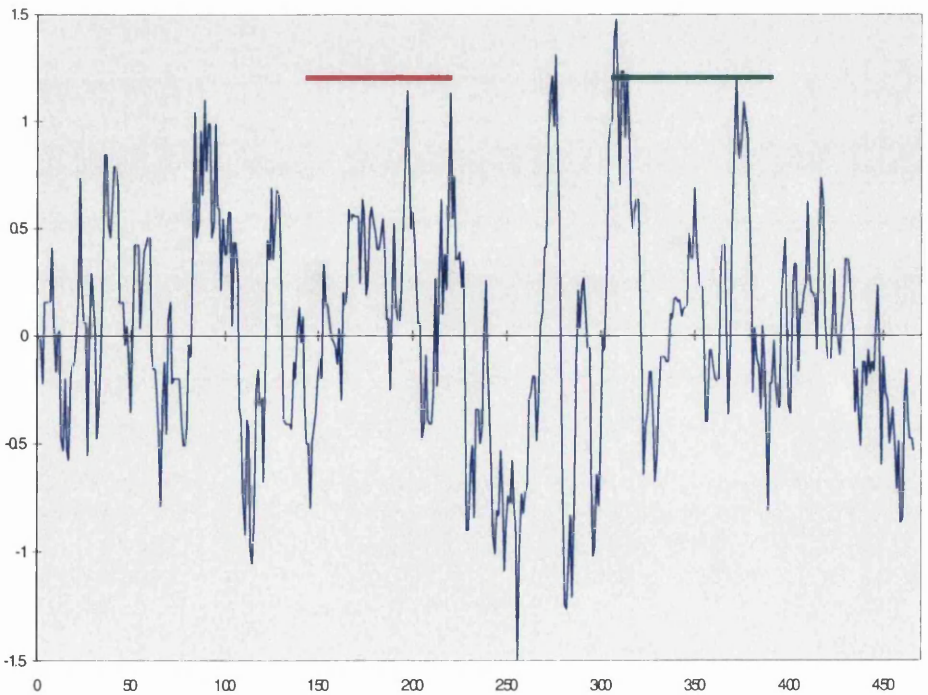

1.5
1
0.5
0
-0.5
-1
-1.5
0
50
100
150
200
250
300
350
400
450

Fig. 3.11: Hydropathy plots of tbCRK4 and cfCRK

Hydropathy plots of the tbCRK4 and cfCRK4 proteins were produced by the method of Hopp & Woods, using the ProSite tool at the ExPasy website. A) tbCRK4: The first insert domain covers residues 140-210 (red line), the second insert domain covers residues 312-390 (green line). B) cfCRK: The first insert domain covers residues 144-210 (red line), the second insert domain covers residues 311-400 (green line). In both cases the X axis represents the sequence of residues, the Y axis represents the relative polarity of those residues. Hopp and Woods plots are essentially hydrophilicity plots, with polar amino acids assigned a positive value and hydrophobic amino acids assigned a negative value. Consequently, in both plots, hydrophilic regions have positive values and hydrophobic regions have negative values.

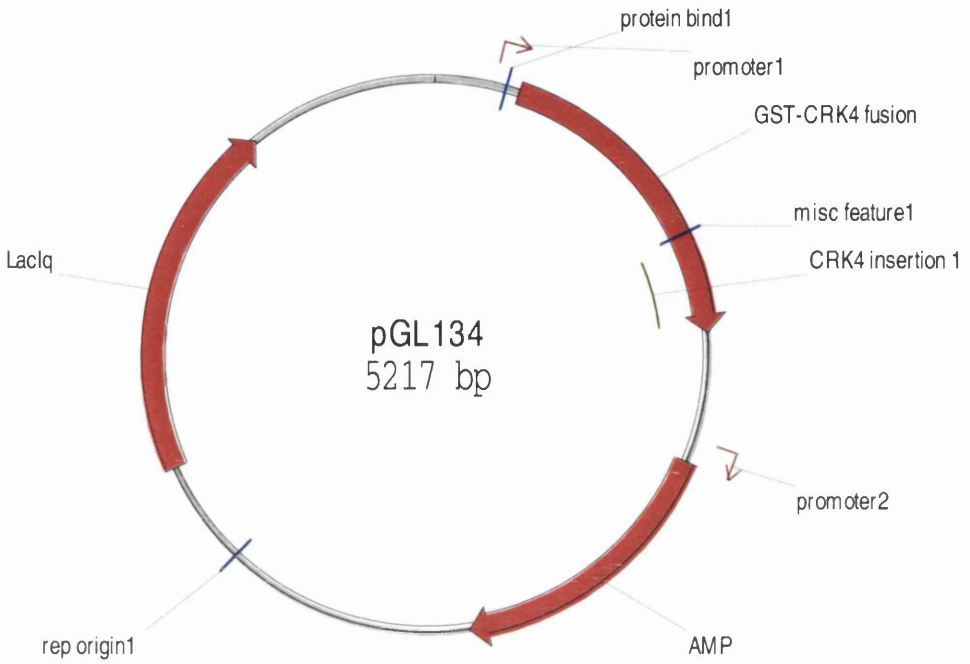


Fig. 3.12: GST fusion expression construct pGL134

The construct for the expression of recombinant GST-CRK4 was generated by excision of the sequence encoding Insert Domain 1 (244 bp) from pGL119 with *Hinc*II-*Hae*III digestion. The 244 bp fragment was cloned into the *Sma*I site of pGEX-5X-2 vector, and screened for orientation by PCR. The ligation junction between the GST and CRK4 fragments was sequenced for one positive, pGL134, before it was transformed to BL21 *E. coli* for pilot expression. The GST fragment is approximately 26 kDa.

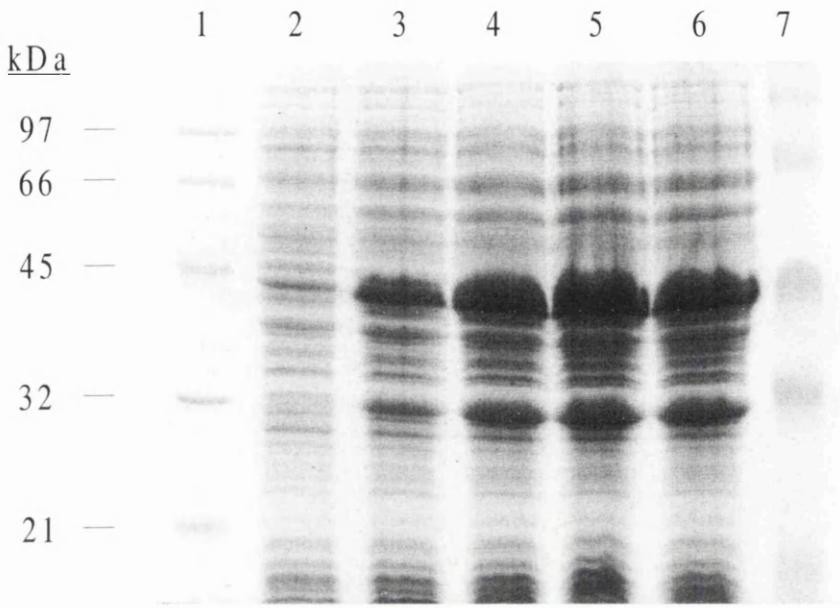


Fig. 3.13: Pilot expression of GST-CRK4

20 ml of LB medium was inoculated at 1:100 with a saturated overnight culture of BL21pGL134 *E. coli* and grown, with shaking, at 37°C to an OD at 600 nm of 0.8. A sample was taken to represent the uninduced control. IPTG was added to a final concentration of 2 mM and induced samples were withdrawn hourly. Crude whole cell lysates of the samples were prepared and electrophoresed on a 12.5% Lammeli gel, stained with Coomassie blue and dried in an EasyBreeze gel dryer (Hoeffer).

Lane 1. Low Range nonprestained markers (BioRad)

Lane 2. T=0 (uninduced)

Lane 3. T=1 hour

Lane 4. T=2 hours

Lane 5. T=3 hours

Lane 6. T=4 hours

Lane 7. Rainbow prestained markers (BioRad)

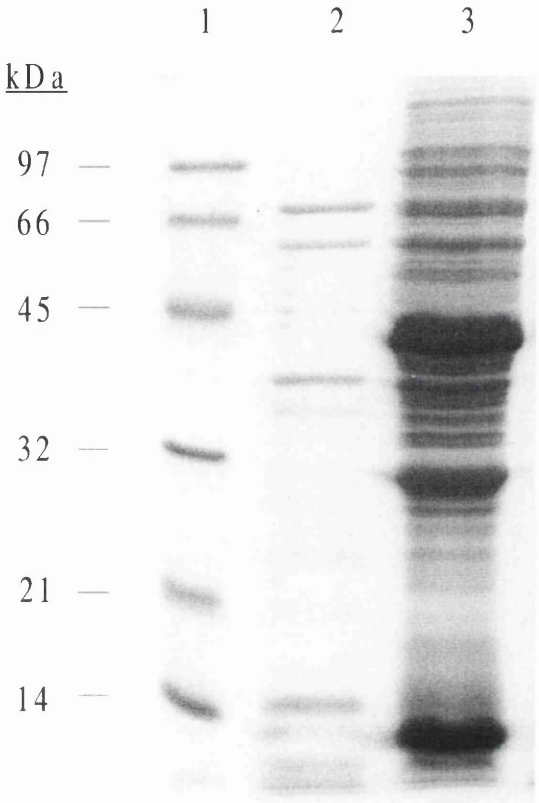


Fig. 3.14: Solubility of recombinant GST-CRK4

Pellets of BL21pGL134 *E. coli* that had been induced for four hours with 2 mM IPTG and subsequently and frozen at -20°C were thawed and resuspended in sonication buffer. Cells were lysed by lysozyme treatment and freeze/thaw cycles, and centrifuged at maximum speed in a bench microfuge for 20 minutes at 4°C to pellet cellular debris. Soluble protein and insoluble protein samples were prepared and electrophoresed on a 12.5% Lammeli gel, stained with Coomassie blue and dried.

Lane 1. Low Range nonprestained markers (BioRad)

Lane 2. Insoluble protein

Lane 3. Soluble protein

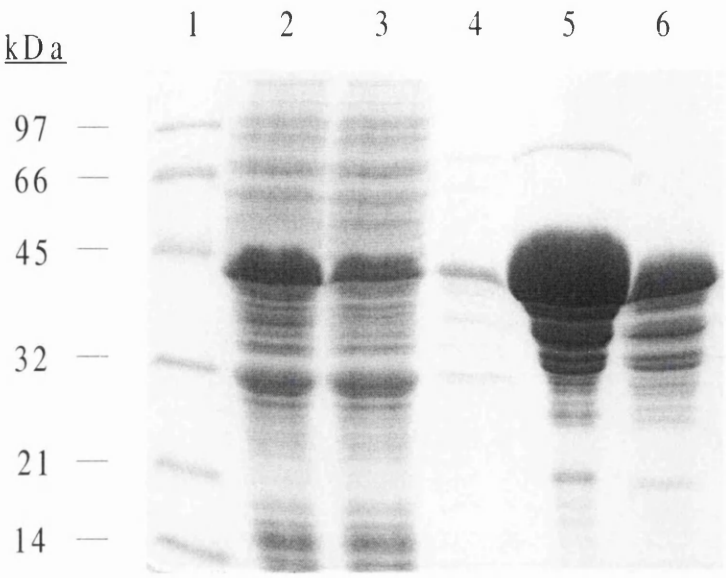


Fig. 3.15: Purification of GST-CRK4

Four flasks each containing 500 ml LB medium with 100 μg ml^{-1} ampicillin were inoculated at 1:100 from a saturated overnight culture of pGL134 in BL21 *E. coli* and grown to an optical density at 600 nm of 0.8. IPTG was added to a final concentration of 0.5 mM and the cultures grown for a further 4 hours before harvesting. Cell pellets were resuspended to 100 ml in 1x PBS pH 7.2 and sonicated using a button tip at 40% amplitude for ten 10 second bursts with an on/off pulse of 1 second, and 2 minutes cooling on ice between bursts. TritonX-100 was added to 1% and the suspension centrifuged at 12 000*g* for 10 minutes at 4°C. The supernatant was passed through a 2 ml bed volume column of GS4B matrix, washed extensively with PBS and eluted with 10 mM reduced glutathione. Eluate fractions were dialysed overnight against a large volume of PBS at 4°C. Samples were electrophoresed on a 12.5% Lammeli gel, stained with Coomassie blue and dried.

Lane 1. Low Range nonprestained markers (BioRad)

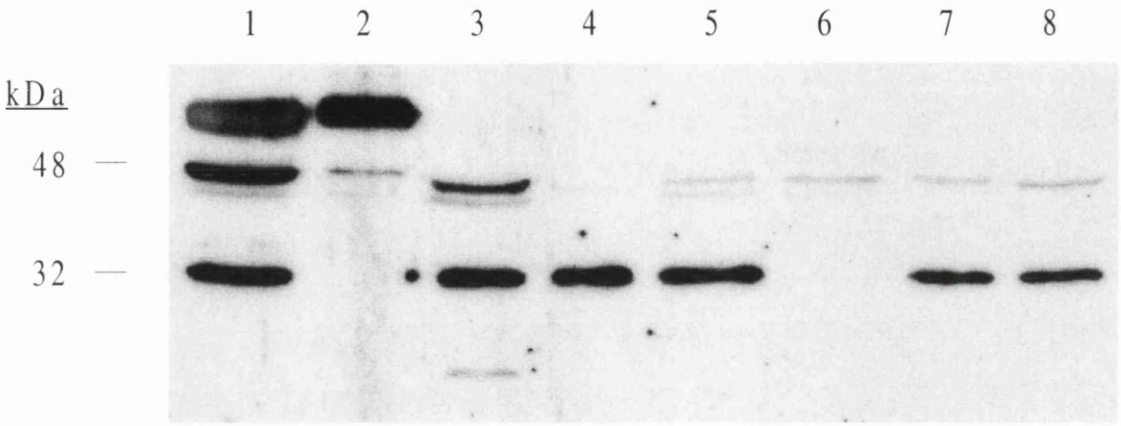
Lane 2. Crude soluble protein

Lane 3. Column flow-through

Lane 4. First wash fraction

Lane 5. Pool 1

Lane 6. Pool 2


1 2 3 4 5 6 7 8
kDa
48
32

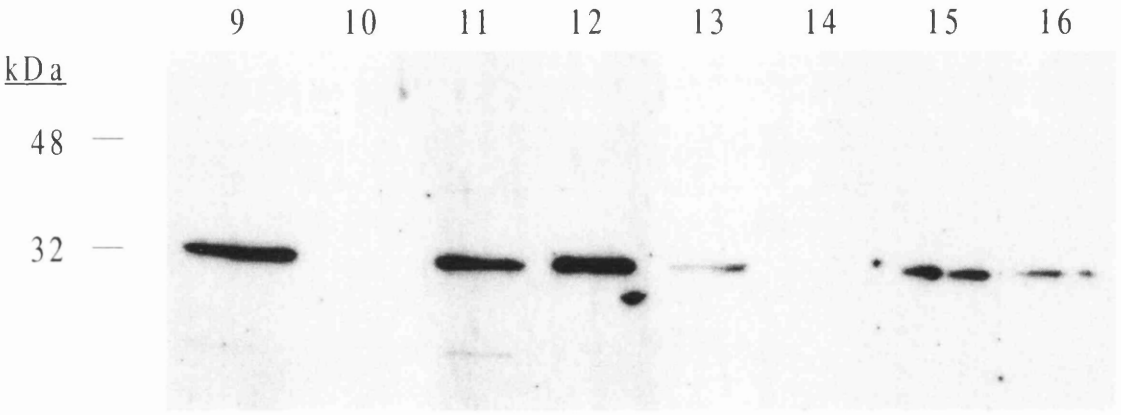

9 10 11 12 13 14 15 16
kDa
48
32

Fig. 3.16: Western blot with partially purified anti-CRK4 antibodies

15 µl per lane of *T. b. brucei* STIB 247 long slender bloodstream, short stumpy bloodstream and procyclic form lysates were electrophoresed on a 12.5% Lammeli gel and transblotted to PVDF membrane at 30 mA overnight at 4°C. Soluble (S-100), insoluble and total protein lysates were made as described in chapter 2.3.4. The anti-CRK4/B antiserum was partially purified by passing an aliquot through an immobilised GST-Sepharose column to remove anti-GST antibodies. After brief staining with Ponceau S the filter was rewetted, blocked and probed with partially purified anti-CRK4 antiserum at a titre of 1:1000. An aliquot of partially purified antiserum was preblocked with 100 µg of purified GST-CRK4 overnight at 4°C. Secondary antibody was anti-rabbit HRP-conjugate (Promega) at a titre of 1:2500, and the blot developed with Supersignal ECL reagents (Pierce). Lanes 1-8 were probed with partially purified anti-CRK4/B antiserum, lanes 9-16 were probed with partially purified anti-CRK4/B antiserum preblocked with fusion protein. As the concentration of protein in the lysates was not known, protein loading between lanes is not equivalent. Markers were the Low Range nonprestained markers (BioRad) and Rainbow prestained markers (BioRad).

Lanes 1, 9. STIB 247 long slender bloodstream form soluble protein

Lanes 2, 10. STIB 247 long slender bloodstream form insoluble protein

Lanes 3, 11. STIB 247 short stumpy bloodstream form soluble protein

Lanes 4, 12. STIB 247 short stumpy bloodstream form insoluble protein

Lanes 5, 13. STIB 247 procyclic form total protein

Lanes 6, 14. STIB 247 procyclic form soluble protein

Lanes 7, 15. *T. b. rhodesiense* TREU 869 long slender bloodstream form soluble protein

Lanes 8, 16. *T. b. rhodesiense* TREU 869 long slender bloodstream form insoluble protein

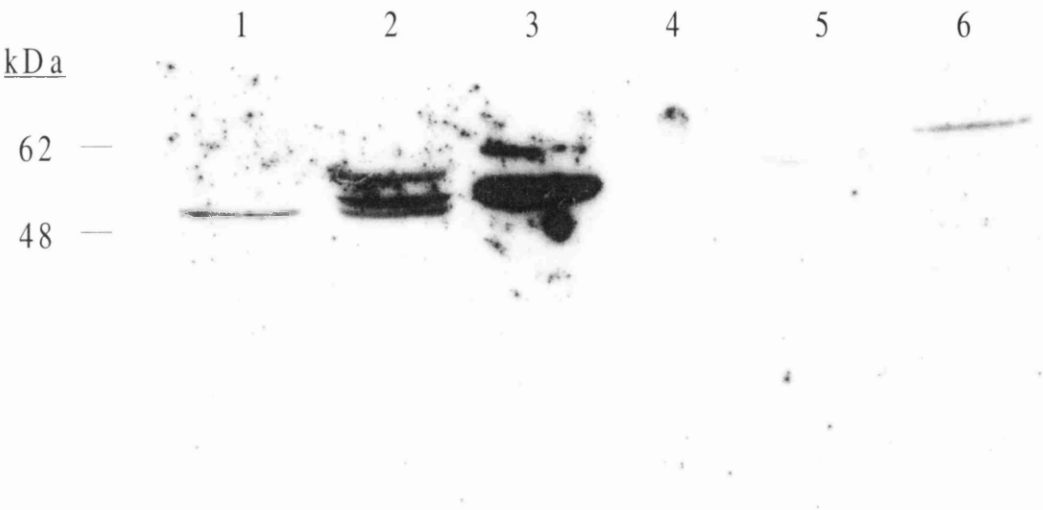


Fig. 3.17: Western blot with purified anti-CRK4 antibodies and soluble fraction protein

Soluble fraction (S-100) protein lysates of *T. b. brucei* STIB 247 long slender bloodstream, short stumpy bloodstream and procyclic forms were made from freshly harvested cells, as described in chapter 2.3.4. 5x 10^6 cells per lane were electrophoresed on a 12.5% Lammeli gel and transblotted to PVDF membrane at 250 mA for 3 hours at 4°C. The partially purified anti-CRK4/B antibodies were further purified by passing an aliquot through an immobilised GST-CRK4 column to specifically purify anti-CRK4 antibodies. Lanes 1-3 were probed with the purified anti-CRK4 antibodies at a titre of 1:50, lanes 4-6 were probed with the purified anti-CRK4 antibodies at a titre of 1:50 that had been preblocked with 100 µg purified GST-CRK4. Secondary antibody was anti-rabbit HRP-conjugate (Promega) at a titre of 1:2500, and the blot developed with Supersignal ECL reagents (Pierce). Markers were the Low Range nonprestained markers (BioRad) and Rainbow prestained markers (BioRad).

Lanes 1, 4. Long slender bloodstream form soluble protein

Lanes 2, 5. Short stumpy bloodstream form soluble protein

Lanes 3, 6. Procyclic form soluble protein

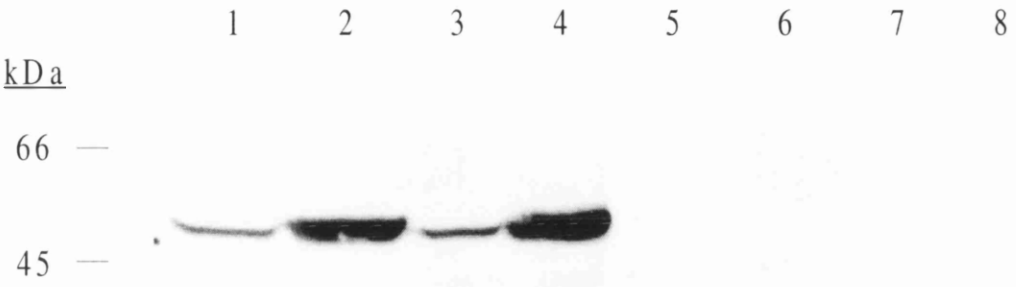
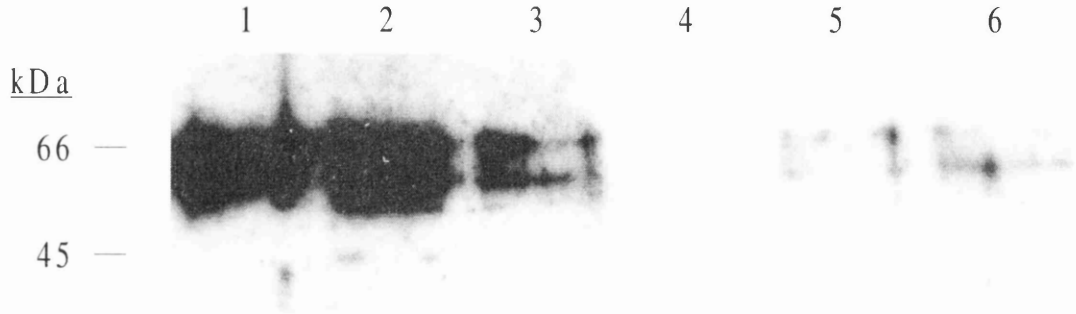


Fig. 3.18: Modified Western blot with anti-CRK4 antibodies

5x 10^6 cells per lane were electrophoresed on a 12.5% Lammeli gel and transblotted to PVDF membrane at 30 mA overnight at 4°C. Soluble (S-100) protein lysates were made as detailed in chapter 2.3.4. The filters were blocked for 8 hours at room temperature on a rolling platform, then probed overnight at 4°C with the purified anti-CRK4 antibodies at a titre of 1:100 (lanes 3 and 4) and 1:500 (lanes 1 and 2), or the purified anti-CRK4 antibodies at a titre of 1:100 and 1:500 that was preblocked with 100 µg GST-CRK4 fusion protein (lanes 7 and 8, and lanes 5 and 6, respectively). Filters were washed extensively with TBST, probed with anti-rabbit HRP-conjugate at a titre of 1:2500 for 45 minutes at room temperature, and again washed extensively with TBST before being developed with a 1:10 dilution of SuperSignal reagents (Pierce). Lanes 1-4 probed with purified anti-CRK4 antibodies, lanes 5-8 probed with the purified anti-CRK4 antibodies preblocked with fusion protein. Markers were the Low Range nonprestained markers (BioRad) and Rainbow prestained markers (BioRad).

Lanes 1, 3, 5, 7. *T. b. brucei* STIB 247 short stumpy bloodstream form soluble protein

Lanes 2, 4, 6, 8. *T. b. brucei* STIB 247 procyclic form soluble protein


1
2
3
4
5
6
kDa
66
45

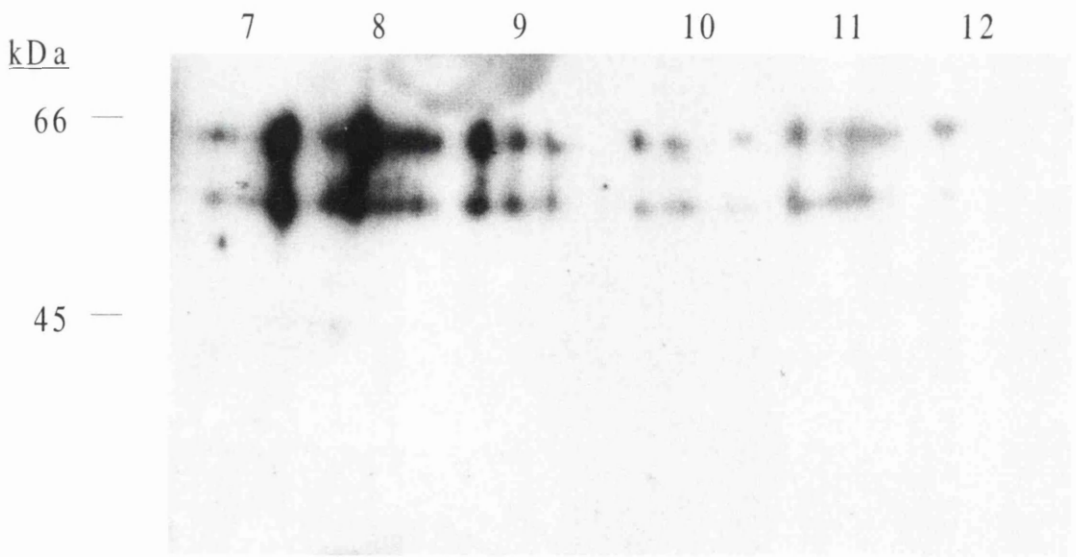

7
8
9
10
11
12
kDa
66
45

Fig. 3.19: Testing CRK4 binding to recombinant p12 and p13

S-100 STIB 247 procyclic extracts were selected using beads that were aminolinked to either TRIS (control beads), purified recombinant $p12^{CKS1}$ or purified recombinant $p13^{SUC1}$. Beads were preincubated in minicolumns with BSA to block nonspecific binding sites and washed in lysis buffer before 100 µl extracts were added to each aliquot of beads (originally, 40 µl of a 50% slurry) and incubated for 3 hours with agitation at 4°C. Columns were extensively washed with lysis solution with glycerol [LSG], high salt lysis solution [HSLS] and lysis solution [LS] before being resuspended in 100 µl SDS-PAGE sample buffer and heated to 100°C for 5 minutes to release bound protein. 10 µl aliquots were electrophoresed on Lammeli gels, transblotted to PVDF membrane and probed with purified anti-CRK4 antibodies or purified anti-CRK4 antibodies preblocked with 140 µg GST-CRK4 fusion protein, both at a titre of 1:50, with the secondary anti-rabbit HRP-conjugate at 1:2500. Blots were developed with Supersignal ECL reagents (Pierce). Lanes 1-6 was a 12.5% Lammeli gel, lanes 7-12 was a 7.5% Lammeli gel.

Lanes 1, 4, 7, and 10 were control beads

Lanes 2, 5, 8, and 11 were $p12^{CKS1}$ beads

Lanes 3, 6, 9, and 12, were $p13^{SUC1}$ beads

Lanes 1-3, 7-9. Purified anti-CRK4 antibodies

Lanes 4-6, 10-12. Purified anti-CRK4 antibodies preblocked with 140 µg GST-CRK4 fusion protein

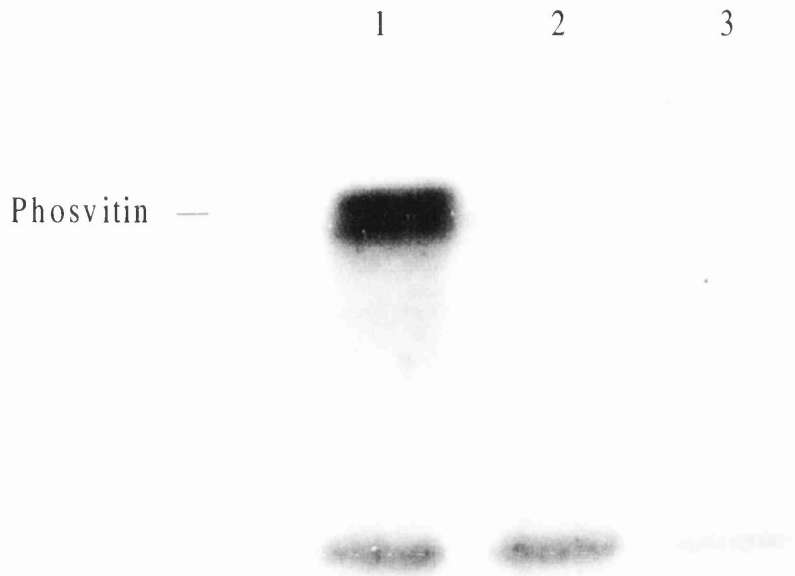


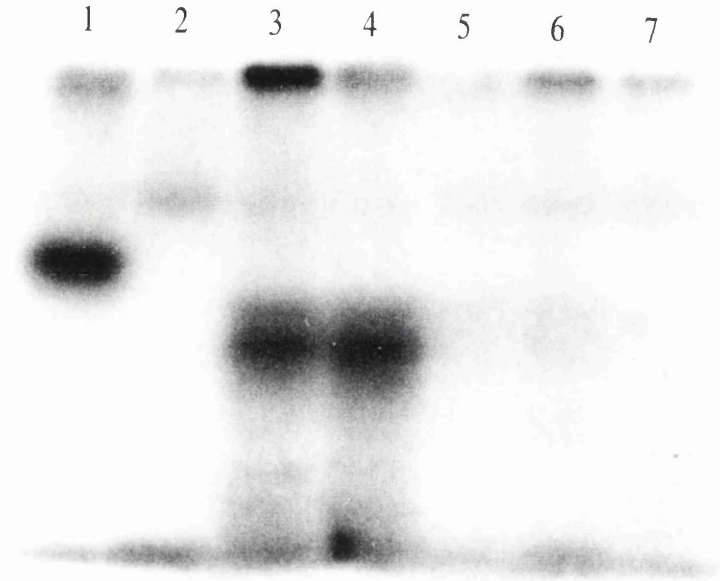
Fig. 3.20: Immunoprecipitation-linked kinase assay with anti-CRK4 antibodies

The purified anti-CRK4 antibodies were used to immunoprecipitate from an S-100 lysate of midlog STIB 247 procyclics (20 µl anti-CRK4/250 µl extract per sample). After binding to protein A-Sepharose beads (100 µl of a 50% slurry per minicolumn) for 2 hours at 4°C on a rolling platform, columns were drained and washed extensively with lysis solution [LS], high salt lysis solution [HSLS] and kinase assay buffer [KAB]. Beads were resuspended in KAB, transferred to eppendorfs and centrifuged briefly to pellet the beads, then resuspended in 20 µl kinase assay mix [KAM] representing 5 µg of substrate protein per reaction, and incubated for 20 minutes at 30°C. Reactions were stopped by addition of 20 µl of TSB and heating to 100°C for 5 minutes. 20 µl of each reaction was electrophoresed on a 12.5% Lammeli gel, which was stained with Coomassie blue, dried and exposed to autoradiography film. Protein substrate was phosvitin. Markers were the Rainbow prestained markers (BioRad).

Lane 1. Anti-CRK4 preimmune serum

Lane 2. Purified anti-CRK4 antibodies

Lane 3. Purified anti-CRK4 antibodies preblocked with 50 µg GST-CRK4 fusion protein


1
2
3
4
5
6
7

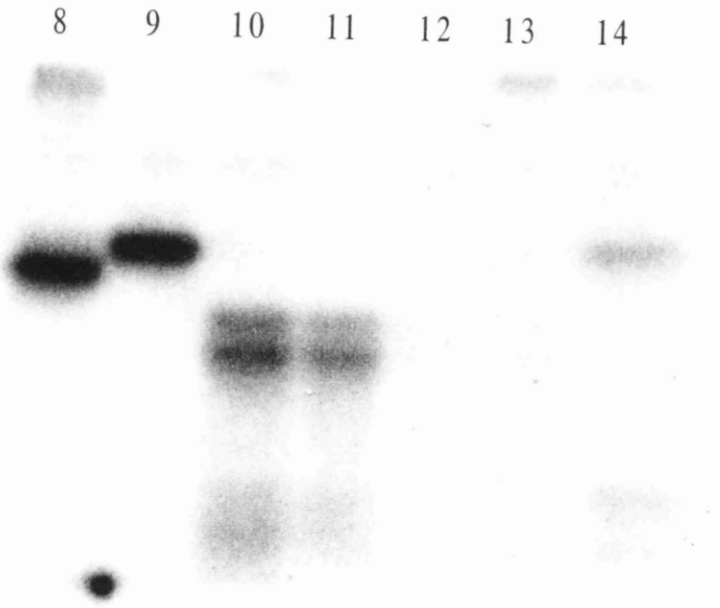

8
9
10
11
12
13
14

Fig. 3.21: Substrate range kinase assay with anti-CRK4 antibodies

The purified anti-CRK4 antibodies were used to immunoprecipitate from an S-100 lysate of midlog STIB 247 procyclics (50 µl anti-CRK4/100 µl extract per sample). After binding to protein A-Sepharose beads and extensive washing with lysis solution [LS], high salt lysis solution [HSLS] and kinase assay buffer [KAB], beads were resuspended in KAB, transferred to eppendorfs and centrifuged briefly to pellet the beads. The beads were then resuspended in 20 µl kinase assay mix [KAM] per sample, representing 5 µg of substrate protein per reaction, for 20 minutes at 30°C. Reactions were stopped by addition of 20 µl of TSB and heating to 100°C for 5 minutes. 20 µl of each reaction was electrophoresed on a 12.5% Lammeli gel, which was stained with Coomassie blue, dried and exposed to autoradiography film. Lanes 1-7 were immunoprecipitates using the purified anti-CRK4 antibodies, lanes 8-14 were immunoprecipitates using the purified anti-CRK4 antibodies preblocked with 50 µg GST-CRK4 fusion protein. Markers were the Rainbow prestained markers (BioRad).

Lanes 1, 8. Phosvitin

Lanes 2, 9. Protamine

Lanes 3, 10. α-casein (batch 1)

Lanes 4, 11. α-casein (batch 2)

Lanes 5, 12. β-casein (batch 1)

Lanes 6, 13. β-casein (batch 2)

Lanes 7, 14. Histone mix

CHAPTER 4: INVESTIGATION OF GENETIC AND BIOCHEMICAL FUNCTION OF THE *tbCYC1* GENE AND tbCYC1 PROTEIN

4.1: Introduction

The *Trypanosoma brucei CYC1* gene was the first cyclin-like gene to be identified from trypanosomatids (Affranchino *et al.*, 1993). It was initially isolated by PCR using two primers. Firstly, a degenerate oligonucleotide corresponding to one of the most conserved regions of amino acid sequence from a range of cyclins (Minshull *et al.*, 1990) taking into account trypanosome codon bias to limit the redundancy, and secondly a sense primer corresponding to the mini-exon splice leader sequence (Boothroyd and Cross, 1982). PCR using these primers on procyclic cDNA of *T. brucei* consistently gave an amplification product of 500 bp. Using this fragment as a probe, the full length gene was isolated from a λgt10-cDNA library prepared from *T. brucei* procyclic poly[A]+ RNA (Affranchino *et al.*, 1993).

Sequence analysis showed the encoded protein, CYC1, to have the highest level of sequence identity to the mitotic-type cyclins (A and B types) and to possess a motif for ubiquitin modification characteristic of such cyclins, called the destruction box (Glotzer *et al.*, 1991). Initial investigation reported that *CYC1* could rescue a fission yeast *CDC*13^{ts} mutant, suggesting a functional role in mitosis control in trypanosomes (Affranchino *et al.*, 1993). Furthermore, immunoprecipitations using a specific anti-CYC1 antiserum coprecipitated a 34 kDa polypeptide reactive to an anti-PSTAIRE monoclonal antibody, and this immunoprecipitated complex displayed kinase activity towards histone HI. In a separate study it was demonstrated that there is a 3- to 5-fold increase in transcript levels of *CYC1* during differentiation from the short stumpy bloodstream form to the procyclic form, suggesting a specific role for tbCYC1 during the differentiation process (Hua *et al.*, 1996).

In this chapter an analysis of the *CYC1* gene and of the CYC1 protein was carried out to attempt to delineate the function of the gene and protein in the procyclic form. A full length clone of *CYC1* was kindly supplied by Dr. K. Matthews (University of Manchester). The gene had been amplified by PCR from *T. b. brucei* STIB 427 genomic DNA and cloned into the *Eco*RV site of pBluescript (SK+) vector (plasmid pGL106; Fig. 4.1).

4.2: Results

4.2.1: Production of recombinant CYC1

The full-length *CYC1* gene was released from pGL106 by digestion with *Bam*HI and *Cla*I, and cloned into the *Bam*HI and *Sma*I sites of the pQE32 vector (Qiagen) to give pGL107. As *Cla*I does not generate blunt ends, pGL106 was treated with the Klenow fragment of *E. coli* DNA Polymerase I after digestion with *Cla*I to generate a blunt end before the gene was released by *Bam*HI digestion. After cloning to pQE32 the ligation joins were then sequenced to confirm the identity and orientation of the insert and to check that the reading frame of the gene was correct with respect to the vector's ATG start codon and the six histidine fusion tag. This construct creates a translational fusion putatively encoding the CYC1 protein with an N-terminal hexahistidine tag.

Plasmid pGL107 was transformed to M15[pREP4] *E. coli* and a series of test inductions and purifications were carried out. Initially, a small-scale pilot native purification was performed as per the manufacturer's protocol, using 1 mM IPTG induction for 2 hours (for full details see chapter 2.3.3). Samples electrophoresed on 12.5% Lammeli gels showed no protein at the predicted size of 36-37 kDa being specifically eluted from the matrix. Poor gel resolution due to an abundance of cellular debris made it difficult to visualise any difference between crude whole-cell lysates of uninduced and induced cultures of M15[pREP4]pGL107, therefore precluding clear identification of CYC1His (data not shown).

A pilot purification geared towards determining subcellular localisation of the fusion protein was carried out, again following the manufacturer's protocol (for details see chapter 2.3.3). The result was unambiguous (Fig. 4.2); no induced protein was seen in the soluble fraction (lane 3) or the periplasmic fraction (lane 2), whereas an induced protein of approximately 36 kDa was seen in the insoluble fraction (lane 4). This indicated that the CYC1His was most likely forming an inclusion body within the *E. coli*, possibly due to over-expression, incorrect folding, or both. The level of expression in this pilot was high; subsequent expression levels for large scale inductions were always much lower, for reasons that remain unclear.

Following subcellular localisation a set of varying conditions were tested, to attempt to increase the levels of soluble protein to make a native purification attempt feasible. Various combinations of lysozyme treatment and sonication were tested with extraction with TritonX-100, extraction with Tween-20 and EGTA, induction with lower concentrations of IPTG and induction at lower temperatures. All failed to achieve any solubilisation of CYC1His (data not shown). It was decided from these results that it would not be possible to isolate soluble native CYC1His so a denaturing purification would be necessary.

A pilot denaturing purification to determine the elution conditions of the recombinant protein was therefore carried out; an induced culture was split, pelleted and one pellet resuspended in QIAexpress buffer A [6 M guanidine hydrochloride, 100 mM sodium phosphate, 10 mM TRIS pH 8.0], the other in QIAexpress buffer B [8 M urea, 100 mM sodium phosphate, 10 mM TRIS pH 8.0]. Following lysis by sonication the samples were centrifuged at 4 000*g* at 4°C for 10 minutes and the supernatants were applied to Ni-NTA columns pre-equilibrated with the same buffers as used for resuspension of the cell pellet. Columns were then washed with buffer A or buffer B, as appropriate, until the absorbance of the flow-through at 280 nm was negligible. All columns were then washed with QIAexpress buffer C [same composition as buffer B, pH 6.3] until the absorbance at 280 nm was negligible. The columns were then eluted, either in a stepwise manner with QIAexpress buffers D, E [Both: same composition as buffer B, pH 5.9 and 4.5, respectively] and finally with QIAexpress buffer F [6 M guanidine hydrochloride, 200 mM acetic acid] or with buffer C containing 200 mM imidazole. Samples of all wash and eluate fractions were electrophoresed on 12.5% Lammeli gels, and the gels stained with Coomassie blue to determine the elution profile of CYC1His (data not shown). This denaturing pilot (for full details see chapter 2.3.3) showed the recombinant protein to have bound the column very efficiently, as only trace amounts eluted in the buffer C wash; this likely represented protein that had only partially bound the column, possibly as a result of partial proteolytic degradation. No CYC1His was observed to be eluted in the buffer D and E elution steps, but a protein of about 36 kDa was seen to elute in the first two wash fractions of the buffer F wash. Little if any protein was observed in further fractions of this wash, therefore it was clear that the bulk of recombinant, non-degraded CYC1His was eluting in the first two eluate fractions. A 36 kDa protein was also observed to elute with the buffer C plus 200 mM imidazole elution fractions, but the protein appeared to elute over a larger number of fraction than observed for the buffer F elution step. Consequently, a larger-scale induction and purification based on denaturing conditions and elution by protonation was adopted.

Larger-scale inductions and purifications were carried out, using a culture volume of 400 ml and a final concentration of 2 mM IPTG for three hours for induction. After harvesting the cell pellets were processed as described (for full protocol see chapter 2.3.3). Following elution, the first two eluate fractions were dialysed; the recombinant protein always precipitated rapidly, and upon resuspension was diluted and assayed. Typically, the final concentration was between 1-3 mg ml^{-1}. Samples of two purified, dialysed CYC1His pools were electrophoresed on a 12.5% Lammeli gel (Fig. 4.3). This showed a smear of protein centred on the 36 kDa mark, indicating that some of the protein was aggregating in some manner as well as some protein being degraded.

4.2.2: Production of anti-CYC1 antiserum and primary testing by Western blotting

A single New Zealand white rabbit was immunised with recombinant protein. Aggregated purified protein in phosphate sonication buffer [50 mM sodium phosphate pH 7.8, 300 mM NaCl] was briefly sonicated prior to immunisation to break up the aggregates, and primary immunisation was done with 200 μg protein homogenised with incomplete Freund's reagent to a final volume of 200 μl. Secondary immunisation was performed six weeks later, again with 200 μg protein homogenised with incomplete Freund's reagent to a final volume of 200 μl. Booster immunisation was performed with 20 μg protein homogenised with PBS by vortexing to a final volume of 200 μl.

The first anti-CYC1 antiserum test bleed, which was taken prior to the first boost, was assessed for cross-reactivity to both purified recombinant CYC1His protein and CYC1His expressed in M15[pREP4] *E. coli* (Fig. 4.4). The anti-CYC1 antiserum from the first bleed produced a strong signal when reacted against purified recombinant CYC1His (lane 1); the almost complete abrogation of this signal upon preincubation of the antiserum with an excess of purified CYC1His (lane 4) demonstrated that the antiserum was specific. Furthermore, the smearing of the signal corroborated the Coomassie stained Lammeli gel (Fig. 4.3), which showed the purified protein to be aggregating. The antiserum also recognised proteins in the whole cell lysate of the M15[pREP4] *E. coli* strain with three major proteins, of approximately 15 kDa, 20 kDa and 35 kDa (lane 2); however, preincubation of the antiserum with CYC1His did not compete out any of these three proteins (lane 5), thus showing them to be nonspecific cross-reactions. The anti-CYC1 antiserum detected a protein at about 36 kDa in the induced whole cell lysate of M15[pREP4]pGL107, the same size as recombinant purified protein (lane 3). This signal was largely competed out with preblocking of the antiserum with purified CYC1His (lane 6). In addition, the antiserum detected a large, strong smeary signal stretching from the top of the gel down to 36 kDa (lane 3); this too was largely abrogated by preblocking of the antiserum, confirming the aggregation of the precipitated recombinant protein (lane 6). Interestingly, the antiserum also detected a protein at around 25 kDa (lane 3) that was competed out by preblocking of the antiserum (lane 6), suggesting that at least some of the induced CYC1His was being proteolysed in a specific manner. Furthermore, as this 25 kDa protein did not appear in purified recombinant protein (lane 1) it can be speculated that the degradation occurred at the N-terminus, possibly by a specific endoproteolytic action that cleaved off the histidine tag on a fragment of approximately 10 kDa.

The first bleed antiserum was tested against total protein lysates for STIB 247 procyclics and *T. b. rhodesiense* TREU 869 long slender bloodstream form (Fig. 4.5 and 4.6). The antiserum detected a protein of approximately 36 kDa in the STIB 247 procyclic extract (Fig. 4.5, lane 2), that was competed by preincubating the antiserum with purified recombinant

CYC1His (Fig. 4.5, lane 3). Large molecular weight proteins detected by the antiserum were not specific as they were not competed by the fusion protein. Preimmune serum did not cross-react with any protein in the procyclic lysate (Fig. 4.5, lane 1). As a control for the quality of the cell extracts and the exact protocol used, anti-CRK1 antiserum was used to probe duplicate filters (Fig. 4.5, lanes 4 and 5). The 33 kDa CRK1 protein was detected (Fig. 4.5, lane 4) and was competed by preincubation of the anti-CRK1 antiserum with CLEHPY peptide (Fig. 4.5, lane 5).

A protein of about 36 kDa was detected in the TREU 869 long slender bloodstream form cell extracts by the anti-CYC1 antiserum (Fig. 4.6, lane 2) which was blocked by preincubating the antiserum with purified recombinant CYC1His (Fig. 4.6, lane 3). In addition there was a protein of 32 kDa that was competed with fusion protein. The pattern of large molecular weight proteins detected was very similar to those detected for the procyclic extract, and were also shown to be nonspecific as they were not competed by the fusion protein. The preimmune serum cross-reacted with a single protein of about 65-70 kDa (Fig. 4.6, lane 1) that was also detected by the antiserum (Fig. 4.6, lane 2). As a control, a duplicate filter was probed with the anti-CRK1 antiserum (Fig. 4.6, lanes 4 and 5). The 33 kDa CRK1 was detected (Fig. 4.6, lane 4) and was competed by preincubation of the anti-CRK1 antiserum with CLEHPY peptide (Fig. 4.6, lane 5).

4.2.3: Western blot analysis with anti-CYC1 antiserum

After booster immunisation and exsanguination of the rabbit the bleedout anti-CYC1 antiserum was tested against protein lysates of all three trypanosome life cycle stages studied, STIB 247 procyclic, long slender bloodstream and short stumpy bloodstream forms, to assess the stage-specific regulation of *CYC1* expression. As a putative mitotic-type cyclin the supposition was that CYC1 would be expressed in the proliferative procyclic and long slender bloodstream forms, but not the cell cycle arrested short stumpy bloodstream form.

The anti-CYC1 antiserum was tested against insoluble and soluble protein lysate fractions of the procyclic and long slender bloodstream forms, and total protein, insoluble and soluble protein lysate fractions of the short stumpy bloodstream form (Fig. 4.7). Total protein lysates and insoluble protein lysates were made as described in chapter 2.3.2, using bench microcentrifugation to obtain the insoluble fraction. Soluble lysates were made by centrifuging whole cell lysates at 100 000*g* at 4°C for 45 minutes in a Beckman ultracentrifuge and aspirating the supernatant, as described in chapter 2.3.2. As previously, preblocking of the anti-CYC1 antiserum was effected by preincubation with an excess of purified recombinant protein (typically, an equivalent volume of purified CYC1His diluted to 1 mg ml^{-1}).

In the procyclic soluble lysate two specific proteins were detected (lane 8) at about 36 kDa and 40 kDa, that were competed by preincubation of the antiserum with purified recombinant CYC1His (lane 12). In addition there was a protein at 30 kDa (lane 8) that was not competed with fusion protein (lane 12). Only one protein was detected in the insoluble procyclic lysate (lane 7) at 140 kDa, but this was not specific as it was not competed with fusion protein (lane 11). The preimmune serum did not cross-react to any proteins in either of the procyclic lysates (lanes 3 and 4).

The pattern of proteins detected in the short stumpy bloodstream form lysates was very similar to those detected in the procyclic lysates. In the short stumpy bloodstream soluble lysate two specific proteins were detected (lane 18) at about 36 kDa and 40 kDa, that were competed by preincubation of the anti-CYC1 antiserum with purified recombinant CYC1His (lane 21). An additional protein at 30 kDa and two proteins at 60 and 75 kDa (lane 18) were not specific as they were not competed with fusion protein (lane 21). Only one protein was detected in the insoluble short stumpy bloodstream lysate (lane 17) at 140 kDa, and this was not competed by fusion protein (lane 20). The preimmune serum did not cross-react to any proteins in any of the short stumpy bloodstream lysates (lanes 13, 14 and 15).

The long slender bloodstream lysates produced a different result compared to the procyclic and short stumpy bloodstream lysates. In the long slender bloodstream soluble lysate two proteins were detected (lane 6) at 30 kDa and 75 kDa, although these were not specific as they were not competed with fusion protein (lane 10). Two proteins were detected in the long slender bloodstream insoluble lysate (lane 5), one at 140 kDa and one very large molecular weight protein. Again, these were not competed with fusion protein (lane 9). The preimmune serum did not cross-react to any proteins in either of the long slender bloodstream fractions (lanes 1 and 2). This is contrary to what was seen with the TREU 869 long slender bloodstream lysates (Fig. 4.6) where the 36 kDa CYC1 protein was detected along with another specific protein at 32 kDa. As well as the fact that *T. b. rhodesiense* and *T. b. brucei* are arguably different subspecies, TREU 869 long slender bloodstream form differs from STIB 247 in that it is a long-term culture-adapted monomorphic isolate, incapable of differentiation.

To try and improve the specificity of the anti-CYC1 antiserum and to attempt to remove nonspecific cross-reactions as well as to lower the background signal, the antibodies were purified on a Protein A/G column. An aliquot of anti-CYC1 was diluted 1:1 with Immunopure IgG Binding Buffer (Pierce), applied to a protein A/G column pre-equilibrated with the same buffer, and allowed to flow completely into the gel matrix. The column was washed with binding buffer until the absorbance of the flow-through at 280 nm was negligible, then the antibodies eluted with Immunopure Elution Buffer (Pierce). The eluate fractions were assessed spectrophotometrically at 280 nm and those showing the greatest absorbance were pooled and desalted by dialysis against a large excess of 30 mM sodium borate pH 7.0 overnight at 4°C. The

purified antibodies were tested against total protein lysates of all three life cycle stages of STIB 247 and the TREU 869 long slender bloodstream form (Fig. 4.8). This showed little improvement in the signal quality. The 36 kDa protein was not detected in the STIB 247 long slender bloodstream lysate either with the nonpurified anti-CYC1 antiserum (lane 1) or the A/G purified antiserum (lane 13). CYC1 was very faintly detected in the TREU 869 long slender bloodstream lysate by the nonpurified antiserum (lane 2) and was detected more strongly by the purified antiserum (lane 14). CYC1 was detected by the nonpurified antiserum in the STIB 247 short stumpy bloodstream lysate (lane 3) and the STIB 247 procyclic lysate (lane 4), and by the purified antiserum (short stumpy bloodstream lysate, lane 15; procyclic lysate, lane 16). No specific proteins were detected in any of the lysates by the column flow-though (lanes 5-8) or by the column wash (lanes 9-12).

4.2.4: Immunoprecipitation with anti-CYC1 antiserum

To try to demonstrate association of CYC1 with a CRK protein, and thus corroborate the function of CYC1 as a cyclin homologue, attempts were made to immunoprecipitate CYC1 with the protein A/G purified anti-CYC1 antiserum. Soluble fraction lysates were made from midlog STIB 247 procyclic cells (for full details see chapter 2.3.5). Aliquots of purified anti-CYC1 antiserum were mixed with S-100 procyclic soluble fraction lysates to bind CYC1; this mix was then added to Sepharose-protein A beads that had been preblocked with BSA, and mixed to allow the protein A to bind the Fc portion of the antibodies. The beads were then extensively washed with lysis solution [LS], high salt lysis solution [HSLS] and low salt lysis solution [LSLS] before being eluted by addition of FSB and heating to 100°C for 5 minutes. Samples were electrophoresed on 12.5% Lammeli gels, transblotted to PVDF membranes and probed with anti-CRK1, anti-CRK2 (purified EVREE antibodies, see chapter 5.2.1), and anti-CRK3 antisera (data not shown). None of the three CRKs could be detected in immunoprecipitates of CYC1. However, as the anti-CYC1 antiserum did not detect CYC1 in the immunoprecipitates it was thought likely that the antibodies were not precipitating CYC1 and therefore could not coprecipitate CRKs. This experiment was repeated several times with identical results.

In order to test if CYC1 was associated with an active kinase subunit in STIB 247 procyclic cells, immunoprecipitations were linked to kinase assays (see chapter 2.3.5). Procyclic S-100 soluble lysates were mixed with aliquots of purified anti-CYC1 antiserum; after an incubation period the lysate-antibody mix was added to Sepharose-protein A beads that had been preblocked with BSA, and incubated with mixing to allow the protein A to bind the Fc portion of the antibodies. The beads were then extensively washed with LS, HSLS and kinase

assay buffer [KAB]. Beads were resuspended in a small volume of KAB, transferred to eppendorfs and briefly centrifuged to pellet the beads. The KAB was aspirated and 20 µl of kinase assay mix [KAM] added, mixed in and incubated for 20 minutes at 30°C, before the reaction was stopped by adding an equivalent volume of FSB and heating to 100°C for 5 minutes. The protein substrate in the KAM was varied between experiments, but per reaction represented approximately 5 µg substrate protein. 20 µl samples were electrophoresed on 12.5% Lammeli gels. Gels were stained with Coomassie blue to check for the presence of the antibody chains (to check the binding of antibodies to Sepharose-protein A), dried down and exposed to autoradiography film (DuPont) in hypercassettes (Amersham).

Initial kinase assays for a CYC1-associated kinase activity were promising; the purified antiserum precipitated an activity towards both histone H1 and an α/β-casein mix (Fig. 4.9). A negative control containing no serum gave no signal for both substrates (α/β-casein, lane 1; histone HI, lane 4). The preimmune serum gave a small signal for both substrates (α/β-casein, lane 2; histone HI, lane 5) and the purified anti-CYC1 antiserum a much stronger signal in both cases (α/β-casein, lane 3; histone HI, lane 6). However, as the preimmune serum had not been purified on a protein A/G column it was not possible to compare the kinase activity between the two samples as the concentration of anti-CYC1 antibodies was diluted in the protein A/G purified antiserum.

The crucial control to assess CYC1-associated kinase activity was competition of activity by preblocking of the anti-CYC1 antiserum with purified CYC1His. The initial attempt at preblocking of the anti-CYC1 antiserum by preincubation with an excess of recombinant CYC1His attenuated the signal by approximately 50% (as judged by scintillation counting) but could not abrogate it (data not shown). Further assays were problematic, as the reproducibility between assays was discovered to be poor. A number of explanations could account for the lack of a clearly demonstrable CYC1-associated kinase activity:

(1) The antibodies, although recognising the denatured protein on Western blots, did not cross-react with the native conformation of the protein and therefore did not precipitate the cyclin.

(2) The antibodies precipitated CYC1 in complex with a CRK, but the substrates used to test for activity, histone HI and α/β-casein, were not suitable substrates for the complex.

(3) CYC1 was precipitated, but it did not associate with a CRK and was not a real cyclin.

One further attempt to resolve these options was made, to try to detect a difference in CYC1-associated kinase activity between midlog procyclics and procyclics at the same cell density that had been artificially cell cycle arrested by serum depletion (Fig. 4.10). If CYC1 was a cyclin involved in cell cycle control, then a comparison of rapidly cycling midlog cells and quiescent cells would be expected to show a relative decrease in CYC1-associated kinase

activity in the latter. The signal for preimmune serum was low for midlog cells (lane 1) and stationary cells (lane 2), and although a strong signal was obtained with the anti-CYC1 antiserum for midlog cells (lane 3) there was also a strong signal for stationary cells (lane 4). Various alterations to the general protocol, including modifications of the lysis buffers and wash buffers to include 0.5% sodium deoxycholate and 0.01% SDS to increase stringency of binding, and varying incubation times and protein substrate all failed to solve the problems of poor reproducibility, low signal and high background. In general, more stringent regimens of washing resulted in lower signal but the proportional difference in signal between preimmune serum and immune antiserum did not change. The conclusion drawn was that no significant kinase activity was associated with anti-CYC1 immunoprecipitates.

4.2.5: Selection using recombinant p12 and p13

Active kinase complexes bind to p12/p13, and if CYC1 could be shown to bind in association with a CRK this would provide evidence that it is a cyclin. This is an alternative method of precipitation based on binding by associated proteins rather than antibodies (Mottram and Grant, 1996). Recombinant *Leishmania* $p12^{CKS1}$ and yeast $p13^{SUC1}$ were expressed, purified and coupled to beads by aminolinkage as described (Mottram and Grant, 1996). *T. b. brucei* CRK1, CRK2 and CRK3 have all been shown to bind p12 beads, although CRK1 and CRK2 do not bind p13 beads and CRK3 binds only weakly (Van Hellemond and Mottram, unpublished). To investigate whether CYC1 is associated with CRK1, 2 or 3, indeed with an as yet unidentified CRK, p12/p13 binding assays from procyclic S-100 extracts were performed. 40 µl aliquots of a 50% slurry of control beads, p12 beads or p13 beads were dispensed to minicolumns, washed and preincubated with a saturating quantity of BSA to block nonspecific binding sites; columns were drained, washed with lysis solution [LS] and 100 µl of extract added per column. After incubation for 3 hours at 4°C columns were drained and washed extensively with lysis solution with glycerol [LSG], high salt lysis solution [HSLS] and LS. Beads were resuspended in 100 µl of FSB and heated to 100°C for 5 minutes to release bound protein. 10 µl samples were electrophoresed on 12.5% Lammeli gels, transblotted to PVDF and the membranes probed with the anti-CYC1 antiserum, anti-CRK2 (EVREE) antiserum and anti-CRK1 antiserum (Fig. 4.11).

No proteins were detected with the anti-CYC1 preimmune serum (lanes 1-3); the purified anti-CYC1 antiserum did not detect any protein on the control beads (lane 4) or the p13 beads (lane 6). A single protein was detected for p12 beads (lane 5) at approximately 12 kDa, that was specific as it was competed with fusion protein (lane 8). This protein likely represents p12 that was present at large quantities, as a 12 kDa protein was visible when the membrane was

stained with Ponceau S, and which forced a spurious cross-reaction. As a control, duplicate filters were probed with anti-CRK2 and anti-CRK1 antisera. The purified EVREE antibodies detected the 39 kDa CRK2 on the p12 beads (lane 11) that was competed with peptide (lane 14), but no CRK2 was detected on the p13 beads (lane 12). No proteins were detected on the control beads. The anti-CRK1 antiserum detected the 33 kDa CRK1 on p12 beads (lane 17) that was competed with peptide (lane 20), but no CRK1 was detected on the p13 beads (lane 18). No proteins were detected on the control beads. The control blot was not very clean, suggesting that stringency of washing could have been higher.

4.2.6: Targeted disruption of *CYC1*

Targeted gene disruption is a useful method in the analysis of the function of a gene in any particular life cycle stage. It is also useful in defining if a gene has an essential function; a gene may be designated essential, with certain provisos, if viable null mutants for the gene cannot be isolated following targeted disruption to delete both wild-type alleles. If the gene is not essential such nulls can be isolated and the function of the encoded protein studied by analysis comparing the resulting null phenotype to the wild-type phenotype. To gain more information on the role of *CYC1* in STIB 247 procyclics, gene targeting constructs were developed based on those designed for *CRK2* (pGL108 and pGL110, Fig. 5.8). These constructs exploit antibiotic resistance marker genes flanked by the 5' and 3' processing signals of the PARP gene, which is constitutively expressed in procyclics. Sequences specific to the target gene are cloned to flank the PARP processing signals, and the knockout cassette introduced to procyclics by electroporation; homologous recombination occurs between the target gene specific sequences of the cassette and the gene itself, and transgenic cells are selected by application of the appropriate antibiotic.

The CYC1 open reading frame is 1005 bp long, including the TGA stop codon, coding for 334 residues. Primers were designed and used in PCR to amplify a 380 bp fragment corresponding to the 5' end of the gene, and a 240 bp fragment corresponding to the 3' end of the gene. The 5' fragment represented the first codon after the ATG start codon to codon 129, the 3' fragment represented the last 80 codons, ending at (and including) the stop codon. Primers 5'5' and 5'3' were used in PCR from STIB 247 genomic DNA to amplify the 5' 340 bp fragment, primers 3'5' and 3'3' were used in PCR from STIB 247 genomic DNA to amplify the 3' 240 bp fragment. Both PCR fragments were cloned into pTAG vector and sequenced to confirm their identity using the pTAG vector-specific pTAG5' and pTAG3' primers.

The pUC18/PARP-*BLE*/*CRK2* 5'3' and the pUC18/PARP-*PAC*/*CRK2* 5'3' constructs, intermediaries in the generation of the *CRK2* targeted disruption constructs pGL108 and

pGL110, were used to make the *CYC1* constructs (Fig. 4.12). The 3' *CRK2* sequence was excised by *KpnI/Eco*RI cleavage and replaced with the *CYC1* 3' fragment. Bacterial colony screening by PCR identified positives, and one of these for each marker was used in the second replacement, in which the 5' *CRK2* sequence was released by *SphI/XbaI* cleavage and replaced with the *CYC1* 5' fragment. Owing to the presence of an internal *Hin*dIII site in the constructs, both had to be cloned into the *Hin*dIII-*Eco*RI sites of pBluescript in order to gain unique restriction sites bounding the construct that could be used to release the entire cassette for transfection. Accordingly, digestion with *Hin*dIII/*Eco*RI released the 745 bp 5' fragment and the 1.0 kb (*BLE*) or 1.2 kb (*PAC*) 3' fragment. The 3' fragment was cloned into the *Hin*dIII-*Eco*RI sites of pBluescript. Subsequently, the 5' fragment was cloned to the *Hin*dIII site and bacterial colony screening by PCR identified positives in which the fragment had inserted in the correct orientation (the *CYC1* 5' sequence flanked by the vector's *Sac*I site, the PARP sequence ligating adjacent to the start of the resistance marker) to give the targeted disruption constructs pGL111 (*BLE*) and pGL112 (*PAC*). The entire cassettes were excised for transfection by *SacI/Xho*I digestion. Typically, 100 μg of DNA was digested to completion, with the extent of digestion followed by electrophoresing small aliquots on 0.8% TAE minigels. Once digestion was complete the DNA was purified by phenol-chloroform extraction and ethanol precipitation. Purified DNA was resuspended to 2 mg ml^{-1} in distilled water and frozen as 25 μg aliquots at -20°C until use. Typically, between 25-50 μg of DNA was used in each transfection.

Transfection of STIB 247 procyclics was carried out with the *SacI/Xho*I cassettes from pGL111 and pGL112, both singly and together. Individual phleomycin or puromycin resistant cultures were recovered, however nothing was recovered from the double-transfection culture. The original culture transfected with the pGL112 *SacI/Xho*I cassette (*PAC* resistance marker), was cloned by the semisolid plate method. Efficacy of cloning was very low, with only two clones surviving the procedure (WUMP 833 and WUMP 834). Both of these clones were judged from PCR to have the replacement cassette successfully integrated at the correct locus (Fig. 4.13) and therefore to be single allele replacements, *Δcyc1::PAC/CYC1*. WUMP 833 was then transfected with the *SacI/Xho*I pGL111 cassette (*BLE* resistance marker). Twenty hours after transfection this was split between five flasks, and after these had grown to midlog all five were cloned on semisolid plates. Again, cloning efficiency was poor and the emerging clones were of two types; large, diffuse and 'wet-looking', and small, tight and 'dry-looking', suggesting uneven distribution of moisture on the surface of the plates. Three clones survived transfer to liquid medium (WUMP 878, WUMP 879, WUMP 880). DNA was prepared from these three clones as well as from wild-type STIB 247. DNA was digested with *Hin*dIII, electrophoresed on a 0.8% TBE agarose gel, transferred to a supported nitrocellulose membrane and probed with the full length *CYC1* gene (Fig. 4.14).

Two strongly hybridising fragments of 2.7 kb and 3.8 kb were detected in wild-type

DNA as expected, as *CYC1* has a single internal *Hin*dIII site. The 2.7 kb and 3.8 kb DNA fragments were also detected in the phleomycin-puromycin doubly resistant clones WUMP 878, 879 and 880, indicating that at least one wild-type copy of the gene was present in all three clones. In WUMP 878 and 879 two additional weakly hybridising fragments at approximately 1.9 kb and 3.25 kb were detected. WUMP 880 differed from WUMP 878 and 879 in that the additional weakly hybridising fragments were at approximately 2.0 kb and 3.4 kb.

4. 3: Discussion

The aims of the experiments in this section were to analyse the stage-specific expression of *CYC1*, to analyse CYC1 function by investigating if CYC1 associated with a CRK to give a kinase activity, and to determine whether the gene is essential to the procyclic form of *T. b. brucei*.

Expression of the full-length gene as a fusion protein using the Qiagen histidine-tag system was successful in that expression of CYC1 with an N-terminal six-histidine tag was achieved (Fig. 4.2). However, the fusion protein formed inclusion bodies and attempts to solubilise enough CYC1His to make a native purification feasible were not successful. Consequently, a denaturing purification strategy had to be adopted, and while this allowed purification of CYC1His (Fig. 4.3) it proved problematic in that the purified protein precipitated when the urea was removed upon dialysis. Western blotting using first bleed anti-CYC1 antiserum from a New Zealand white rabbit that had been immunised with the precipitated protein showed that the antiserum was specific as it recognised the purified fusion protein (Fig. 4.4).

Western blotting of the first bleed anti-CYC1 antiserum against *T. brucei* protein lysates detected a protein at the predicted size of 36 kDa (Fig. 4.5, 4.6) that was specific as it could be competed by preincubation of the antiserum with a saturating amount of purified fusion protein. Initial expectations that *CYC1* would be expressed in proliferative life cycle stages of the parasite - if it was a functional cyclin homologue involved in the mitotic aspect of cell cycle control - were bolstered by the presence of the 36 kDa protein in STIB 247 procyclics and TREU 869 long slender bloodstream form. However, after booster immunisation and exsanguination the anti-CYC1 antiserum was tested against lysates of procyclic, long slender bloodstream and short stumpy bloodstream forms of STIB 247 (Fig. 4.7). This clearly showed that CYC1 is expressed in the short stumpy bloodstream form and in procyclics, but not in the long slender bloodstream form. It has been demonstrated (Hua *et al.*, 1996) that levels of *CYC1* transcript and CYC1 protein rise 3- to 5-fold between the short stumpy bloodstream and procyclic form *T. b. brucei* STIB 427, and that this rise in expression correlates closely with the

disappearance of variable surface glycoprotein (VSG) and the appearance of procyclic acidic repetitive protein (PARP). The hypothesis developed was therefore that CYC1 plays a role in the differentiation between the two life cycle stages and is not involved in mitosis in trypanosomes. The pattern of expression of CYC1 observed would support this, at least for STIB 247; as previously mentioned, TREU 869 long slender bloodstream form is a long-term culture-adapted monomorphic isolate that is not capable of differentiation and therefore patterns of gene expression may differ between pleomorphic and monomorphic cell lines.

Key to proving that CYC1 is a functional cyclin is the demonstration of association with a CDK homologue (CRK protein in the case of trypanosomes). Immunoprecipitations with purified anti-CYC1 antiserum from procyclic soluble (S-100) lysates were probed with an anti-CRK1 antiserum and purified EVREE antibodies; in the absence of an anti-tbCRK3 antiserum an anti-lmmCRK3 antiserum was used. None of these antisera detected any specific proteins. Indeed, probing the filters with the anti-CYC1 antiserum did not detect the 36 kDa CYC1 either. The antiserum may contain antibodies to linear peptide sequence and spurious conformational epitopes corresponding to portions of the protein's interior, rather than to native conformational epitopes on the protein's exterior. It is therefore not surprising that such an antiserum would prove good for detecting denatured protein in the context of a Western blot but could be poor at binding native protein for immunoprecipitation.

To further investigate the possible association of CYC1 with a CRK, the recombinant proteins $p12^{CKS1}$ and $p13^{SUC1}$ were used to specifically select CRKs and associated proteins from midlog STIB 247 procyclic S-100 extracts. After selection and extensive washing, bound protein was eluted and used in Western blots that were probed with the anti-CYC1 antiserum (Fig. 4.11). No CYC1 was found to bind to p12 or p13. CRKs 1-3 have been shown to bind p12, and CRK3 only has been shown to weakly bind p13 (Van Hellemond and Mottram, unpublished), therefore if CYC1 was associated with any of these CRKs in procyclics it should have been detected. The fact that it was not, along with the fact that duplicate filters probed with anti-CRK1 and anti-CRK2 antisera clearly showed CRK1 and CRK2 binding, shows that CYC1 does not associate with CRK1, CRK2 or CRK3 in the procyclic form of STIB 247.

Immunoprecipitations linked to kinase assays initially gave indications that a precipitated activity towards α/β-casein and histone HI was present (Fig. 4.9). However, as was noted previously the preimmune serum and purified anti-CYC1 antiserum were not directly comparable. Accordingly, the best criterion for a specific associated kinase activity was competition of precipitable activity by preblocking of the antiserum with a saturating quantity of purified CYC1His fusion protein. Although some attenuation of precipitable activity was seen, this effect was not large and varied considerably between individual assays. The overall conclusion from these experiments was that no significant histone HI or α/β-casein kinase activity was associated with the immunoprecipitates.

The interpretation of this data is problematic, chiefly due to the poor ability of the anti-CYC1 antiserum to immunoprecipitate CYC1. Western blotting showed the protein to be present in the short stumpy bloodstream form and procyclic form but not the long slender bloodstream form, consistent with the hypothesis that CYC1 functions in the differentiation process from short stumpy bloodstream form to procyclic form (Hua *et al.*, 1996). This argues against (though does not disprove) a role for CYC1 in mitosis in proliferative life cycle stages of the parasite. Selection experiments showed that CYC1 is not associated with CRK1, CRK2 or CRK3 in the procyclic form, but does not preclude association of CYC1 with any of these CRKs in the short stumpy bloodstream form or during the differentiation process, although persistence of gene expression and protein production in the procyclic is then puzzling if it is not meaningfully functional after the differentiation process is complete. An alternative explanation is that CYC1 binds an as yet unidentified CRK that has a novel biological activity and does not bind the *Leishmania* p12 protein. A third interpretation is that CYC1 is not a cyclin homologue and that it has a non-cyclin function during the differentiation from short stumpy bloodstream form to procyclic form. Unfortunately, in the absence of further data, no firm conclusions can be drawn.

The genetic experiments may suggest that *CYC1* is an essential gene in the procyclic form, by the criteria that a gene can be deemed essential if it is not possible to delete both wild-type alleles and obtain viable null mutants. It was possible to delete one of the two wild-type alleles (Fig. 4.13) as judged by PCR; however, the primers used sat within the disruption construct. Although viable cells were obtained from transfection to delete the remaining wild-type allele, Southern blotting showed that a wild-type allele was present in all double-resistant clones isolated (Fig. 4.14). Interpretation of this blot is difficult; *CYC1* contains a single *Hin*dIII site at position 721 (from the ATG start codon) and the blot showed hybridising bands of approximately 2.7 kb and 3.8 kb in wild-type STIB 247. The *Hin*dIII site in the disruption constructs is at position 745 from the ATG start codon, which is the beginning of the replacement cassette, therefore the difference between the 5' fragments produced on *Hin*dIII digestion of wild-type genomic DNA as compared to WUMP 878, 879 and 880 genomic DNA would be only 24 bp, too small to be resolved on a Southern blot. However, the size difference between the 3' fragments of these internal *Hin*dIII sites was calculated to be 721 bp (*BLE*) and 921 bp (*PAC*) from the wild-type, therefore a mobility shift of one of the two hybridising fragments produced from wild-type by these sizes was expected. This was not what was observed, as the extra hybridising fragments seen for WUMP 878, 879 and 880 are all smaller then wild-type. Furthermore, WUMP 880 showed a different pattern of hybridising fragments from WUMP 878 and 879. An explanation for these results is difficult, but it is possible that the clones isolated were those in which the cassettes integrated at a locus or loci other than the CYC1 locus. Regardless of what has occurred, the key point is that all double-resistant clones

still contain a wild-type copy of *CYC1*. However, an important consideration is the lack of information concerning the *CYC1* locus. The PARP promoter is powerful, and without knowledge of the locus downstream of *tbCYC1* it is not possible to say that the failure to delete both alleles was due to a requirement for *tbCYC1* i.e. transcriptional read-through could provide abnormal expression of a nearby gene that is deleterious to the procyclic-stage cells. Assurance of *tbCYC1* as an essential gene would require demonstration that the disruption constructs did not directly alter the expression of any nearby gene, and rescue of phenotype by ectopic expression to demonstrate the specificity of the effect (in this case, if *tbCYC1* wild-type alleles could be deleted with ectopic expression of *tbCYC1*). In retrospect, disruption of *tbCYC1* in the long slender bloodstream stage of STIB 247 would have been a better approach, as tbCYC1 is not detected in STIB 247 long slender form cells; in addition, this would have provided an opportunity to analyse the function of tbCYC1 in the transformation form bloodstream to procyclic-form cells.

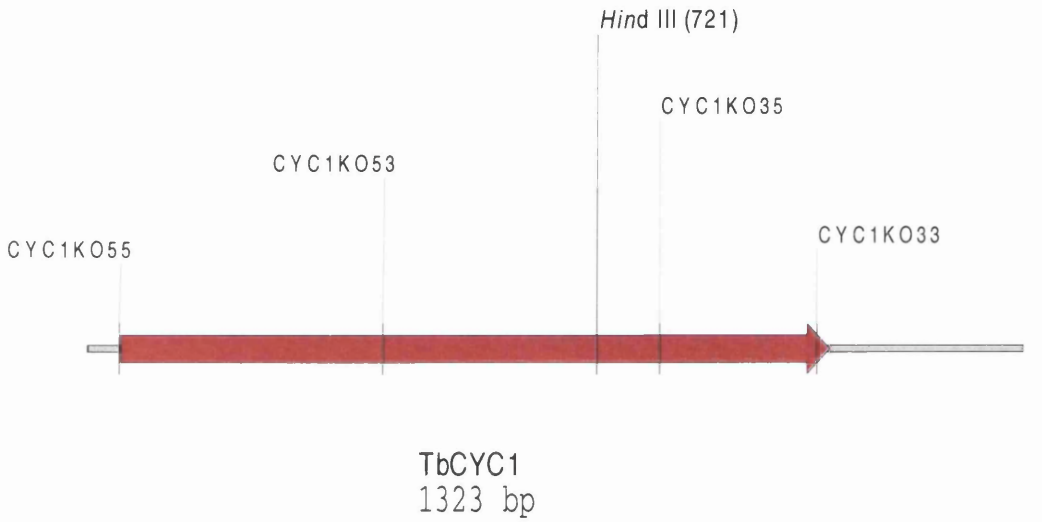


Fig. 4.1: The *T. brucei CYC1* gene

Schematic representation of the CYC1 gene showing relevant restriction sites and primers. The *Hin*dIII restriction site was used to digest DNA for Southern blotting of transgenic clones. The primer combinations of CYC1KO55 with CYC1KO53 and CYC1KO35 with CYC1KO33 were used to amplify fragments corresponding to the 5' and 3' end of the gene, respectively, for the generation of targeted disruption constructs.

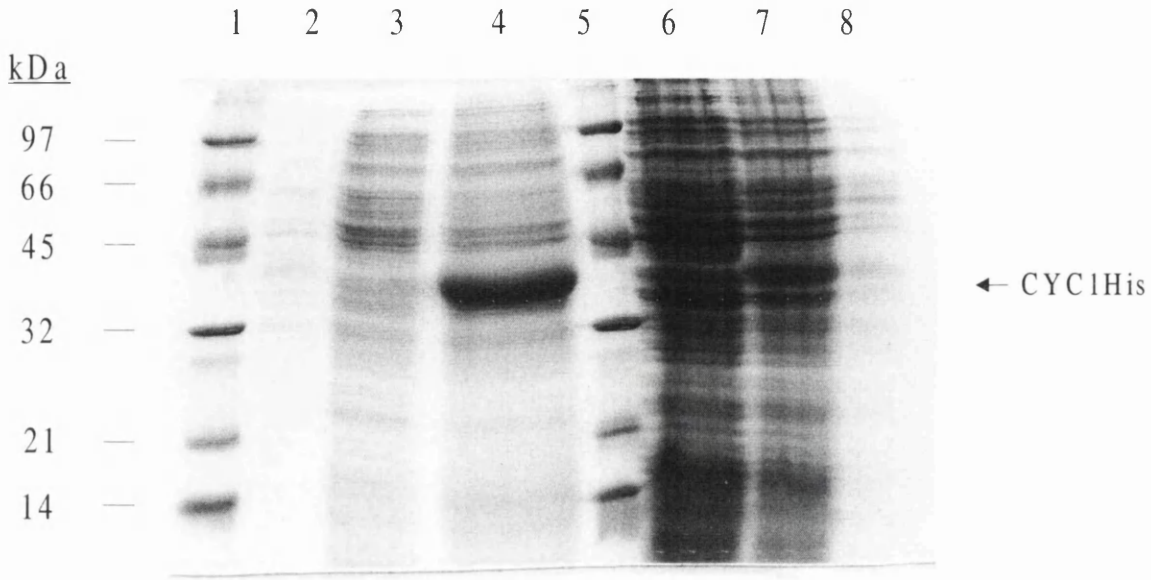


Fig. 4.2: Assessment of the subcellular distribution of CYC1His in *E. coli*

12.5 % Lammeli gel stained with Coomassie blue, showing distribution of CYC1His in different subcellular compartments. The control samples were the M15[pREP4]pGL107 culture immediately prior to induction by the addition of 2 mM IPTG (uninduced control) and a sample of the same culture after induction, immediately prior to harvesting (induced control). The M15[pREP4]pGL107 culture was grown to an OD at 600 nm of 0.8 before addition of IPTG, and left to grow for a further two hours before harvesting.

Lane 1. Low Range nonprestained markers (BioRad)

Lane 2. Periplasmic fraction

Lane 3. Soluble fraction

Lane 4. Insoluble fraction

Lane 5. Low Range nonprestained markers (BioRad)

Lane 6. Uninduced control

Lane 7. Induced control

Lane 8. Periplasmic wash fraction

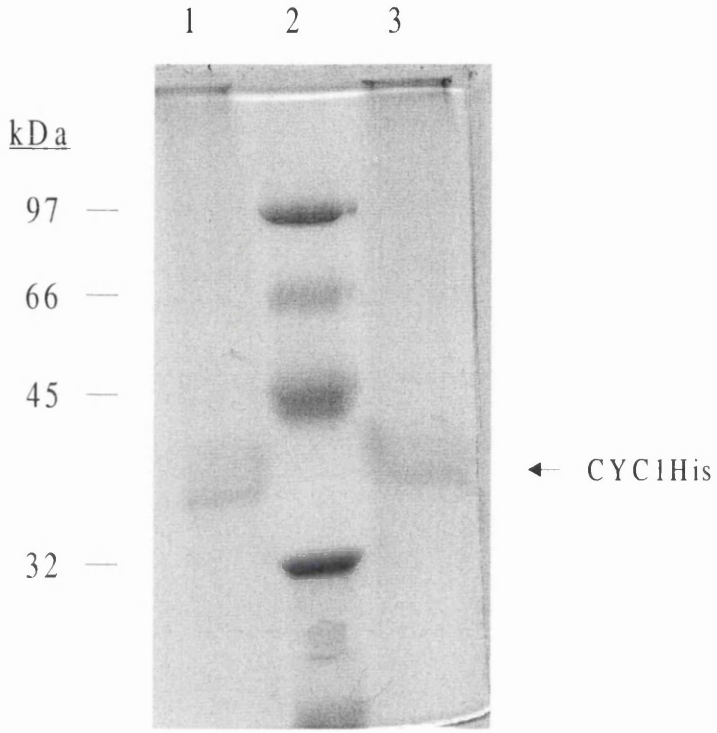


Fig. 4.3: Purified recombinant CYC1His

Samples of two different pools of purified, dialysed recombinant fusion protein, electrophoresed on a 12.5% Lammeli gel and stained with Coomassie blue.

Lane 1. Pool A (approximately 15 µg CYC1His)

Lane 2. Rainbow markers (Amersham)

Lane 3. Pool B (approximately 20 µg CYC1His)

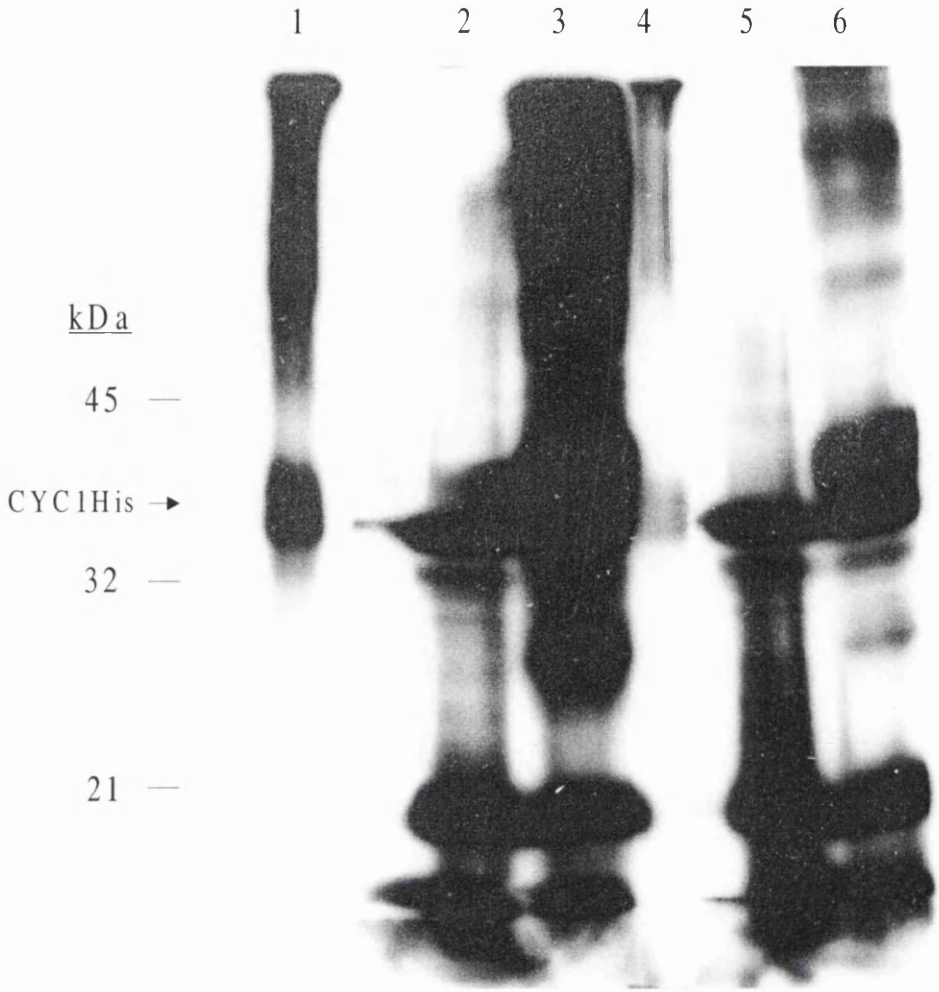


Fig. 4.4: Western blot with first bleed anti-CYC1 antiserum

Crude whole-cell *E. coli* lysates and purified recombinant CYC1His (approximately 5 μg) were electrophoresed on a 12.5% Lammeli gel, blotted to PVDF membrane at 25 mA overnight at 4°C, and the filter probed with first bleed anti-CYC1 antiserum at a titre of 1:100 (lanes 1-3). Duplicate lanes were probed with the first bleed anti-CYC1 antiserum that had been preincubated with an equivalent volume of 1 mg ml^{-1} purified fusion protein (lanes 4-6). Secondary anti-rabbit HRP-conjugate (Promega) was used at 1:2500, and the blot developed with Supersignal ECL reagents (Pierce). Markers were the Kalidescope prestained markers (Amersham) and Low Range nonprestained markers (BioRad).

Lanes 1, 4. Purified CYC1His (approximately 5 μg

Lanes 2, 5. M15[pREP4] *E. coli*, grown for 2 hours with 1 mM IPTG

Lanes 3, 6. M15[pREP4]pGL107 *E coli*, grown for 2 hours with 1 mM IPTG

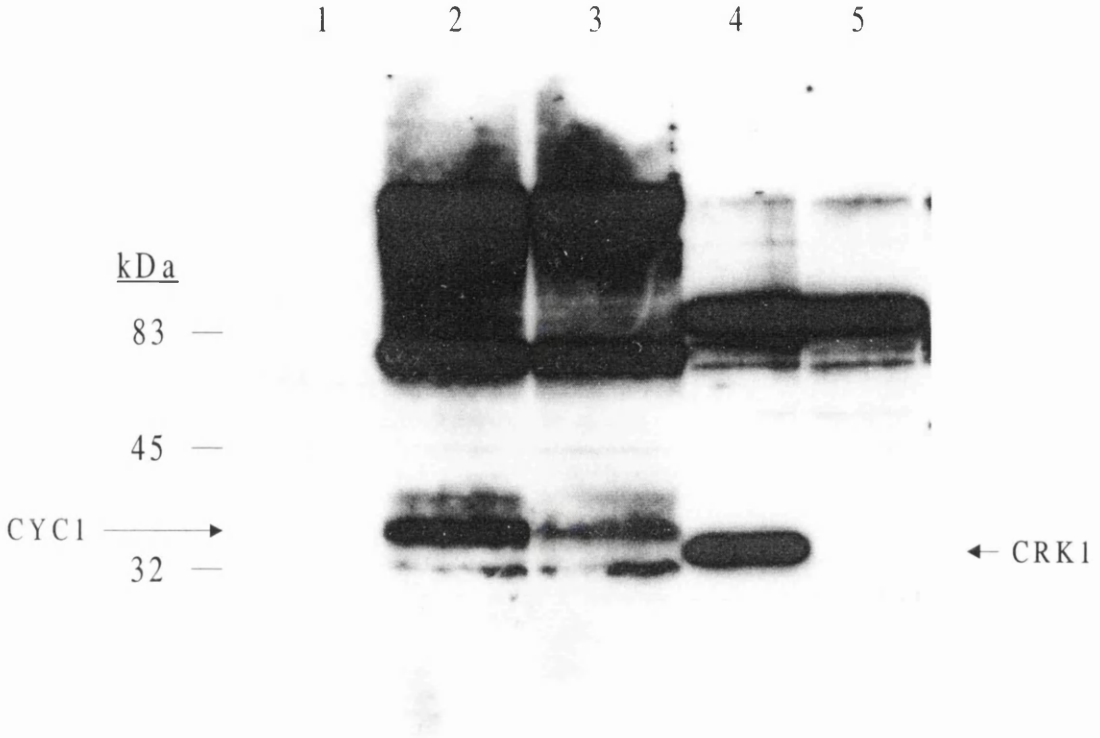


Fig. 4.5: Western blot of STIB 247 procyclic total protein lysate with first bleed anti-CYC1 antiserum

5x 10^6 cells per lane of *T. b. brucei* STIB 247 procyclic total protein lysate was electrophoresed on a 12.5% Lammeli gel at 130 V at room temperature. Protein was transblotted to PVDF membrane using a wet-blot tank (BioRad) at 25 mA overnight at 4°C. The first bleed anti-CYC1 antiserum was used at a titre of 1:2500, the anti-CRK1 antiserum at 1:100. Secondary anti-rabbit HRP-conjugate (Promega) was used at 1:2500, and the blot developed with Supersignal ECL reagents (Pierce). Markers were the Kalidescope prestained markers (Amersham) and Low Range nonprestained markers (BioRad).

Lane 1. Preimmune serum

Lane 2. First bleed anti-CYC1 antiserum

Lane 3. First bleed anti-CYC1 antiserum preblocked with 1 mg ml^{-1} CYC1His

Lane 4. Anti-CRK1 antiserum

Lane 5. Anti-CRK1 antiserum preblocked with 1 mg ml^{-1} CLEHPY peptide

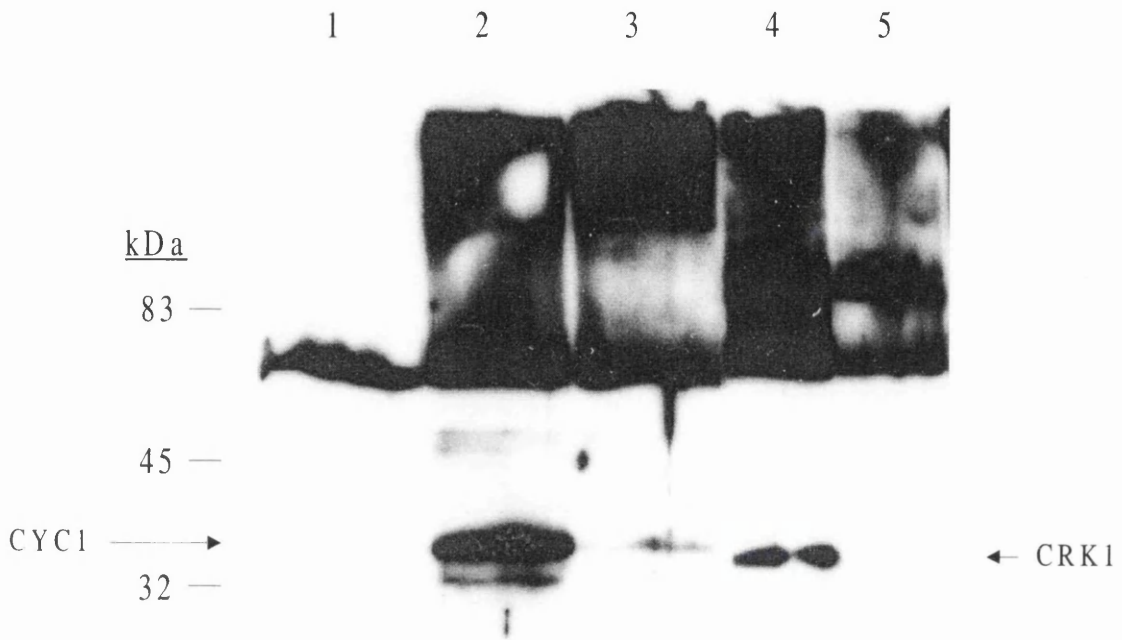


Fig. 4.6: Western blot of TREU 869 long slender bloodstream form total protein lysate with first bleed anti-CYC1 antiserum

5x 10^6 cells per lane of *T. b. rhodesiense* TRUE 869 long slender bloodstream form total protein lysate per lane was electrophoresed on a 12.5% Lammeli gel at 130 V at room temperature. Protein was transblotted to PVDF membrane using a wet-blot tank (BioRad) at 25 mA overnight at 4°C. The first bleed anti-CYC1 antiserum was used at a titre of 1:2500,. Secondary anti-rabbit HRP-conjugate (Promega) was used at 1:2500, the anti-CRK1 antiserum at 1:100, and the blot developed with Supersignal ECL reagents (Pierce). Markers were the Kalidescope prestained markers (Amersham) and Low Range nonprestained markers (BioRad).

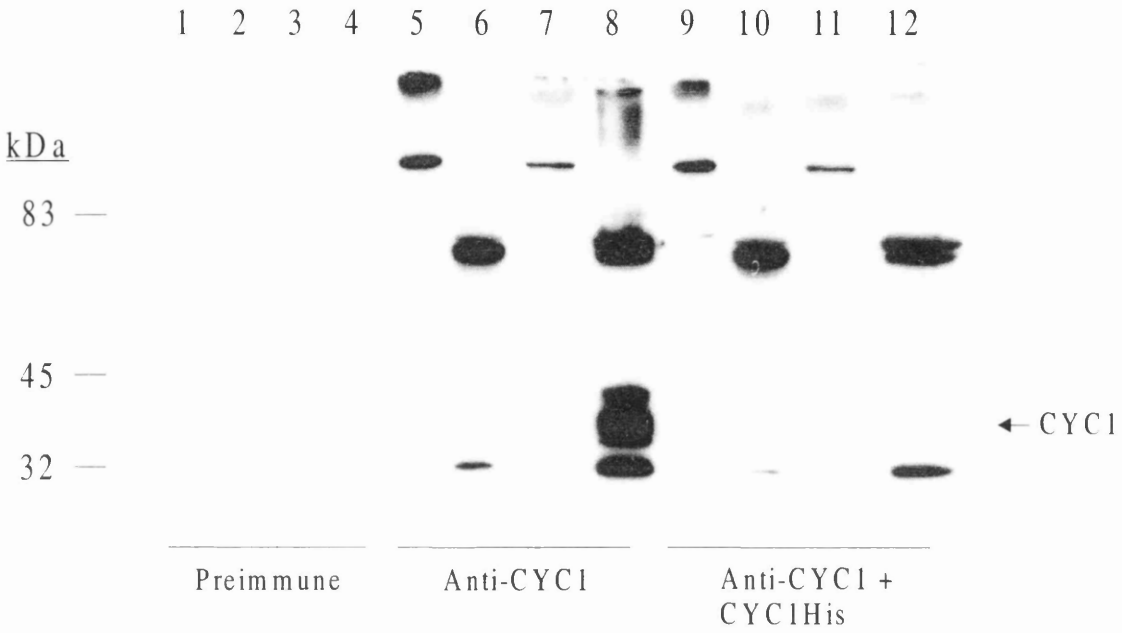
Lane 1. Preimmune serum

Lane 2. First bleed anti-CYC1 antiserum

Lane 3. First bleed anti-CYC1 antiserum preblocked with 1 mg ml^{-1} CYC1His

Lane 4. Anti-CRK1 antiserum

Lane 5. Anti-CRK1 antiserum preblocked with 1 mg ml^{-1} CLEHPY peptide


1 2 3 4 5 6 7 8 9 10 11 12
kDa
83
45
32
CYC1
Preimmune
Anti-CYC1
Anti-CYC1 +
CYC1His

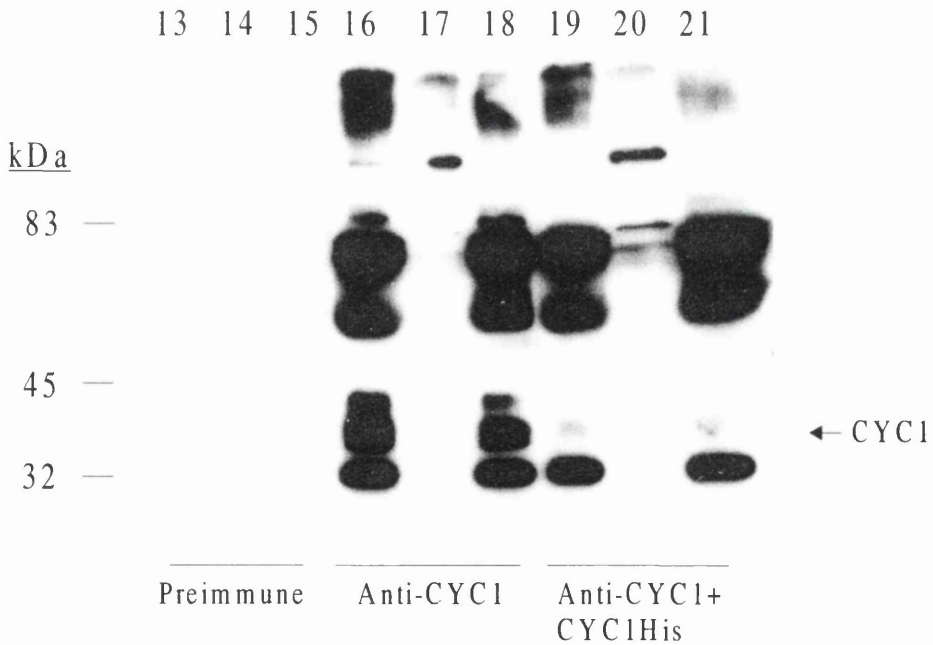

13 14 15 16 17 18 19 20 21
kDa
83
45
32
CYC1
Preimmune
Anti-CYC1
Anti-CYC1+
CYC1His

Fig. 4.7: Western blot of STIB 247 procyclic, long slender and short stumpy bloodstream form lysates with anti-CYC1 antiserum

5x 10^6 cells per lane of *T. b. brucei* STIB 247 procyclic, long slender and short stumpy bloodstream form lysates per lane were electrophoresed on a 12.5% Lammeli gel and the protein transblotted to PVDF membrane by wet-blotting at 250 mA for 3 hours at 4°C. The membranes were probed with final bleed anti-CYC1 antiserum at a titre of 1:100, and secondary anti-rabbit HRP-conjugate at 1:2500, and the blot developed with Supersignal ECL reagents (Pierce). Markers were the Kalidescope prestained markers (Amersham) and Low Range nonprestained markers (BioRad).

Lanes 1-4 and 13-15. Preimmune serum

Lanes 5-8 and 16-18. Anti-CYC1 antiserum

Lanes 9-12 and 19-21. Anti-CYC1 antiserum preblocked with 1 mg ml^{-1} CYC1His

The protein lysates used in the lanes were as follows.

Lanes 1, 5 and 9. Long slender bloodstream insoluble fraction

Lanes 2, 6 and 10. Long slender bloodstream soluble fraction

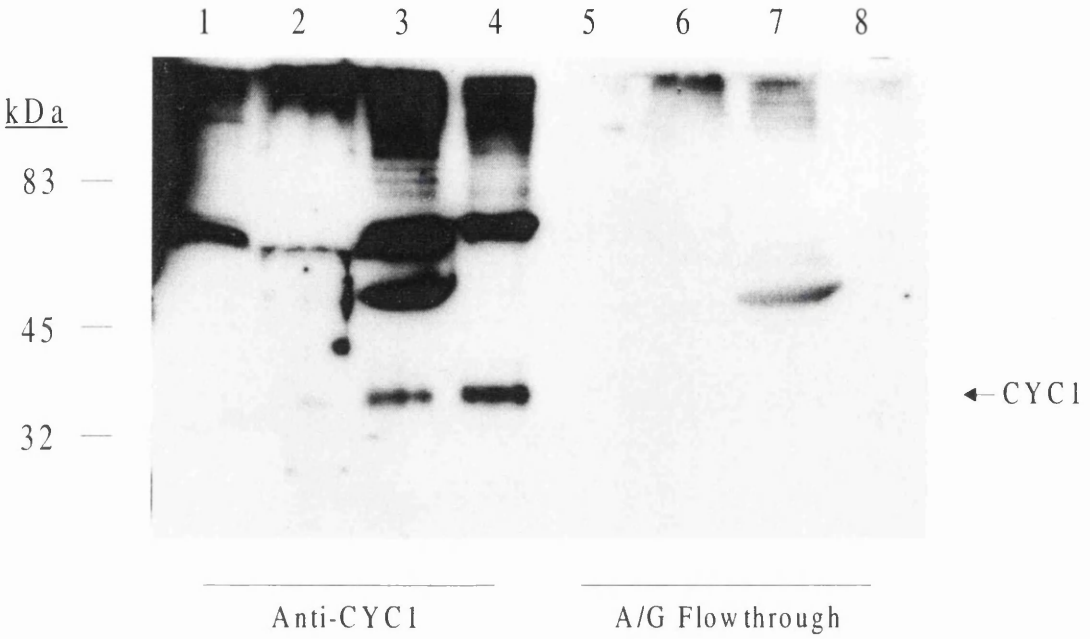
Lanes 3, 7 and 11. Procyclic insoluble fraction

Lanes 4, 8 and 12. Procyclic soluble fraction

Lanes 13, 16 and 19. Short stumpy bloodstream total protein

Lanes 14, 17 and 20. Short stumpy bloodstream insoluble fraction

Lanes 15, 18 and 21. Short stumpy bloodstream soluble fraction


1 2 3 4 5 6 7 8
kDa
83
45
32
CYC1
Anti-CYC1
A/G Flowthrough

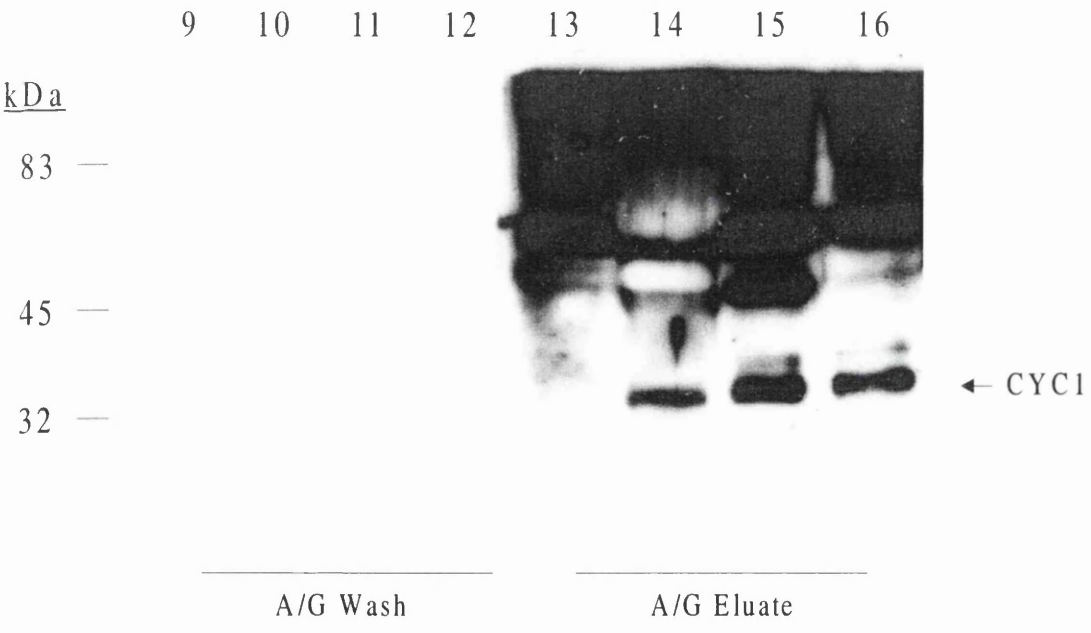

9 10 11 12 13 14 15 16
kDa
83
45
32
CYC1
A/G Wash
A/G Eluate

Fig. 4.8: Western blot with protein A/G purified anti-CYC1 antiserum

5x 10^6 cells per lane of *T. b. brucei* STIB 247 procyclic, long slender and short stumpy bloodstream form lysates and *T. b. rhodesiense* TREU 869 long slender bloodstream form lysate per lane were run on a 12.5% Lammeli gel, and the protein transblotted to PVDF membrane by wet-blotting at 50 mA at 4°C overnight. The anti-CYC1 antiserum was partially purified by passing through a recombinant protein A/G column (see chapter 2.3.4). The filters were probed with the anti-CYC1 antiserum, the A/G column flow-through, the A/G column wash and the A/G column eluate, all at a titre of 1:50. Secondary anti-rabbit HRP-conjugate was used at a titre of 1:2500, and the blot developed with Supersignal ECL reagents (Pierce). Markers were the Kalidescope prestained markers (Amersham) and Low Range nonprestained markers (BioRad).

Lanes 1-4. Anti-CYC1 antiserum

Lanes 5-8: A/G column flow-through

Lanes 9-12. A/G column wash

Lanes 13-16. A/G column eluate

The protein lysates used in the lanes were as follows.

Lanes 1, 5, 9, 13. STIB 247 long slender bloodstream total protein lysate

Lanes 2, 6, 10, 14. TREU 869 long slender bloodstream total protein lysate

Lanes 3, 7, 11, 15. STIB 247 short stumpy bloodstream total protein lysate

Lanes 4, 8, 12, 16. STIB 247 procyclic total protein lysate

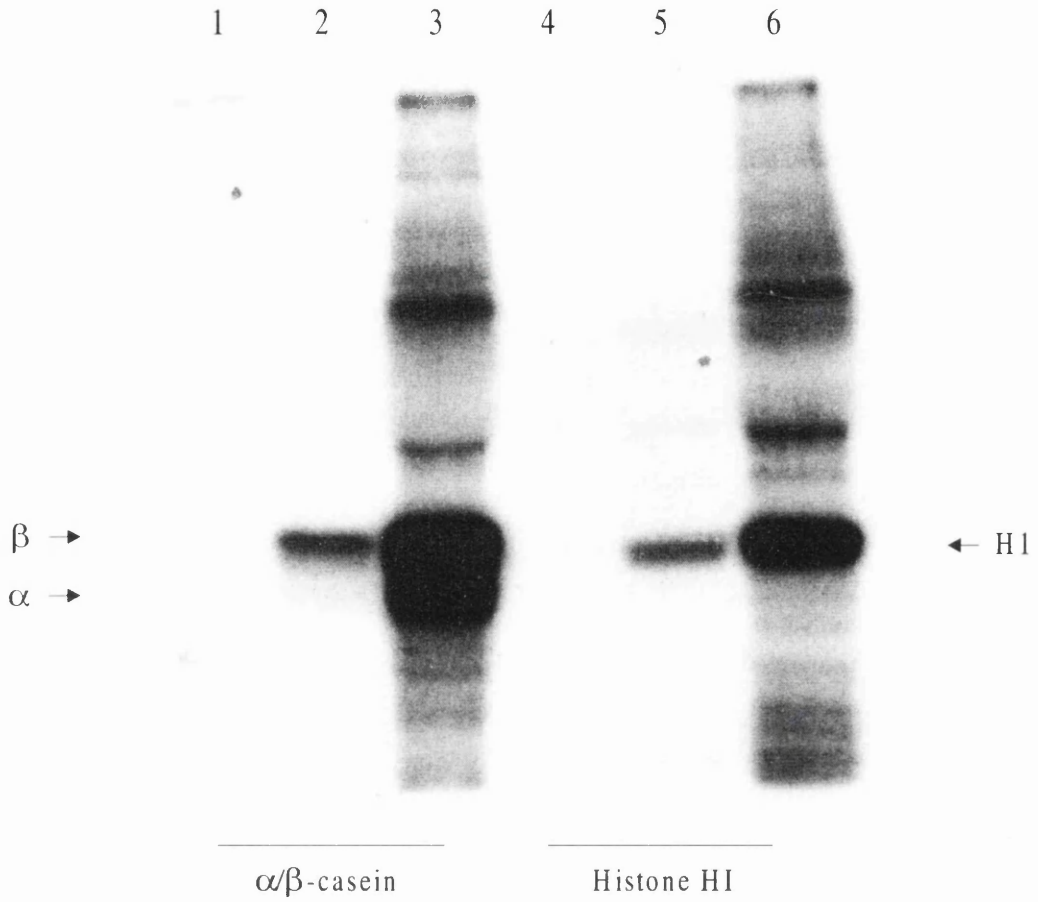


Fig. 4.9: Immunoprecipitation-linked kinase assay with anti-CYC1 antiserum

The protein A/G purified anti-CYC1 antiserum was used to immunoprecipitate from an S-100 lysate of midlog STIB 247 procyclics (20 µl anti-CYC1/200 µl extract per sample). After binding to Sepharose-protein A and extensive washing each sample was incubated with 20 µl KAM, that included 5 µg of substrate protein per reaction, for 20 minutes at 30°C. Reactions were stopped by addition of 20 µl of FSB and heating to 100°C for 5 minutes. 20 µl of each reaction was electrophoresed on a 12.5% Lammeli gel, which was stained with Coomassie blue, dried and exposed to autoradiography film. The substrate in lanes 1-3 was an equal mix of α/β-casein, and in lanes 4-6 was histone HI. Markers were the Low Range nonprestained markers (BioRad).

Lanes 1, 4. No serum

Lanes 2, 5. Preimmune serum

Lanes 3, 6. Protein A/G purified anti-CYC1 antiserum

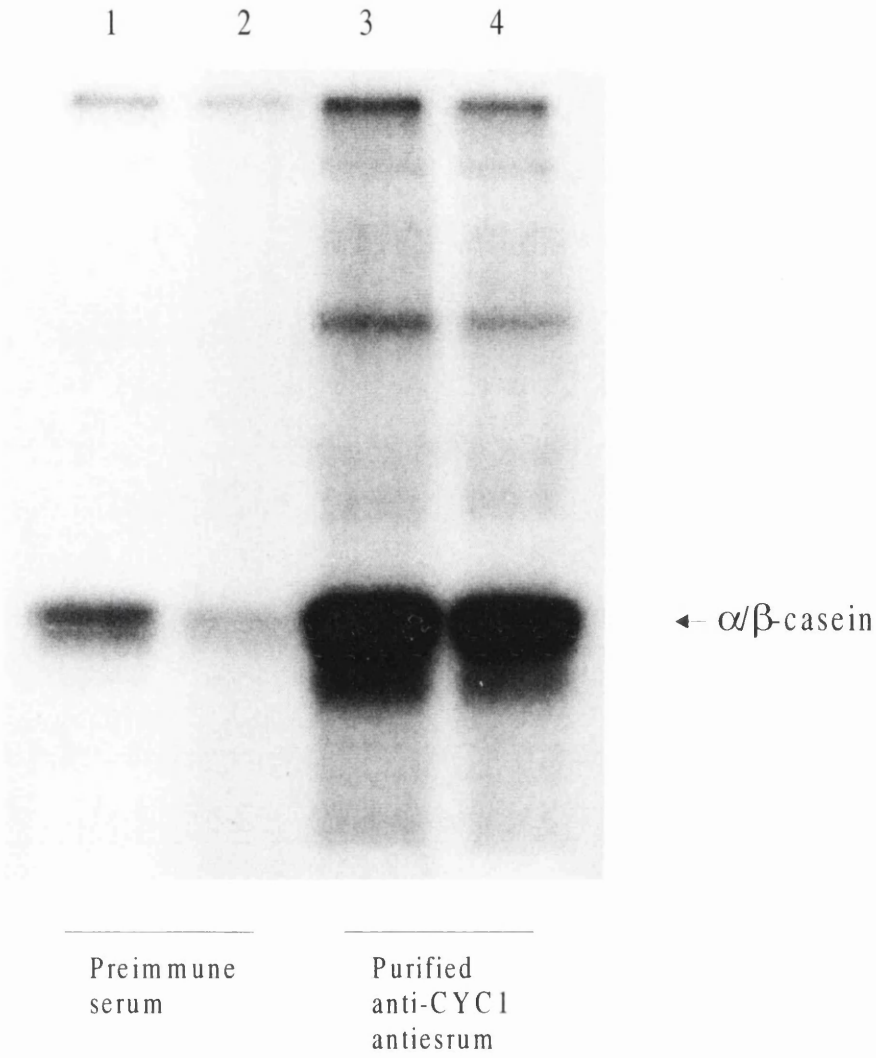


Fig. 4.10: Immunoprecipitation-linked kinase assay comparing midlog and cell cycle arrested cells

The protein A/G purified anti-CYC1 antiserum was used to immunoprecipitate from S-100 lysates of midlog STIB 247 procyclics (cell density of 6x 10^6 cells ml^{-1}) and STIB 247 procyclics that had been artificially cell cycle arrested by serum depletion. The immunoprecipitation and kinase assay protocols were as described in chapter 2.3.5, with 20 µl anti-CYC1/200 µl extract per sample and 5 µg substrate protein per reaction. The substrate was a mix of α/β-casein (2.5 µg of each casein per reaction). Markers were the Low Range nonprestained markers (BioRad).

Lanes 1, 2. Preimmune serum

Lanes 3, 4. Protein A/G purified anti-CYC1 antiserum.

Lanes 1, 3. Midlog procyclics

Lanes 2, 4. Cell cycle arrested procyclics

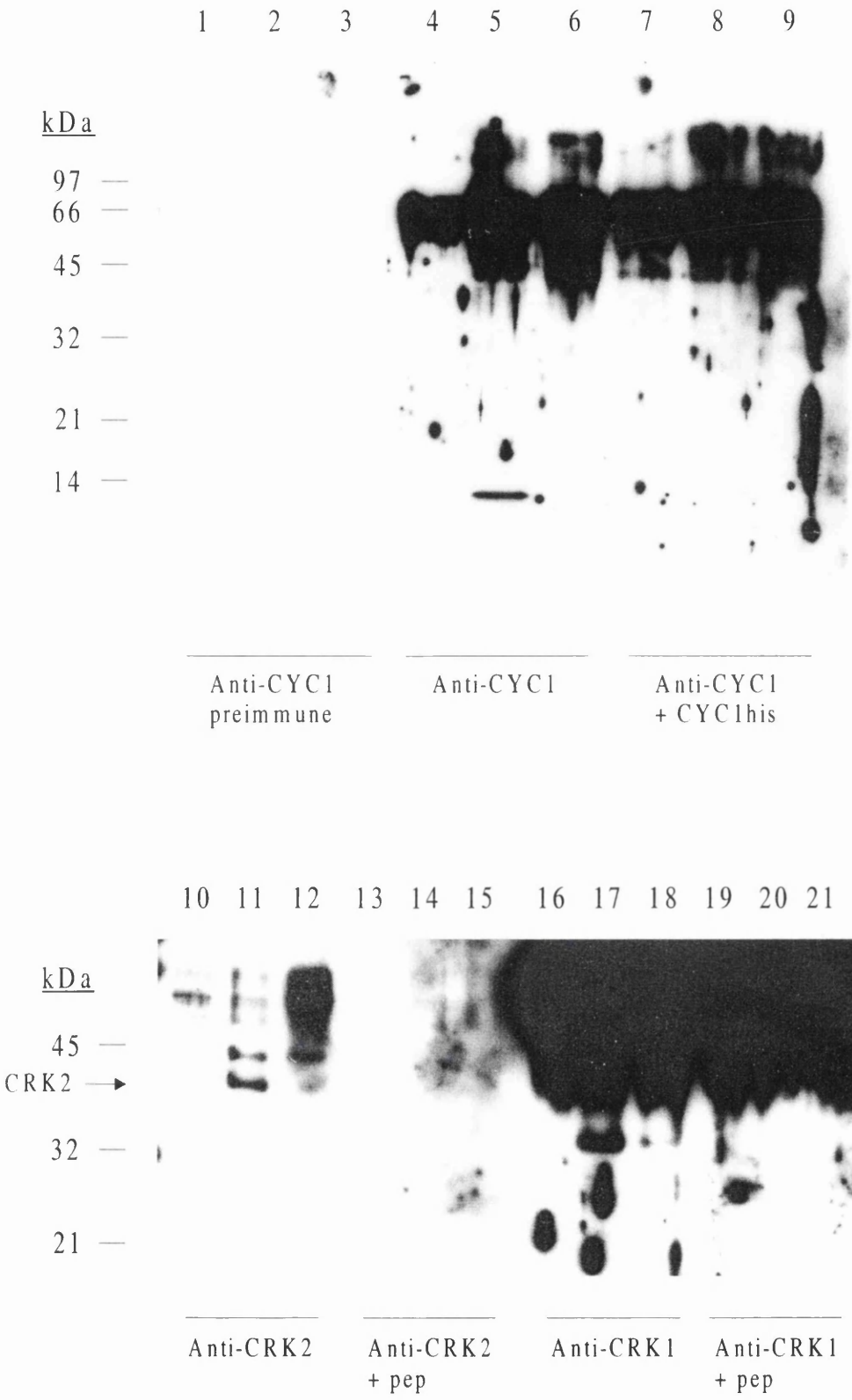

1 2 3 4 5 6 7 8 9
kDa
97
66
45
32
21
14
Anti-CYC1 preimmune
Anti-CYC1
Anti-CYC1 + CYC1his
10 11 12 13 14 15 16 17 18 19 20 21
kDa
45
CRK2
32
21
CRK1
Anti-CRK2
Anti-CRK2 + pep
Anti-CRK1
Anti-CRK1 + pep

Fig. 4.11: Selection of CRK complexes from STIB 247 procyclics using recombinant p12 and p13

S-100 STIB 247 procyclic extracts were selected using beads that were aminolinked to either TRIS (control beads), purified recombinant $p12^{CKS1}$ or purified recombinant $p13^{SUC1}$. 10 µl aliquots of selected samples were electrophoresed on 12.5% Lammeli gels, transblotted to PVDF membrane and probed with antisera at a titre of 1:100, with the secondary anti-rabbit HRP-conjugate at 1:2500; the blot was developed with Supersignal ECL reagents (Pierce). Markers were the Kalidescope prestained markers (Amersham) and Low Range nonprestained markers (BioRad).

Lanes 1, 4, 7, 10, 13, 16 and 19 were control beads.

Lanes 2, 5, 8, 11, 14, 17 and 20 were $p12^{CKS1}$ beads.

Lanes 3, 6, 9, 12, 15, 18 and 21 were $p13^{SUC1}$ beads.

Lanes 1-3. Anti-CYC1 preimmune serum

Lanes 4-6. Anti-CYC1 antiserum

Lanes 7-9. Anti-CYC1 antiserum preblocked with 100 µg CYC1His fusion protein

Lanes 10-12: Purified EVREE antibodies

Lanes 13-15: Purified EVREE antibodies preblocked with 50 µg EVREE peptide

Lanes 16-18. Anti-CRK1 antiserum

Lanes 19-21. Anti-CRK1 antiserum preblocked with 50 µg CLEHPY peptide

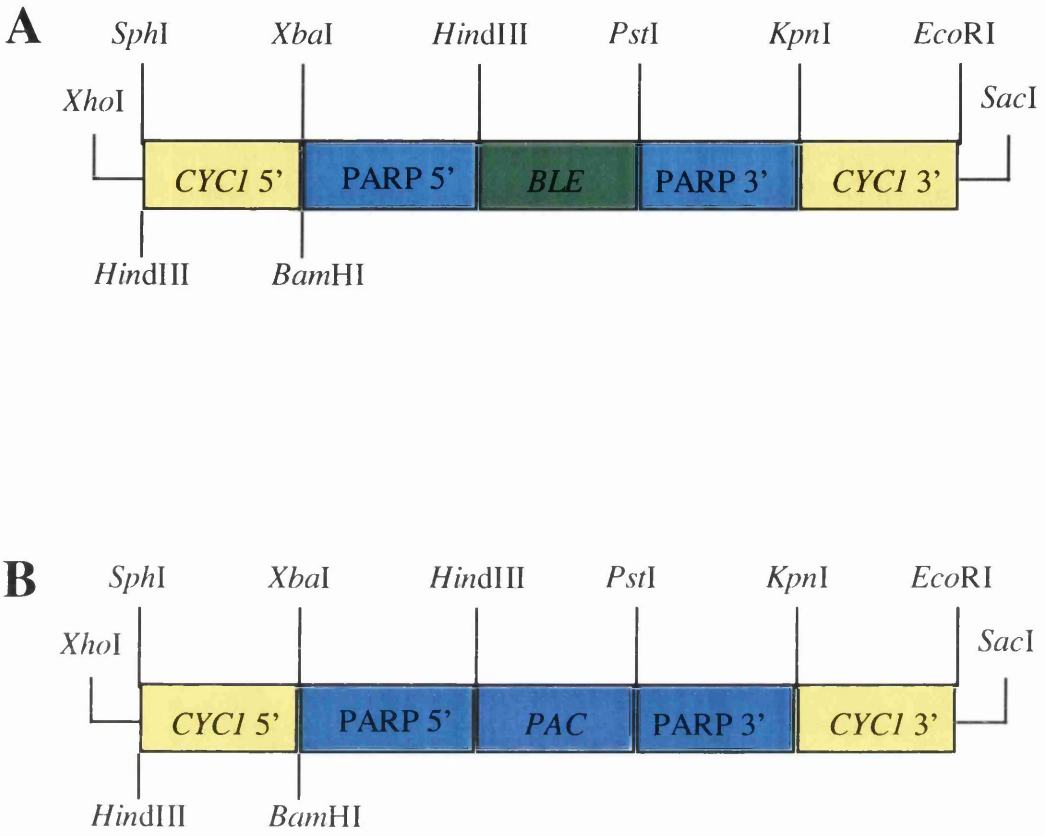


Fig. 4.12: Generation of CYC1 gene disruption constructs

The *CYC1* 5' and 3' sequence fragments were amplified from STIB 247 genomic DNA by PCR and cloned into pTAG as described. The 5' and 3' fragments were released from pTAG using the restriction sites shown and used to replace the *CRK2* 5' and 3' sequences in the intermediary constructs pUC18/PARP-*BLE*/*CRK2*[5'-3'] and pUC18/PARP-*PAC*/CRK2[5'-3']. After specific fragment replacement the *CYC1* disruption cassettes were shuttled to pBluescript, using the *Eco*RI and *Hin*dIII sites shown, to generate the *CYC1* disruption constructs pGL111 (*BLE*) and pGL112 (*PAC*). The cassettes was released for transfection using the pBluescript polylinker sites *Sac*I and *Xho*I, which are not present within the disruption cassettes.

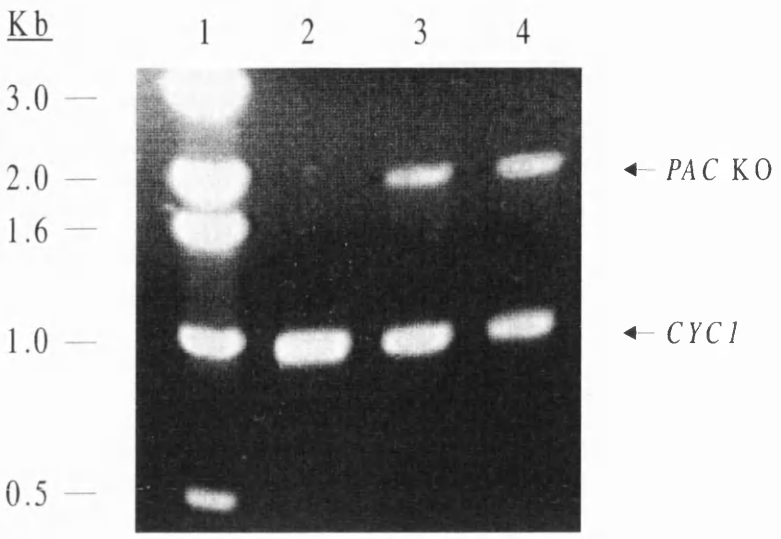


Fig. 4.13: PCR from *Δcyc1::PAC/CYC1* clones genomic DNA

PCR was performed using genomic DNA in 10 µl reactions with 50 pg each of primers 5'5' and 3'3'. DNA was diluted 1:10 in distilled water and 1 µl of this dilution used per reaction. Annealing temperature was 50°C, and the PCR end-point reactions were electrophoresed on a 1% TAE gel, and visualised by staining with 0.5 µg ml^{-1} ethidium bromide.

Lane 1. 1 kb ladder markers (GibcoBRL)

Lane 2. STIB 247 wild-type

Lane 3. WUMP 834

Lane 4. WUMP 833

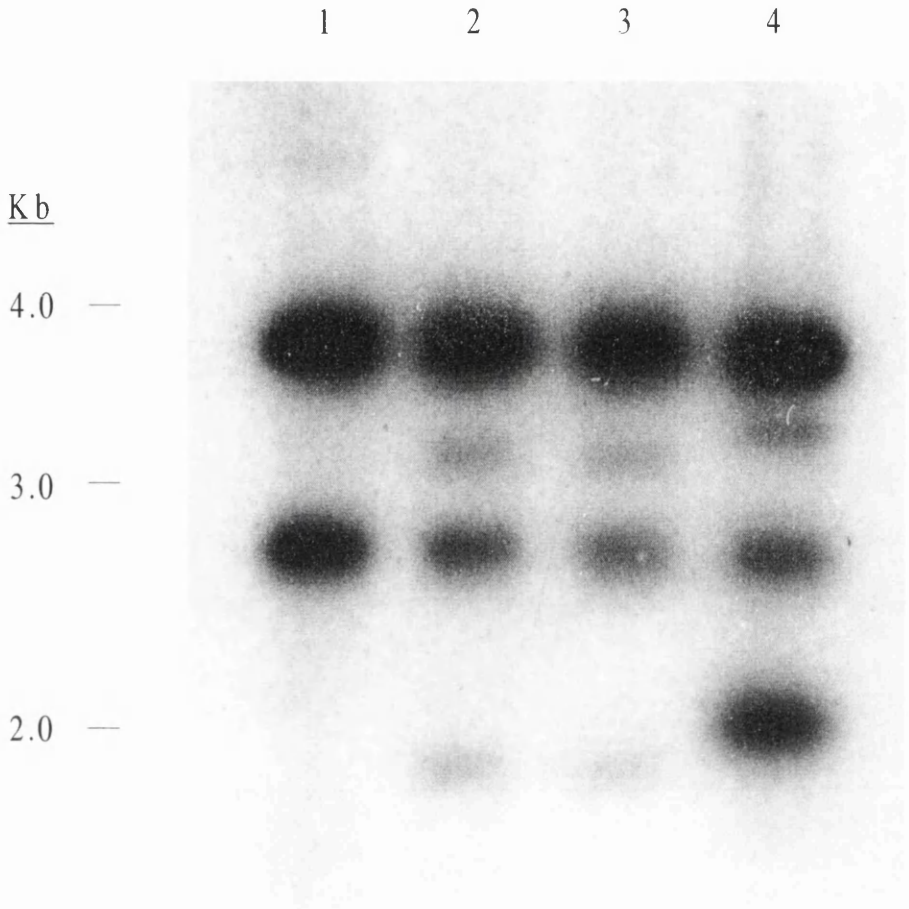


Fig. 4.14: Southern blot of putative *Δcyc1::PAC/Δcyc1::BLE* clones genomic DNA

2.5 µg of genomic DNA was digested with *Hin*dIII and electrophoresed on a 0.8% TBE agarose gel, blotted to supported nitrocellulose membrane by the capillary method and probed with full-length *CYC1*. Markers were the 1 kb ladder (GibcoBRL).

Lane 1. STIB 247 wild-type

Lane 2. WUMP 878

Lane 3. WUMP 879

Lane 4. WUMP 880

CHAPTER 5: ANALYSIS OF THE BIOCHEMICAL AND GENETIC FUNCTION OF THE *tbCRK2* GENE AND tbCRK2 PROTEIN

5.1: Introduction

The *CRK2* gene was initially isolated as two overlapping clones from an EMBL4 λ library of *T. b. brucei* genomic DNA using a low-stringency hybridisation screen with the *lmmCRK1* gene (Mottram and Smith, 1995). A fragment was subcloned from each to pBluescript(+) vector, to give two subclones that overlapped by approximately 200 bp. Subclone pPH2 (pGL116) contained the 5' end of the gene, from the internal *Hin*dIII site at position 675 (relative to the ATG start codon) plus approximately 2 kb of upstream sequence. Subclone pGB1 (pGL117) contained the majority of the gene, from the internal *Bam*HI site at position 229 (relative to the ATG start codon) to a *Bam*HI site approximately 1 kb downstream of the TAA stop codon. The complete open reading frame (ORF) was sequenced plus some 5' and 3' flanking sequence (Fig. 5.1), a schematic is shown along with the relevant primers and restriction sites. *CRK2* differs from other *CDK* homologues from other organisms, as it codes for a 41 amino acid N-terminal leader sequence; only CRK3 has a similar although shorter extension, of 19 amino acids (Mottram and Smith, 1995), which has no significant homology to the CRK2 N-terminal leader. The function of these leaders, if any, is unknown. Phylogenetic analysis does not place CRK2 in strong association with any of the other trypanosomatid CRKs, nor with any of the CDKs from other eukaryotes (Mottram and Smith, 1995), therefore little may be inferred about its role *in vivo*. The aims of the experiments described in this chapter were: 1) To identify CRK2 in different life cycle stages of *T. brucei* by Western blotting with anti-CRK2 antisera. 2) To use immunoprecipitations linked to kinase assays to demonstrate CRK2-associated kinase activity. 3) To use targeted gene disruption to generate null mutants and observe any phenotypic changes. 4) To use inducible systems to overexpress CRK2 and epitope-tagged CRK2 *in vivo* to study phenotypic changes, examine coprecipitating proteins and demonstrate kinase activity.

5.2: Results

5.2.1: Western blot analysis with anti-CRK2 peptide antisera

To attempt to raise specific antisera to CRK2 for use in Western blotting and immunoprecipitations linked to kinase assays, rabbits were immunised with a number of

different peptides deduced from the gene sequence. Two of these peptides were 16-mers and the third an 11-mer. The sequences of the peptides was as follows:-

1) ^{N}E-V-R-E-E-E-V-E-L-K-M-R-F-N-G-A^{C} 'EVREE' peptide

2) ^{N}G-G-V-P-S-T-A-V-R-E-V-S-L-L-R-E^{C} 'PSTAVR' peptide

3) ^{N}C-V-E-L-K-M-R-F-N-G-A^{C} 'CVELK' peptide

'EVREE' corresponds to the C-terminal 16 amino acids. 'PSTAVR' corresponds to the 16 amino acid region of tbCRK2 that is the homologue of the hsCDK1 motif known as the PSTAIRE box (EGVPSTAIREISLLKE); the PSTAIRE box homologue in CRK2 is degenerate, with four amino acid substitutions (GGVPSTAVREVSLLRE) as compared to hsCDK1 (Mottram and Smith, 1995). The third represents the C-terminal 11 amino acids. All three were conjugated to keyhole limpet haemocyanin (KLH) as a carrier protein, and the immunisation schedule followed was as described in chapter 2.3.4, with exsanguination of the animals after the second booster immunisation. Antisera were stored in aliquots at -20°C.

'EVREE' antiserum was purified in a two-step process; firstly, an aliquot of the antiserum was purified on a Protein A/G column (for full details see chapter 2.3.4) to remove contaminating serum proteins and specifically purify the IgG fraction of the antibodies. The EVREE peptide (not conjugated to KLH) was used to make an Sulfolink column as described in chapter 2.3.4, and an aliquot of the partially purified antiserum was passed through this column, thus specifically purifying only antibodies that recognised the peptide. The purified EVREE antibodies were tested against insoluble, soluble (S-100) and total protein lysates of *T. b. brucei* STIB 247 procyclic, long slender and short stumpy bloodstream forms. A number of proteins were detected in all three life cycle stages (Fig. 5.2). For each life cycle stage the pattern of proteins was identical between the soluble and total protein lysates; the purified EVREE antibodies did not detect any proteins in the insoluble lysates (lanes 2, 8, 14). In the procyclic form lysates a protein of approximately 46 kDa was detected, and another protein of 38 kDa was detected more weakly (lanes 1, 3), both of which were competed with peptide (lanes 4, 6). The 38 kDa protein was a doublet. Two proteins were detected weakly at approximately 66 kDa and 90 kDa (lanes 1, 3) that were not competed with peptide (lanes 4, 6). A similar pattern of pattern of proteins were detected in the short stumpy bloodstream and long slender bloodstream form lysates. These proteins were identical to those detected in the procyclic lysates except that the doublet of 38 kDa in procyclics (lanes 1, 3) was a single protein in the short stumpy bloodstream (lanes 7, 9) and long slender bloodstream (lanes 13, 15) form lysates. In the long slender bloodstream form lysates a protein of approximately 60 kDa (lanes 13, 15) was also detected

and could be competed with peptide (lanes 16, 18). The size of CRK2, as predicted from the open reading frame, is 39 kDa, therefore the 38 kDa protein detected in all three life cycle stages lysates was a good candidate for CRK2.

'PSTAVR' antiserum was purified in a two-step process in the same manner as the 'EVREE' antiserum. Partial purification on a protein A/G column of an aliquot of 'PSTAVR' was followed by affinity purification on a Sulfolink PSTAVR peptide column. The purified PSTAVR antibodies were then tested by Western blotting against insoluble, soluble (S-100) and total protein lysates of *T. b. brucei* STIB 247 procyclic, long slender and short stumpy bloodstream forms (Fig. 5.3). A similar pattern of proteins was observed in all three life cycle stages. A protein of between 50-60 kDa was detected very strongly in the total cell extracts from the three life cycle stages (lanes 1, 7, 13) and insoluble lysates (lanes 2, 8, 14) that was competed with peptide (lanes 4, 5, 10, 11, 16, 17). Another protein was detected as a diffuse signal of approximately 65-70 kDa that was present in the total (lane 1, 7, 13) and insoluble (lane 2, 8, 14) lysates and was competed with peptide (lanes 4, 5, 10, 11, 16, 17). A protein was also detected at approximately 43 kDa in the total (lane 1) and soluble (lane 3) procyclic lysates that was nonspecific as it could not be competed with peptide (lanes 4, 6). In the long slender bloodstream form total (lane 13) and soluble (lane 15) lysates a protein of approximately 41 kDa was detected weakly and was competed with peptide (lanes 16, 18).

The results from the two purified 16-mer antibodies were largely irreconcilable; both detected a number of proteins, but the proteins recognised differed between the two antibodies. As the size of CRK2, as predicted from the gene sequence, is 39 kDa it was most likely that the 38 kDa band detected by the EVREE antibodies was CRK2. In addition, this 38 kDa protein was located in the soluble fraction lysates. The proteins recognised by PSTAVR were of the wrong size and located in the insoluble fraction lysates, with the exception of the 41 kDa protein located in the soluble lysate of the long slender bloodstream form. At this juncture, this protein could not be ruled out as CRK2.

The 11-mer peptide representing the C-terminal 11 amino acids was conjugated to KLH and used to immunise two rabbits (see chapter 2.3.4). The exsanguination antisera obtained, labelled 'CVELK' (R421) and 'CVELK' (R507), were tested in Western blots against soluble (S-100), insoluble and total protein lysates of *T. b. brucei* STIB 247 procyclic, long slender and short stumpy bloodstream forms. Neither antiserum detected any proteins in any of the lysates of the three life cycle stages. Consequently, these antisera were abandoned for use in CRK2 analysis. It was unclear why the 11-mer peptide failed to provoke a useful immune response, as an 11-mer corresponding to the C-terminal 11 amino acids of CRK1 was conjugated to KLH and used in the immunisation of two rabbits, producing two antisera that both clearly and specifically recognised the 34 kDa CRK1 (data not shown). Therefore the short size of the peptide is unlikely to have been the reason *a priori*.

5.2.2: Production of recombinant CRK2

As the approach to raising antisera using peptides deduced from the gene sequence had not produced sufficiently specific antibodies, it was decided to attempt to use bacterial expression systems to produce recombinant CRK2 for immunisation. The vector chosen was pQE60, which adds a C-terminal hexahistidine tag (Qiagen). Full-length *CRK2* was generated by PCR using primers that engineered the appropriate restriction sites for the vector's polylinker (*Nco*I at the 5' end of the amplified *CRK2*, and *Bgl*II at the 3' end) and cloned to pQE60 to generate pGL217 (construct generated by G. Smith). Following transformation to the M15[pREP4] *E. coli* cell line, pilot inductions were performed using standard protocols with varying concentrations of IPTG to induce expression. No induced protein of the expected size of 40 kDa was detected (data not shown). A second pilot followed a time-course of induction over four hours for three different cultures also failed to produce a detectable 40 kDa protein. In an attempt to determine if any significant induction of CRK2His was taking place, a small-scale pilot denaturing purification was carried out as previously described (chapter 2.3.3). However, no purified 40 kDa CRK2His was detected (not shown). The conclusion from these experiments was that CRK2His was not expressed to any significant level. Poor expression is often the result of the recombinant protein being toxic to the bacteria, but this usually results in slow growth and cell death and the induced cultures of M15[pREP4]pGL217 appeared to grow normally. As the gene was not sequenced after PCR and cloning it is possible that a premature stop codon was generated by the amplification process. To test this theory, crude whole cell lysates of uninduced and induced M15[pREP4]pGL217 were prepared and electrophoresed on a 12.5% Lammeli gel, and transblotted to nitrocellulose membrane, which was then probed with the purified EVREE and PSTAVR antibodies (Fig. 5.4). The PSTAVR antibodies detected a protein at approximately 40 kDa in both the uninduced (lane 5) and induced (lane 6) lysates that was competed with PSTAVR peptide (lanes 7, 8). There was no difference in the strength of the signal between the uninduced (lane 5) and induced (lane 6) lysates. The EVREE antibodies detected a protein of approximately 40-42 kDa in both the uninduced (lane 1) and induced (lane 2) lysates that was competed with EVREE peptide (lanes 3, 4). There was no difference in the strength of the signal between the uninduced and induced lysates. In addition, the EVREE antibodies detected a small molecular weight protein in both lysates (lanes 1, 2) that was competed with peptide (lanes 3, 4) and a protein of approximately 60 kDa was weakly detected in both lysates (lanes 1, 2) that was competed with peptide (lanes 3, 4). The conclusion was that full-length CRK2His was expressed only to a very low level, for whatever reason, that was sufficient to be detected on a Western blot but not by Coomassie staining. Construct pGL217, therefore, was not used to express and purify CRK2His. This experiment also showed that both EVREE and PSTAVR antibodies recognised CRK2, but that the EVREE antibodies had a

higher affinity for CRK2 in the context of a Western blot. As pGL217 could not be made to express significant levels of recombinant protein in *E. coli*, use of the Qiagen histidine tag system was abandoned and a new approach taken using a yeast system.

The yeast system *Pichia pastoris* was used to attempt to produce recombinant full-length CRK2. *Pichia pastoris* is a methylotrophic yeast that can use methanol as a sole carbon source, although this is not the preferred source. Methanol utilisation is based upon induction of the alcohol oxidase genes *AOX1* and *AOX2*, in a process of derepression (lack of glucose in the culture medium) and specific induction (presence of methanol). *AOX1* accounts for the vast majority of transcript produced in methanol-grown cells and this can be as much as 5% of poly[A]+ RNA, corresponding to 30% total soluble protein. The strong, highly-inducible P_{AOX1} promoter is used for heterologous expression of the gene of interest, either by gene replacement or gene insertion to one of the two yeast strains, using a vector that contains the *AOX1* promoter and 3' sequences. The particular vector used in this case was pPIC9, which contains the α-factor secretion signal immediately upstream of the polylinker. This expression system was tested because it has advantages over bacterial systems, including a greater likelihood of correct protein folding and post-translational modification, low levels of native protein secretion, and a tendency not to hyperglycosylate secreted proteins all combined with high expression of the cloned gene of interest.

A full-length copy of *CRK2* with the appropriate restriction sites for cloning to the pPIC9 vector was produced by PCR and cloned to pPIC9. Following sequencing to ensure that the gene was in-frame with respect to the α-factor secretion signal sequence, the construct was transformed to the yeast strains involved, selected, integration checked by PCR and pilot expressions carried out. Transformation, selection, PCR and pilot expressions were performed by Dr. P. Neuville. No expression of CRK2 was observed, either intracellular or secreted. For this reason this expression system was also abandoned.

5.2.3: Expression of *CRK2* in *T. b. brucei* AnTat1.1 procyclics

As heterologous expression of *CRK2* had been a failure in both *E. coli* and yeast, it was decided to use the Clayton expression vector pHD430 (Wirtz and Clayton, 1995) to express *CRK2* in *T. b. brucei* AnTat 1.1 procyclics. This should allow a determination of the biochemical identity of CRK2 in Western blots and provide a phenotype for overexpression of CRK2. The AnTat 1.1 cell line used was kindly provided by Prof. C Clayton and was termed '360'. It contained the cassette from construct pHD360 stably integrated at the repetitive αβ-tubulin locus, to give constitutive expression of the tetracycline repressor protein (Wirtz and Clayton, 1995).

The vector pHD430 contains a cassette with the luciferase (*LUC*) gene under the control of a modified PARP promoter that contains the operator sequence from the *E. coli* tetracycline resistance operon. Downstream of the luciferase gene and the PARP 3' processing elements a bleomycin (*BLE*) resistance marker, flanked by the aldolase 5' and 3' processing elements, provides selection for transgenic cells by transcriptional read-through from the *LUC* gene. The construct contains a segment of the ribosomal DNA (rDNA) intergenic spacer, with a unique *Not*I site in the middle of this region being used for linearisation of the construct for transfection. The cassette is orientated with respect to the rDNA sequence such that integration at the rDNA spacer by homologous recombination results in the cassette orientated in a reverse (nonsense) orientation with respect to the 18S rRNA locus (Wirtz and Clayton, 1995). As the rDNA spacer is transcriptionally silent and the cassette integrates in a reverse orientation, there can be no transcriptional read-through from the 18S rRNA genes. Control of expression is therefore as shown (Fig. 5.5). The pHD430 construct integrates at the rDNA locus and the *E. coli* repressor protein binds to the operator sequence in the modified promoter, preventing transcription. Addition of as little as 5 ng ml^{-1} of tetracycline results in the repressor protein becoming bound to the tetracycline, causing a conformational change that precludes binding to the operator, and allowing transcription of the *LUC* and *BLE* genes. Addition of tetracycline and phleomycin to the recovery culture medium following transfection therefore selects for transgenic cells with stably integrated pHD430 cassettes.

As there was no full-length *CRK2* available on one subclone, the full-length *CRK2* required was generated by exploiting the 200 bp overlap of pGL116 and pGL117, and the fact that there were *Acc*I sites approximately 150 bp upstream of the ATG start codon in pGL116, and at approximately 55 bp downstream of the TAA stop codon in pGL117. *Acc*I cleavage of pGL116 and pGL117 was followed by Klenow treatment to fill in the 3' recessed ends left by the enzyme cleavage. Subsequently, *Hin*dIII cleavage liberated a 0.58 kb fragment from pGL116 that contained some 5' flank sequence, the ATG start codon and 5' gene sequence to the internal *Hin*dIII site at position 675, and from pGL117 a 0.65 kb fragment that contained the 3' gene sequence from the internal *Hin*dIII site at position 675 to the TAA stop codon and a small amount of 3' flank sequence. These two contiguous fragments were cloned at the same time to the *Eco*RV site of pBluescript (pGL130; schematic, Fig. 5.1).

The *LUC* gene was excised from pHD430 by *Bam*HI/*Hin*dIII cleavage and Klenow treated to fill in the 3' recessed termini and created blunt ends. The vector was then treated with CIP to dephosphorylate the termini. Full-length *CRK2* was released from pGL130 by *Cla*I/*Sma*I cleavage, and was subsequently Klenow treated to fill in the 3' recessed end left by *Cla*I cleavage. The vector and gene were then ligated and transformants screened by restriction mapping and PCR. A clone with the *CRK2* gene ligated in the correct orientation, plasmid pGL115, was linearised for transfection by cleavage with *Not*I. The 360 cell line was transfected

with approximately 25 μg of the purified linearised construct (for full details see chapter 2.1.2) and recovered in 10 ml of complete SDM-79 medium with 10 μg ml^{-1} phleomycin and 10 ng ml^{-1} tetracycline, to give the nonclonal population WUMP 827. The concentration of tetracycline used was sufficient to induce maximal expression of the cloned *CRK2* gene from the pHD430 cassette. Once the transfected culture had grown to midlog (approximately 6x 10^6 cell ml^{-1}), a total protein lysate was made and compared on Western blots with wild-type STIB 247 procyclic total protein lysate made from a midlog culture. Purified EVREE antibodies and nonpurified EVREE antiserum were tested first (Fig. 5.6). The crude EVREE antiserum detected a 25 kDa protein in the STIB 247 lysate (lane 1), a protein of the predicted size of 38-39 kDa (lane 1) as well as several other proteins. All of these proteins except the high and low molecular weight proteins were competed with EVREE peptide (lane 3). A similar pattern of proteins was detected for WUMP 827 (lane 2). The 38-39 kDa protein (lane 2) that was competed with peptide (lane 4) gave a much stronger signal than from STIB 247 (lane 1), showing overexpression of CRK2 in the transgenic cell population. The equivalence in signal from the 25 kDa protein between the two lysates shows equivalence of loading of protein between the lysates, therefore providing an additional control validating the difference in signal strength for the 38-39 kDa protein. The purified EVREE antibodies also gave a far stronger signal for the 38-39 kDa protein (lane 5) that was competed with peptide (lane 6). This was the expected result as maximal expression from the pHD430 cassette has resulted in significant overexpression of CRK2. Therefore, the EVREE antiserum and purified antibodies recognise endogenous CRK2 in wild-type cells and in cells overexpressing CRK2 to abnormal cellular levels (WUMP 827). Purified PSTAVR antibodies gave the same result, but the signal was much weaker (data not shown). As well as confirming the biochemical identity of CRK2 by Western blotting this experiment also provided the preliminary indications that large overexpression of CRK2 was neither deleterious to the cell cycle progression of *T. brucei* procyclics nor did it accelerate progress through the cell cycle. Wild-type STIB 247 and WUMP 827 both appeared by visual inspection to have similar cell morphologies and growth rates, although detailed analyses were not performed.

5.2.4: Immunoprecipitation with anti-CRK2 antisera

The failure of the peptide approach to produce specific antisera and the failure of the expression systems to produce recombinant protein meant that immunoprecipitations to profile CRK2 expression and activity would be difficult to perform and interpret. A pilot immunoprecipitation-linked kinase assay, however, was carried out from a STIB 247 procyclic soluble (S-100) lysate using both purified EVREE and PSTAVR antibodies, and the 11-mer

polyclonal CVELK (R421) antiserum, with histone HI as the substrate (Fig. 5.7). For the purified EVREE antibodies, little difference was observed between the preimmune serum (lane 1) and the purified antibodies (lane 2). The signal could not be competed with peptide (lane 3). The purified PSTAVR antibodies (lane 5) gave a slightly stronger signal than the preimmune serum (lane 4), and this appeared to be slightly competed with peptide (lane 6) but the difference was marginal. The preimmune serum (lane 7) and immune (lane 8) CVELK (R421) antiserum gave equivalent signals, with some competition with peptide observed (lane 9); however, as there was no difference between the preimmune serum and immune CVELK (R421) antiserum this was not a specific effect. These results confirmed the unsuitability of all the antisera for analysis to detect CRK2 histone H1 kinase activity.

5.2.5: Targeted disruption of *CRK2*

To investigate the role of the gene *in vivo* a regimen of targeted gene disruption was followed, to determine if the gene was essential and to analyse any resultant phenotype on deleting both chromosomal copies of *CRK2*. Three constructs were developed for this, identical except for the resistance markers employed. For this analysis only the procyclic form was used.

Primers CRK2N5' and CRK2N3' were designed for use in PCR to amplify the 5' end of the gene as a fragment of approximately 490 bp, corresponding to approximately 250 bp coding sequence and 240 bp flanking (Fig. 5.1). Primers CRK2C5' and CRK2C3' were designed to for use in PCR to amplify the 3' end of the gene as a fragment of approximately 525 bp, corresponding to approximately 300 bp of coding sequence and 225 bp of flanking sequence (Fig. 5.1). The primers engineered the following restriction sites on to the ends of the amplification products: CRK2N5' - *Sph*I; CRK2N3' - *Xba*I; CRK2C5' - *Kpn*I; CRK2C3' - *Eco*RI. PCR was performed using these two primer combinations on pGL116 (CRK2N5' and CRK2N3') and pGL117 (CRK2C5' and CRK2C3'). The PCR amplification products were digested after amplification with the relevant enzyme pair, electrophoresed on 1% TAE agarose gels, excised and gel purified using either the SpinX minicolumns (Costar) or the QIA Gel Extraction kit (Qiagen) before being cloned to the disruption constructs.

The starting construct was based on the vector pJB44 (Sherman *et al.*, 1991) and consisted of the *BLE* gene cloned into pBluescript and flanked by the PARP promoter, splice acceptor and 3' expression sequences. The whole cassette was excised by *Bam*HI-*Kpn*I cleavage and cloned into pUC18 vector. The PCR products were digested with the enzymes corresponding to the engineered sites and cloned into the relevant sites in the pUC18/PARP-*BLE* construct polylinker. The complete cassette was then shuttled into pBluescript in two stages, firstly the *Hin*dIII-*Eco*RI fragment (containing the BLE gene, the PARP 3' expression

sequences and the *tbCRK2* 3' sequence), then the *Hin*dIII-*Hin*dIII fragment (containing the *tbCRK2* 5' sequence and the PARP promoter), to generate pGL108 (Fig. 5.8). Excision of the cassette for transfection was effected by *Sac*I-*Xho*I cleavage.

The puromycin acetyltransferase (*PAC*) gene was excised from a construct designed for use in *Leishmania mexicana* by *Bam*HI/*Spe*I cleavage. Klenow treatment of the gene generated blunt ends and the gene was then cloned to the *Eco*RV site of pBluescript (SK-). Positive clones were screened by PCR for those that contained the *PAC* gene orientated such that the 5' end lay adjacent to the *Hin*dIII site in the vector's polylinker. The *Eco*RI site lying between the polylinker *Pst*I site and the 3' end of the *PAC* gene was then destroyed by cutting with *Eco*RI, Klenow treating the termini and religating. The gene was then excised by *Hin*dIII/*Pst*I cleavage and cloned to the pB/PARP-*BLE* starting construct in place of the *BLE* gene. The *CRK2*-specific sequences were cloned to flank the PARP-*PAC* cassette and the finished construct shuttled into pBluescript in two stages, exactly as for pGL108, to give pGL110 (Fig. 5.8).

The streptathricin acetyltransferase (*SAT*) gene was excised from a construct designed for use in *Leishmania mexicana* by *Bam*HI/*Spe*I cleavage and cloned in precisely the same way as the *PAC* cassette pGL110 to give pGL109 (Fig. 5.8).

Primary transfection was carried out as described (chapter 2.1.2) to STIB 247 wild-type procyclics using the *Sac*I/*Xho*I insert from pGL108, and selected using phleomycin at 20 μg ml^{-1} in liquid culture. Two nonclonal populations of cells resistant to the antibiotic grew up, WUMP 563 and WUMP 564. Genomic DNA was isolated from both cell populations and a Southern blot was performed to assess integration of the *BLE* resistance cassette (Fig. 5.9). Genomic DNA was digested with *Bam*HI, as the position of this site around the *CRK2* locus had been mapped. The blot was probed with full-length *CRK2*, released from pGL130 by *Xba*I/*Xho*I cleavage. Given the position of two *Bam*HI sites within the gene and a further *Bam*HI site 3' to the gene, wild-type genomic DNA was expected to give two hybridising fragments of approximately 0.28 kb and 1.8 kb. The *BLE* resistance cassette contains two *Bam*HI sites and so hybridising fragments of approximately 0.22 kb, 1.8 kb and 2.45 kb were expected. The smallest fragment ran off the end of the gel, but the two larger fragments were clearly present in both WUMP 563 (lane 3) and WUMP 564 (lane 2), with a diminished signal from the 1.8 kb fragment as compared to wild-type consistent with replacement of one of the wild-type alleles. Wild-type STIB 247 (lane 4) and wild-type EATRO 795 (lane 1) genomic DNA showed only the 1.8 kb wild-type hybridising fragment. Both WUMP 563 and WUMP 564 had a normal morphology and growth rates equivalent to wild-types. FACs analysis comparing both these nonclonal cell populations with wild-type STIB 247 showed no difference in DNA content (data not shown).

WUMP 563 and WUMP 564 were transfected with the *Sac*I/*Xho*I insert from pGL110 to attempt to generate null mutants, and selected with puromycin and phleomycin. WUMP 563

was also transfected with the *SacI/XhoI* insert from pGL110 and selected with puromycin alone. Three healthy populations of cells grew, genomic DNA was isolated from each as well as from WUMP 563, and PCR was performed using primers CRK2N5' and CRK2C3' to assess the integration of the resistance cassettes, with predicted PCR product fragments of approximately 1.5 kb for the wild-type gene, approximately 2.1 kb for the *BLE* resistance cassette and 2.3 kb for the *PAC* resistance cassette (Fig. 5.10). Wild-type STIB 247 DNA showed only a 1.5 kb fragment representing wild-type allele (lane 5). WUMP 563 showed the 1.5 kb wild-type allele amplification product and a fragment at approximately 2.1 kb, the predicted size of the amplification product from the *BLE* resistance cassette integrated at the *CRK2* locus (lane 4). WUMP 563 transfected and selected with puromycin only showed the wild-type 1.5 kb and a band of approximately 2.3 kb, the predicted size of the amplification product from the *PAC* resistance cassette integrated at the *CRK2* locus (lane 2). Given the selection applied this was consistent with replacement of the *BLE* resistance cassette with the *PAC* resistance cassette. WUMP 563 and WUMP 564 transfected and selected with phleomycin and puromycin showed the 2.1 kb *BLE* resistance cassette and the 2.3 kb *PAC* resistance cassette, but also the 1.5 kb wild-type allele amplification product. As *CRK2* is a single-copy gene the simplest explanation of this anomalous result was that there had been a change in ploidy as a result of the transfection with the *SacI/XhoI* insert from pGL110, possibly due to cell-cell fusion promoted by electroporation. In *Leishmania*, ploidy changes resulting from targeted disruption of an essential gene have been observed (Cruz *et al.*, 1993). FACs analysis was therefore carried out on the two nonclonal *Δcrk2::BLE/Δcrk2::PAC/CRK2* populations. This showed no change in cell ploidy between the STIB 247 parent wild-type isolate and the *Δcrk2::BLE/Δcrk2::PAC/CRK2* populations (data not shown); the possibility that the presence of the wild-type allele was due to a change in ploidy (as was seen for the *lmmCRK1* gene; Mottram *et al.*, 1996) could therefore be discounted, although duplication of a single chromosome could not be discounted. Initial Southern blotting of these two nonclonal *Δcrk2::BLE/Δcrk2::PAC/CRK2* populations genomic DNA gave poor results, but seemed to show that the wild-type allele present in the correct locus and that both the *BLE* resistance and *PAC* resistance cassettes were present at the *CRK2* locus (data not shown). However, as these populations were nonclonal this analysis was discontinued pending the establishment of clonal isolates. Transfection of the *PAC* resistance cassette to WUMP 564 was repeated, and the transfected culture cloned as soon as was feasible by the semisolid plate method. As more than one specific type of recombination event may have been taking place, it was possible that any null mutants being generated could have been outgrown by incorrect integrants, especially if *CRK2* was involved in some aspect of cell cycle progression and deletion of the gene compromised the cell cycle. Therefore immediate cloning following transfection would result in emergent clones accurately reflecting the genetic arrangement at the *CRK2* locus as a result of secondary transfection with the *PAC* resistance cassette. Cloning by

the semisolid plate method was found to be very inefficient for the procyclic form of *T. brucei*, and several attempts were required to isolate four putative *Δcrk2::BLE/Δcrk2::PAC* clones, WUMP 837, WUMP 838, WUMP 839 and WUMP 840. In retrospect, a limiting dilution method would have been a better approach to cloning transfectants. Genomic DNA was prepared from the four clones and used in PCR with primers CRK2N5' and CRK2C3' to assess the genetic organisation at the CRK2 locus (Fig. 5.11). The control showed the 1.5 kb wild-type allele amplification product (lane 1). WUMP 837, WUMP 839 and WUMP 840 showed the 2.1 kb *BLE* resistance cassette and the 2.3 kb *PAC* resistance cassette, and also the 1.5 kb wild-type allele band (lanes 5, 3, 2). WUMP 838 showed only the 1.5 kb wild-type allele fragment and the 2.1 kb *BLE* resistance cassette, but not the 2.3 kb *PAC* resistance cassette (lane 4). To corroborate this data a Southern blot was performed (Fig. 5.12). Genomic DNA was digested with *Bam*HI and probed with full-length *CRK2* released from pGL130 by *XbaI/XhoI* cleavage (Fig. 5.8). The STIB 247 wild-type control showed only the 1.8 kb fragment corresponding to the wild-type *CRK2* allele (lane 2). WUMP 838 (lane 5) showed the 1.8 kb fragment corresponding to the wild-type *CRK2* allele and a hybridising fragment at approximately 2.4 kb representing the *BLE* resistance cassette integrated at the *CRK2* locus. WUMP 837 (lane 6), WUMP 839 (lane 4) and WUMP 840 (lane 3) showed the 1.8 kb wild-type allele fragment, as well as a fragment at approximately 2.4 kb representing the *BLE* resistance cassette integrated at the *CRK2* locus, and a fragment at approximately 2.6 kb representing the *PAC* resistance cassette integrated at the *CRK2* locus. The supposition for WUMP 838 was that nuclease activity had degraded the *CRK2* sequences at the termini of the *PAC* resistance cassette, thus facilitating integration at an erroneous locus, most likely the PARP locus due to the PARP-specific 5' and 3' processing elements flanking the *PAC* gene. This type of recombination event has been observed to occur in transfection experiments, although it is a low-frequency event (Blundell *et al.*, 1996). The data for WUMP 837, WUMP 839 and WUMP 840 showed three genetic elements apparently occupying only two physically discrete, though identical, homologous loci. In *Leishmania* species, strategies to avoid the deletion of essential genes include changes in ploidy (Cruz *et al.*, 1993) and the translocation of tens of kilobases of DNA bounding the locus in question to elsewhere on the same chromosome, or even on another nonhomologous chromosome (Dumas *et al.*, 1997). Although this kind of genetic event has not been demonstrated in trypanosomes, to investigate the possibility that a similar event had occurred, Southern blotting of WUMP 837, WUMP 839 and WUMP 840 was performed, using a variety of enzymes to digest the genomic DNA and probed with full-length *CRK2* released from pGL130 by *XbaI/XhoI* cleavage. No differences between the three clones and STIB 247 wild-type genomic DNA was observed (data not shown) and further analysis was not performed. FACs analysis of the four clones as compared to STIB 247 wild-type showed no differences.

Another possibility regarding the curious genetic arrangement at the *CRK2* locus in

WUMP 837, WUMP 839 and WUMP 840 was that STIB 247 is trisomic at the chromosome containing *CRK2*. Such partial trisomy would give the observed results by PCR and Southern blotting, and would not show on FACs analysis or Southern blotting. To test this possibility, WUMP 837 was transfected with the *SAT* resistance cassette (insert from pGL109). The transfection culture was split between five flasks and selected with 20 μg ml^{-1} phleomycin, 10 μg ml^{-1} puromycin and 40 μg ml^{-1} nourseothricin, and each flask cloned by the semisolid plate method. Only one clone was isolated, WUMP 877, genomic DNA was prepared and PCR reactions using CRK2N5' and CRK2C3' was performed. Although the results were difficult to interpret (data not shown) and varied between reactions, the 1.5 kb wild-type allele was always shown to be present. The implications are that regardless of the mechanisms involved, repeated transfection to attempt to delete wild-type *CRK2* was not successful and that *T. brucei* retains a wild-type copy of the gene.

5.2.6: Targeted disruption of *CRK2* in AnTat 1.1 procyclics expressing *CRK2*

The *in vivo* function of *CRK2* was addressed by utilisation of the tetracycline-inducible system developed by Wirtz and Clayton (1995). Deletion of *CRK2* had been shown to be not possible (section 5.2.5), therefore deletion of wild-type CRK2 alleles in a background of inducible expressed CRK2 should be possible and provide a tool for subsequent manipulation of cellular levels of CRK2, leading to analysis of phenotype on downregulation or overexpression of *CRK2* as compared to wild-type.

AnTat 1.1 '360' cell line procyclics were cotransfected with pGL115 (the inducible *CRK2* expression construct, see section 5.2.3) and the *SAT* resistance cassette released from pGL109 by digestion with *Sac*I/*Xho*I (Fig. 5.8). Approximately 50 μg of each transfection construct was used, and under selection with tetracycline, phleomycin and nourseothricin a population of resistant cells was obtained. These were cloned by the semisolid plate method as described in chapter 2.1.2, and two clones isolated, WUMP 835 and WUMP 836. Each clone was transfected with approximately 50 μg of the *PAC* resistance cassette (insert from pGL110; Fig. 5.8), and the transfected cultures cloned by the semisolid method. Four clones were isolated, genomic DNA prepared and PCR using primers CRK2N5' and CRK2C3' performed to assess the state of the *CRK2* locus. Neither of the resistance cassettes were detected in any of the clones, only wild-type *CRK2* (data not shown). Southern blot analysis of the WUMP 835 and WUMP 836 (which have the predicted genotype *Δcrk2::SAT/CRK2*/pHD430-*CRK2*) showed the characteristic hybridising fragments of the wild-type allele (Fig. 5.13, lane 2) but not the expected fragment of approximately 2.6 kb corresponding to integration of the *SAT* cassette. Instead, they both showed an identical pattern of hybridising fragments: 1.6 kb, 0.9 kb and a

weakly hybridising fragment at approximately 5.0 kb (lanes 3 and 4). PCR on genomic DNA of WUMP 835 and WUMP 836 with primer combination SAT5' and SAT3' demonstrated the presence in both of the *SAT* gene, but primer combinations CRK2N5' and SAT3', or CRK2C3' and SAT5', did not give any reaction (data not shown), thus confirming that the *SAT* cassette had not integrated correctly at the *CRK2* locus in either clone. Protein samples were prepared from both clones as well as wild-type STIB 247 and equivalent loadings electrophoresed on a gel that was blotted and probed with protein A/G partially purified anti-EVREE antibodies (Fig. 5.14). CRK2 was detected in the wild-type sample (lane 1) and at elevated levels in the samples of WUMP 835 and WUMP 836 (lanes 2 and 3). In all three lanes CRK2 was competed with peptide (lanes 4-6); therefore both WUMP 835 and WUMP 836 displayed ectopic expression of *CRK2* from the rRNA locus.

5.2.7: Expression *in vivo* of histidine-tagged CRK2

Histidine-tagged CRK2 was generated by digestion of pGL217 with *Hin*dIII to release the histidine-tagged 3' end of the gene on a fragment of approximately 630 bp, which was cloned into the *Hin*dIII site in pBluescript. Screening by PCR identified a clone with the fragment orientated with the tag adjacent to the *Xba*I site in the vector's polylinker. This intermediary pB/630 plasmid was then digested with *Sph*I and *Xba*I, releasing the 3' end of the gene and the tag sequence on a small fragment of approximately 150 bp. The 150 bp fragment was gel purified and ligated with pGL130, which had been digested with *Sph*I and *Xba*I and gel purified. Transformants were screened by PCR and one positive - which represented reformed, full-length CRK2 with a hexahistidine C-terminal tag (pGL233) - was sequenced over the last 200 bp at the 3' end. This was because the copy of CRK2 in pGL217 was generated by PCR and had not been sequenced subsequently, therefore the portion of sequence of the pGL233 insert that was derived from pGL217 had to be sequenced to ensure that there had been no PCR-generated mutations. Once the sequence was confirmed, *CRK2His* was released from pGL233 by digestion with *Bsp*1201 and *Not*I, and cloned into the *Bsp*1201 site in the polylinker of expression plasmid pHD675 (Biebinger *et al.*, 1997) to give pGL234. Construct pGL234 was linearised by digestion with *Not*I, purified and transfected to the STIB 427 MiTat 1.1 strain containing the stably integrated pHD449 ('449' cell line) as described (chapter 2.1.2), with selection under phleomycin and hygromycin.

A nonclonal population of doubly-resistant cells were recovered (WUMP 1163) and tested for induction of CRK2His with 0.2 ng ml^{-1}, 1.0 ng ml^{-1} and 10.0 ng ml^{-1} of tetracycline. 10 mls of each induced culture were harvested, resuspended in LSGI and S-100 extracts made. Specific selection of CRK2His was as described (Fig. 5.15) and Western blotting performed

using purified EVREE antibodies. The hexahistidine tag binds efficiently to the nickel ions in the charged NTA resin; minimal washing removes most contaminating endogenous prokaryotic proteins. Relatively little induction of CRK2His was seen with 0.2 ng ml^{-1} and 1.0 ng ml^{-1} (lanes 1 and 2) as compared to the induction that was observed with 10.0 ng ml^{-1} tetracycline (lane 3). A very weak signal for CRK2 was seen for the column flow-through of all three (lanes 4, 5, 6) and was presumably endogenous non-tagged CRK2. Significant CRK2His was found to have bound the Ni-NTA resin from the extract induced with 10.0 ng ml^{-1} (lane9), in contrast to the extracts induced with 1.0 ng ml^{-1} (lane 8) and 0.2 ng ml^{-1} (lane 7). Preblocking of the antibodies with EVREE peptide partially abrogated the signal from CRK2His but did not abolish it entirely (lanes 11-19; compare lanes 1-3 with 11-13, and lane 9 with lane 19). The multiplicity of bands in lanes 1-3 and 11-13 are a result of the primary antibody being used at too high a concentration.

As a follow-up experiment, S-100 extracts were made from midlog STIB 247, WUMP 1163 grown in the absence of tetracycline for several subpassages, and WUMP 1163 grown in 10 ng ml^{-1} tetracycline. CRK2His was selected on Ni-NTA beads and aliquots used for protein samples and kinase assay (Fig. 5.16). The Western blot (Fig. 5.16A) showed that nothing was specifically purified from STIB 247 wild-type (Fig. 5.16A, lane 1) and that a small but significant amount of CRK2His was purified from WUMP 1163 grown in the absence of tetracycline (Fig. 5.16A, lane 2). A very strong signal was obtained from the WUMP 1163 eluate grown in the presence of 10.0 ng ml^{-1} tetracycline (Fig. 5.16A, lane 3), indicating that relatively large amounts of CRK2His had been selected on the Ni-NTA. The histone H1 kinase assay, however (Fig. 5.16B) showed little difference between the three samples. A slightly stronger signal was obtained from WUMP 1163 grown in the presence of 10.0 ng ml^{-1} tetracycline (Fig. 5.16B, lane 3) but a strong signal was obtained from 247 wild-type (Fig. 5.16B, lane 1) indicating that background in this experiment was too high for the results to be significant.

5.3: Discussion

The initial attempts to raise specific antisera to CRK2 by immunising rabbits with short peptides deduced from the gene sequence proved more complicated than anticipated. The 11-mer corresponding to the C-terminal 11 amino acids (CVELK peptide) produced an antiserum (CVELK antiserum) that detected no proteins by Western blotting in trypanosome lysates. The 16-mer corresponding to the C-terminal 16 amino acids (EVREE peptide) raised an antiserum (EVREE antiserum) that was affinity purified and used to detect a number of proteins by Western blotting in the soluble fraction of lysates from procyclic, long slender and short stumpy

bloodstream forms (Fig. 5.2). The 16-mer corresponding to the CRK2 PSTAIRE box homologue (PSTAVR peptide) raised an antiserum (PSTAVR antiserum) that was also affinity purified and used to detect two proteins that gave strong signals by Western blotting in the insoluble fraction lysates from procyclic, long slender and short stumpy bloodstream forms (Fig. 5.3). In addition, a protein was detected in the procyclic soluble fraction lysate and another in the long slender bloodstream form soluble fraction lysate. The size of CRK2, as deduced from the open reading frame, is predicted to be approximately 39 kDa. Therefore although the 38 kDa protein detected by the EVREE antibodies was a good candidate for CRK2, the 41 kDa protein detected only in the long slender bloodstream form by the PSTAVR antibodies could not be discounted.

Further evidence that both sets of antibodies were capable of detecting CRK2 on a Western blot came from attempts to express CRK2 in *E. coli* as a hexahistidine-tagged fusion protein. Pilot inductions and purifications suggested that CRK2His was not expressed. However, a Western blot of induced and uninduced M15[pREP4]pGL217 showed a protein of about 40 kDa detected by both EVREE and PSTAVR antibodies, in both the induced and uninduced lysates, that was competed with peptide (Fig. 5.4). This suggested that transcriptional control of pGL217 in *E. coli* was slightly leaky, leading to a basal level of CRK2His expression, but that relief of transcriptional repression did not result in any increase in CRK2His expression. The reason for this is unclear as M15[pREP4]pGL217 *E. coli* showed no difference in growth kinetics in the presence or absence of the inducer IPTG, as would be expected if the fusion protein was highly toxic to the cells. Whatever the reason, it was clear that the pQE system was not a viable approach to the expression of CRK2 as a fusion protein. However, these experiments did demonstrate that both EVREE and PSTAVR antibodies could detect CRK2 on a Western blot, thus confusing the issue of why the two sets of antibodies did not detect a protein in common in *T. brucei* protein lysates. As the EVREE antibodies gave a stronger signal (Fig. 5.4) it was possible that these antibodies have a higher affinity for CRK2; thus if CRK2 is present at very low cellular levels in *T. brucei* it is possible that EVREE antibodies can detect it but PSTAVR antibodies cannot. The *P. pastoris* yeast system was used for the attempted expression of CRK2, on the basis that yeast can sometimes cope better with ectopic expression of eukaryotic proteins than prokaryotes, with a greater chance of correct folding of the expressed protein and a lower chance of hyperglycosylation. However, no detectable expression of CRK2 was found and the yeast expression studies were abandoned.

Conclusive proof for the biochemical identity of CRK2 came from the expression of a cloned copy of *CRK2* in *T. brucei* using the pHD430 vector system (chapter 5.2.3). EVREE antibodies detected specifically a protein of about 38-39 kDa in a total protein lysate of STIB 247 wild-type that was more abundant in a WUMP 827 total protein lysate (Fig. 5.6). WUMP 827 had been grown in 10 ng ml^{-1} tetracycline before being harvested to make the lysate. This level of tetracycline was sufficient to have induced maximal expression of the cloned copy of

CRK2, therefore the difference in signal strength for the same protein between the two cell lines (coupled with the fact that none of the other proteins detected differed in signal strength between the two) demonstrated that the 38-39 kDa protein initially detected by the purified EVREE antibodies (Fig. 5.2) was CRK2. Overexpression of *CRK2* did not interfere with normal cell cycle progression, as WUMP 827 appeared to have the same morphology and growth kinetics as STIB 247 wild-type, thus cell cycle progression was neither accelerated nor retarded.

The purified EVREE and PSTAVR antibodies, as well as one of the nonpurified 11-mer anti-CRK2 antisera (R421) were tested in a pilot immunoprecipitation-linked kinase assay (chapter 5.2.4; Fig. 5.7). However, no specifically precipitated histone HI kinase activity was detected. There are a number of possible interpretations for this. Firstly, it is possible that although the antibodies reacted to denatured CRK2 on a Western blot they were not capable of recognising native protein, and thus did not precipitate native CRK2. Secondly, the PSTAIRE box of CDKs is involved in cyclin binding (Ducommun *et al.*, 1991; Endicott *et al.*, 1994; Jeffrey *et al.*, 1995). Therefore the PSTAVR box is likely to be involved in cyclin binding, so it is possible that the PSTAVR box was occluded by a cyclin partner and thus the PSTAVR antibodies could not precipitate CRK2. Likewise, the C-terminus of CRK2 may not be available for binding in the native protein, therefore the EVREE antibodies and CVELK (R421) antiserum could not precipitate CRK2. A third possibility regards the substrate specificities of CDKs.

The 'consensus sequence' for CDK phosphorylation (S/T-P-X-K/R) was largely determined by comparison of the sites phosphorylated *in vitro* by CDK1 on various candidate substrates (Nigg, 1993). There is no shortage of such candidates, including nuclear lamins, histone H1 and H3, some transcription factors, non-receptor tyrosine kinases and structural proteins (Nigg, 1992; Bradbury, 1992; Nigg, 1993). The initial assumption was that this 'consensus sequence' would hold for all CDKs. However, this is not necessarily the case, for the following reasons.

1) Some CDK1 substrates can be phosphorylated *in vitro* on the same residues by other kinases, like MAPKs (Litchfield *et al.*, 1991; Drewes *et al.*, 1992; Peter *et al.*, 1992). Therefore, it is difficult to tell which might be the relevant kinase *in vivo*.
2) Activity is modified by cofactor (cyclin) binding. For example, CDK1-cyclin B and CDK1-cyclin A phosphorylate different sites on histone H1; CDK2-cyclin A but not CDK2-cyclin B can phosphorylate the retinoblastoma-related protein p107; CDK2-cyclin A but not CDK2-cyclin E can phosphorylate DP1 (Schulman *et al.*, 1998). Although cyclin binding to CDKs modifies the active site it does not remodel it for a particular substrate (Homes and Solomon, 1996). Rather, the cyclin partner seems to direct CDK activity by binding substrates, for example, cyclin D binds pRB directly and so directs CDK4 activity to pRB (Sherr, 1993). The cyclin D motif

involved in binding is conserved in cyclin E, presumably directing CDK2 activity to pRB.

3) There is a temporal aspect involved. For example, CDK1-cyclin B phosphorylates pRB very efficiently (Lin *et al.*, 1991; Lees *et al.*, 1991) but cyclin B is not expressed until late G1 and does not begin to associate with CDK1 until S phase, therefore pRB is not a relevant substrate *in vivo*. Conversely, CDK4/6 are active in mid-G1 and phosphorylate pRB, but cannot efficiently phosphorylate histone H1 (Matsushime *et al.*, 1994; Ewen *et al.*, 1993; Kato *et al.*, 1993), a protein which is regarded as the 'canonical' CDK substrate.
4) Sequestration may influence substrate specificity. The CDK7-cyclin H dimer or 'free' CDK7-cyclin H-MAT1 trimer has a strong preference for CDK2 as a substrate, but if recruited (via MAT1) to the basal transcription factor TFIIH it has a strong preference for the C-terminal domain of RNA pol II and CDK2 is no longer a substrate (Rossignol *et al.*, 1997).

If any of the antibodies did precipitate CRK2, therefore, failure to phosphorylate histone H1 could be attributable to several different reasons. It could be that histone H1 is simply not a substrate, *in vivo* or *in vitro*, for CRK2, in the same manner that histone H1 is not a substrate for CDK4/6. A second possibility is that CRK2 is capable of phosphorylating histone H1 in other life cycle stages, but that in procyclics it is inactive, or possibly complexed to a cyclin that directs its activity to other substrates. The 'consensus sequences' for CDK1 phosphorylation may be divergent between organisms separated by large phylogenetic distances, so mammalian histone H1 is not a CRK substrate. Alternatively, it is possible that in procyclics CRK2 is recruited to a complex that modifies its activity, by analogy to CDK7 recruitment to TFIIH.

Targeted disruption of the *CRK2* gene was carried out in the hope of isolating null mutants with a phenotype that would provide information on the function of CRK2 in the cell cycle of *T. brucei*. Targeted gene disruption to identify 'essential' genes i.e. ones that cannot be deleted, has been a popular approach to the identification of potential drug targets in parasites, particularly considering the ease with which this can be done with procyclic and bloodstream form trypanosomes. However, trypanosomatid genetic plasticity (considered a criterion for 'essential' genes; Cruz *et al.*, 1993; Mottram *et al.*, 1996; Dumas *et al.*, 1997) has meant that characterisation of essential genes has proven difficult, and it is important to remember that failure to delete a gene means only that it serves an essential function in the particular life cycle stage studied under the culture conditions used. In addition, as the different life cycle stages require specific temperatures for growth and survival, isolation of temperature-sensitive conditional mutants has proven impossible. Moreover, attention has recently been drawn to the shortcomings of targeted gene disruption as a validation of viable drug targets (Barrett *et al.*, 1999). Genes that cannot be deleted do not always represent valid drug targets (Eisenthal and

Cornish-Bowden, 1998), and conversely some genes that can be deleted - at least in some life cycle stages - have proven to be valid drug targets (Li *et al.*, 1996; Li *et al.*, 1998; Wiese, 1998). Trypanosome CRKs are likely to be valid targets as it has not proven possible to isolate null mutants for the *Leishmania CRK1* and *CRK3* (Mottram *et al.*, 1996; Hassan, 1999). Furthermore, if CRKs are functional CDK homologues involved in cell cycle progression, they are likely to be essential in many of the proliferative stages of the trypanosomatid parasites, and therefore to fulfil one important criterion for a possible drug target.

Primary transfection of STIB 247 wild-type procyclics was successful in disruption of one of the alleles of CRK2 (Fig. 5.9), showing that the *BLE* resistance cassette integrated correctly. Transfection of the single allele *BLE* replacement (WUMP 563) with the *PAC* cassette and selected with puromycin alone showed replacement of the *BLE* cassette with the *PAC* cassette (Fig. 5.10, lane 2), showing that the *PAC* cassette functions. Under the selection conditions this is consistent with the preservation of a wild-type allele that has an essential function in procyclics. There would be no selective pressure for maintenance of a wild-type copy of *CRK2* if it were nonessential in procyclics, therefore if integration was random some null mutants should have been obtained. Secondary transfection of a single allele *BLE* replacement (WUMP 564) with the *PAC* cassette and selection with both phleomycin and puromycin yielded four clones, WUMP 837-840. Analysis of the CRK2 locus by PCR (Fig. 5.11) and Southern blotting (Fig. 5.12) showed that at least two different events had occurred. WUMP 838 was demonstrated to possess the wild-type allele and the *BLE* cassette at the correct locus, and the *PAC* cassette not present at the *CRK2* locus. An explanation for this result is that nuclease activity had degraded the *CRK2*-specific termini of the *PAC* cassette, permitting integration at the PARP locus due to the PARP-specific 5' and 3' elements. This type of event has been demonstrated to occur at a low frequency (Blundell *et al.*, 1996). WUMP 837, WUMP 839 and WUMP 840 all showed the same genotype; the presence of a wild-type allele as well as both the *BLE* and *PAC* cassettes correctly integrated at the *CRK2* locus. There were several possibilities that could explain this curious genetic arrangement. Firstly, that there had been a change in DNA content of these clones, possibly resulting in triploid or tetraploid cells. Triploid cells are frequently found after mating and are viable (Tait and Turner, 1990). In *Leishmania* species, triploid or tetraploid cells are frequently found after attempts to delete an essential gene (Cruz *et al.*, 1993; Mottram *et al.*, 1996; Hassan, 1999). Indeed, chromosome plasticity in Leishmania following attempted gene disruption has been proposed as a marker of essential gene function (Cruz *et al.*, 1993; Mottram *et al.*, 1996; Dumas *et al.*, 1997). However, FACs analysis of WUMP 837, WUMP 839 and WUMP 840 disproved this as all three clones were found to have the same DNA content as wild-type STIB 247. A second possibility was that STIB 247 is trisomic at the chromosome containing *CRK2*. This was judged to be unlikely, as CRK2 has been used as a syntenic marker for molecular karyotyping of *T. brucei* (Melville *et*

al., 1998). Alternatively, these clones may have become trisomic following transfection. To address these possibilities WUMP 837 was transfected with the *SAT* resistance cassette and cloned. PCR analysis of the single clone isolated (WUMP 877) demonstrated wild-type allele to be present, although the *SAT* resistance cassette was not detected at the *CRK2* locus (data not shown). If WUMP 837 was or had become trisomic, WUMP 877 may be an example of the low-frequency event of incorrect integration due to the processing elements in the construct (Blundell *et al.*, 1996). In this case, the culture conditions would have selected only for those clones in which the *SAT* cassette had integrated incorrectly. A third possibility was that a large section of DNA around the *CRK2* locus had been transposed to another chromosome; this has been demonstrated to occur in Leishmania (Dumas *et al.*, 1997) but has not yet been proven to occur in *T. brucei*. Regardless of the mechanisms involved, deletion of *CRK2* to generate nulls was not possible, which may suggest that *CRK2* is an essential gene. This is the case for procyclics, however it may also be the case for the bloodstream forms too, as Western blotting has shown that CRK2 is present in the long slender and short stumpy bloodstream forms (Fig. 5.2). However, as transfection of the long slender bloodstream form was not performed, the possibility remains that there is stage-specific expression of another *CRK* or other kinase that can replace the loss of CRK2 in these life cycle stages.

To investigate the possibility that a strict gene dosage requirement operates on *CRK2*, and to demonstrate the specificity of the effect for the failure of targeted disruption to delete *CRK2*, attempts were made to delete the wild-type alleles in a background of ectopic expression of *CRK2* (chapter 5.2.6). The tetracycline-inducible system was used (Wirtz and Clayton, 1995). The particular vector used, pHD430, was reported to achieve induction over four levels of magnitude within 6 hours of induction with 50 ng ml^{-1} tetracycline, and that intermediate levels of induction were achieved with lower concentrations of tetracycline (Wirtz and Clayton, 1995). Cotransfection of the AnTat 1.1 '360' cell line with the *SAT* cassette and pGL115 (see section 5.2.3) followed by immediate cloning yielded two independent clones, WUMP 835 and WUMP 836, which were transfected with the *PAC* cassette. Four clones were isolated; however, PCR demonstrated the only the wild-type allele was present at the *CRK2* locus. Reinvestigation of WUMP 835 and WUMP 836 by Southern blotting demonstrated the presence of the wild-type allele and confirmed that in neither clone had the *SAT* cassette integrated correctly at the *CRK2* locus (Fig. 5.13), although PCR showed the *SAT* gene to be present in both clones, and both were resistant to nourseothricin. Again, an explanation of this is that both clones represented the low-frequency event of incorrect integration due to the processing elements in the construct (Blundell *et al.*, 1996), although the reason why it was not possible to isolate clones with the *SAT* cassette integrated at the *CRK2* locus is not known. The failure to isolate clones with the *PAC* cassette integrated at the *CRK2* locus following transfection with the insert from pGL110 is equally puzzling, since disruption of *CRK2* with the *PAC* cassette was demonstrated to be

possible in STIB 247. The reason why the wild-type alleles could not be disrupted in the AnTat 1.1 '360' cell line is not clear. Successful deletion of *CRK2* wild-type alleles with ectopic expression of CRK2 would be a requirement to demonstrate that *CRK2* is an essential gene in the procyclic-stage cell i.e. that the failure of the targeted gene disruption approach was specifically due to loss of *CRK2*. The lack of success in deleting wild-type alleles in a background of ectopic expression precludes designation of *tbCRK2* as an essential gene.

As an alternative approach to assess the kinase activity of CRK2, instead of immunoprecipitation with antibodies, histidine-tagged CRK2 was expressed from the rRNA spacer using the pHD675 construct (Biebinger *et al.*, 1997). Expression of CRK2His in the '449' cell line (Biebinger *et al.*, 1997) should allow affinity purification of CRK2His by Ni-NTA resin selection; such an approach has been used successfully in Leishmania (Mottram *et al.*, 1996; Grant *et al.*, 1998). As the resistance marker in this construct is constitutively expressed and independent of the cloned copy of *CRK2His*, selection was performed without induction of the cloned *CRK2His* (chapter 5.2.7). Induction of the heterogeneous population that arose (WUMP 1163) with different amounts of tetracycline showed relatively little CRK2His induced with 0.2 ng ml^{-1} or 1.0 ng ml^{-1} tetracycline, but much larger quantities induced with 10.0 ng ml^{-1} (Fig. 5.15). The level of induction achieved with this system depends on the particular gene used (Biebinger *et al.*, 1997), and in this case appeared to require relatively large amounts of tetracycline to get reasonable expression. A comparison of STIB 247 with WUMP 1163 grown with and without tetracycline was carried out, with selection on Ni-NTA resin. EVREE antibodies detected no endogenous CRK2 purified from STIB 247, as was expected as no CRK2His is present in this cell line. A small but significant amount of CRK2His was purified from WUMP 1163 grown in the absence of tetracycline, suggesting that transcriptional control of the pHD675 construct was not total, and a large amount of CRK2His was detected from WUMP 1163 grown in the presence of 10 ng ml^{-1} tetracycline. Curiously, the purified CRK2His demonstrated by Western blotting could not be visualised by silver staining, suggesting that CRK2 is present at very low cellular levels. No kinase activity was detected with purified CRK2His with histone H1 as a substrate, which may be due to the potential problems with substrate choice mentioned previously. Alternatively, the hexahistidine tag may interfere with the kinase activity of CRK2. However, the successful expression and purification of CRK2His in *T. brucei* has promise for the future biochemical characterisation of tbCRK2.

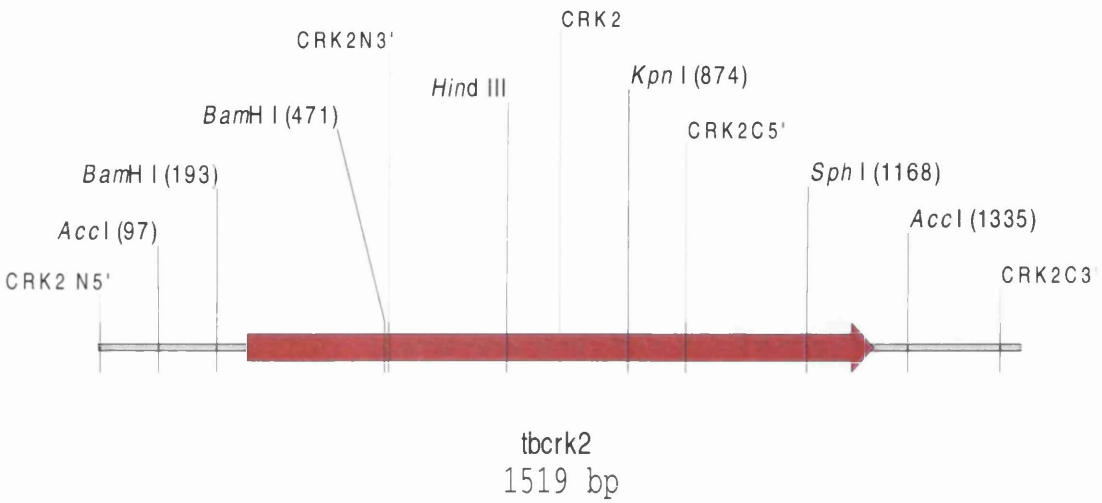
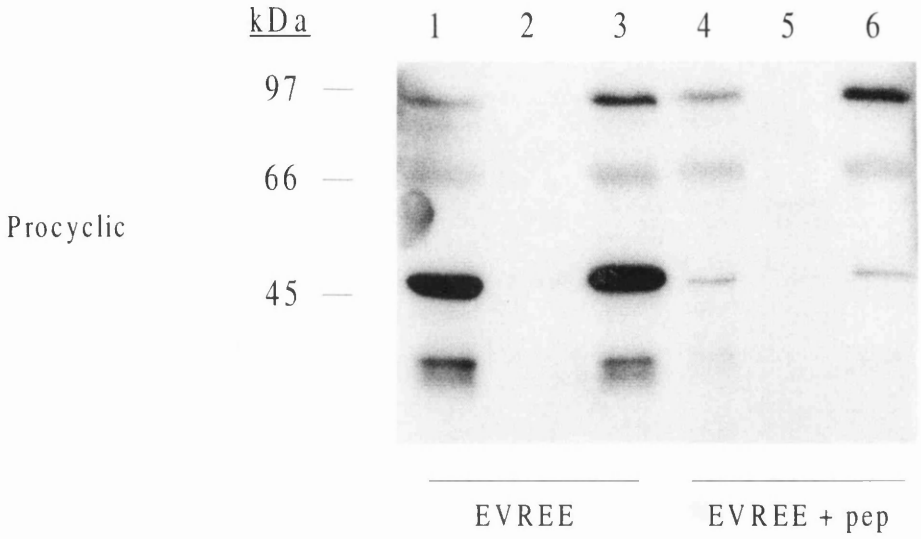
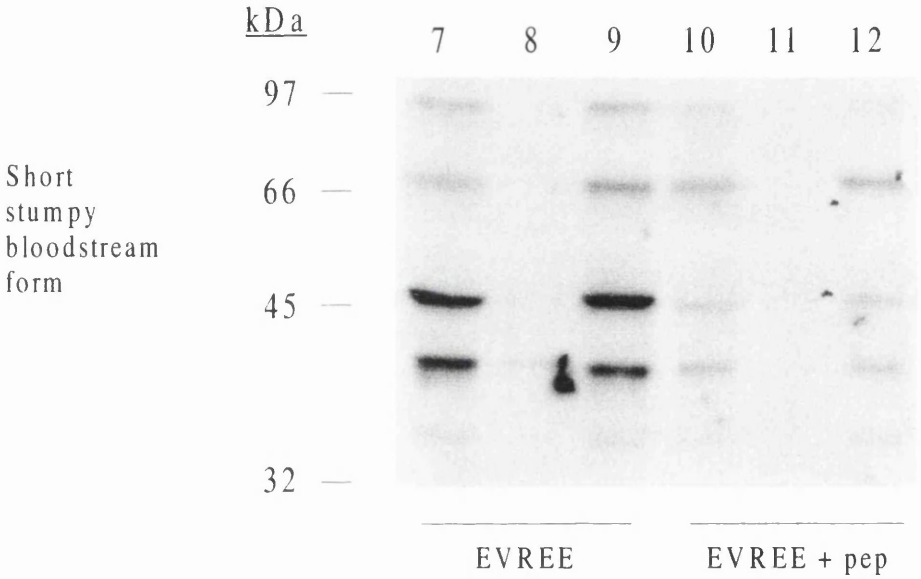


Fig. 5.1: The *T. brucei CRK2* gene

Schematic showing the open reading frame of CRK2 with respect to the relevant primers and restriction endonuclease sites. The *Acc*I sites and the *Hin*dIII site were used in the generation of a cloned copy of full-length *CRK2*. The *Bam*HI sites were used as diagnostic sites for the correct integration of resistance cassettes, and the *Hin*dIII and *Sph*I sites were used in the generation of histidine-tagged *CRK2*. Primer combination CRK2N5' with CRK2N3', and CRK2C5' with CRK2C3' were used to amplify fragments of DNA corresponding to the 5' and 3' ends of the gene, respectively, for use in the generation of the disruption resistance cassettes. Primer combination CRK2N5' with CRK2C3' was frequently used as a diagnostic for correct integration of the resistance cassettes.


kDa
1 2 3 4 5 6
97
66
45
Procyclic
EVREE
EVREE + pep


kDa
7 8 9 10 11 12
97
66
45
32
Short stumpy bloodstream form
EVREE
EVREE + pep

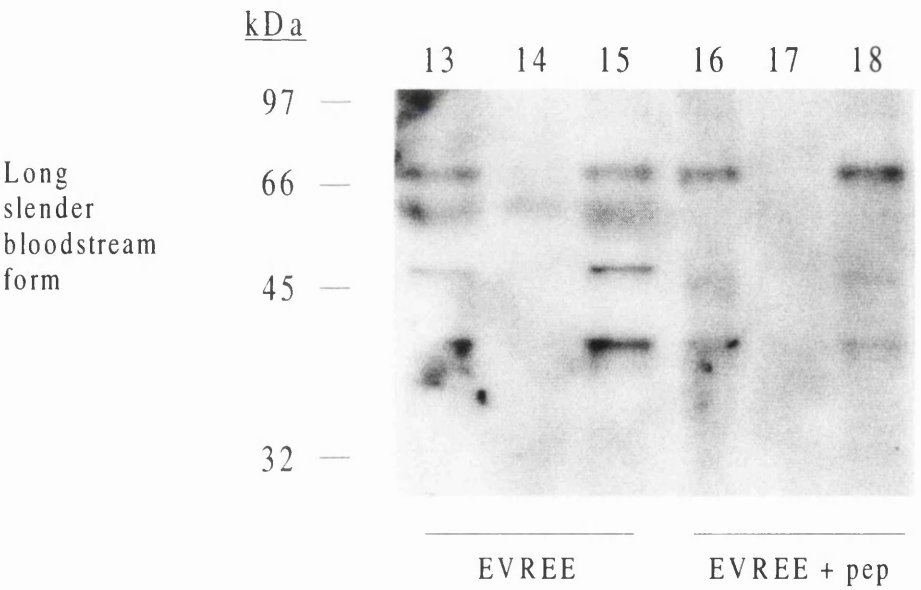

kDa
13 14 15 16 17 18
97
66
45
32
Long slender bloodstream form
EVREE
EVREE + pep

Fig. 5.2: Western blot with EVREE antibodies and STIB 247 procyclic, short stumpy and long slender bloodstream form protein lysates

For each lane 5x 10^6 *T. b. brucei* STIB 247 procyclic, long slender and short stumpy bloodstream form lysates were electrophoresed on a 12.5% Lammeli gel at 130 V at room temperature. The protein was transblotted to PVDF membrane by wet-blotting at 250 mA for 3 hours at 4°C. The membranes were probed with purified EVREE antibodies at a titre of 1:250, and secondary anti-rabbit HRP-conjugate at 1:2500. The blot was developed with ECL luminol reagents (Amersham). Markers were the Low Range nonprestained markers (BioRad).

Lanes 1-3, 7-9, 13-15. Purified EVREE antibodies

Lanes 4-6, 10-12, 16-18. Purified EVREE antibodies with 10 μg EVREE peptide

Lanes 1, 4. STIB 247 procyclic total protein lysate

Lanes 2, 5. STIB 247 procyclic insoluble protein lysate

Lanes 3, 6. STIB 247 procyclic soluble protein lysate

Lanes 7, 10. STIB 247 short stumpy bloodstream form total protein lysate

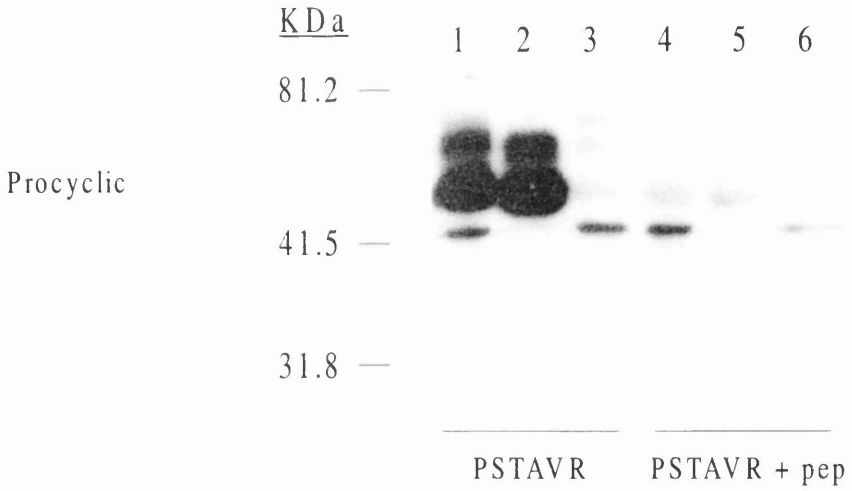
Lanes 8, 11. STIB 247 short stumpy bloodstream form insoluble protein lysate

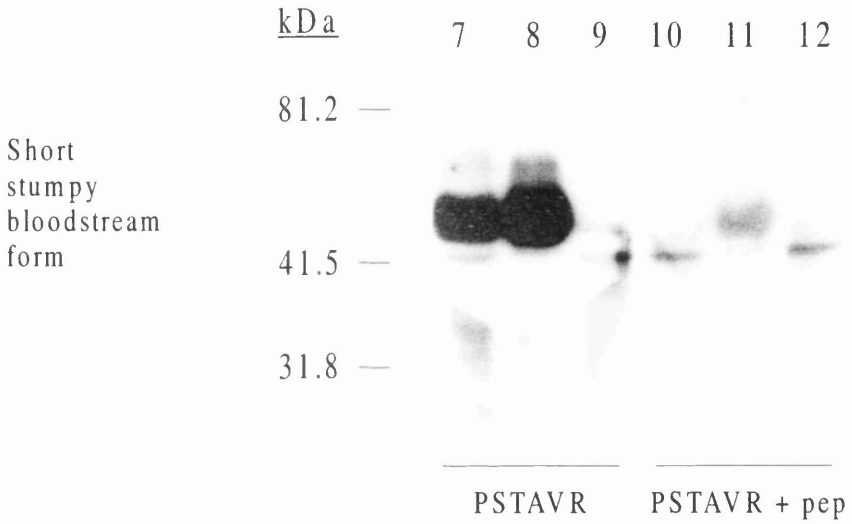
Lanes 9, 12. STIB 247 short stumpy bloodstream form soluble protein lysate

Lanes 13, 16. STIB 247 long slender bloodstream form total protein lysate

Lanes 14, 17. STIB 247 long slender bloodstream form insoluble protein lysate

Lanes 15, 18. STIB 247 long slender bloodstream form soluble protein lysate


KDa
1 2 3 4 5 6
81.2
41.5
31.8
Procyclic
PSTAVR
PSTAVR + pep


kDa
7 8 9 10 11 12
81.2
41.5
31.8
Short stumpy bloodstream form
PSTAVR
PSTAVR + pep

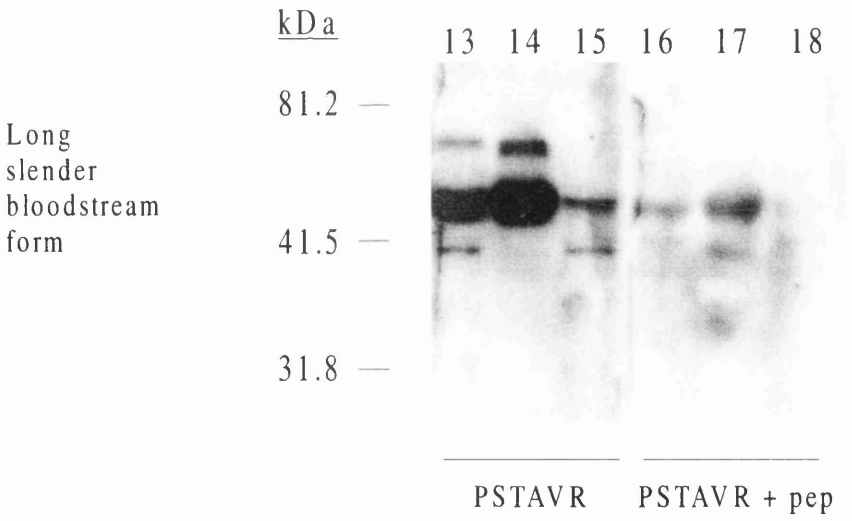

kDa
13 14 15 16 17 18
81.2
41.5
31.8
Long slender bloodstream form
PSTAVR
PSTAVR + pep

Fig. 5.3: Western blot with PSTAVR antibodies and STIB 247 procyclic, short stumpy and long slender bloodstream form protein lysates

For each lane 5x 10^6 *T. b. brucei* STIB 247 procyclic, long slender and short stumpy bloodstream form lysates were electrophoresed on a 12.5% Lammeli gel at 130 V at room temperature. The protein was transblotted to PVDF membrane by wet-blotting at 250 mA for 3 hours at 4°C. The membranes were probed with affinity purified PSTAVR antiserum at a titre of 1:250, and secondary anti-rabbit HRP-conjugate at 1:2500. The blot was developed with ECL luminol reagents (Amersham). Markers were the Kalidescope prestained markers (Amersham).

Lanes 1-3, 7-9, 13-15. Purified PSTAVR antibodies

Lanes 4-6, 10-12, 16-18. Purified PSTAVR antibodies with 10 μg PSTAVR peptide

Lanes 1, 4. STIB 247 procyclic total protein lysate

Lanes 2, 5. STIB 247 procyclic insoluble protein lysate

Lanes 3, 6. STIB 247 procyclic soluble protein lysate

Lanes 7, 10. STIB 247 short stumpy bloodstream form total protein lysate

Lanes 8, 11. STIB 247 short stumpy bloodstream form insoluble protein lysate

Lanes 9, 12. STIB 247 short stumpy bloodstream form soluble protein lysate

Lanes 13, 16. STIB 247 long slender bloodstream form total protein lysate

Lanes 14, 17. STIB 247 long slender bloodstream form insoluble protein lysate

Lanes 15, 18. STIB 247 long slender bloodstream form soluble protein lysate

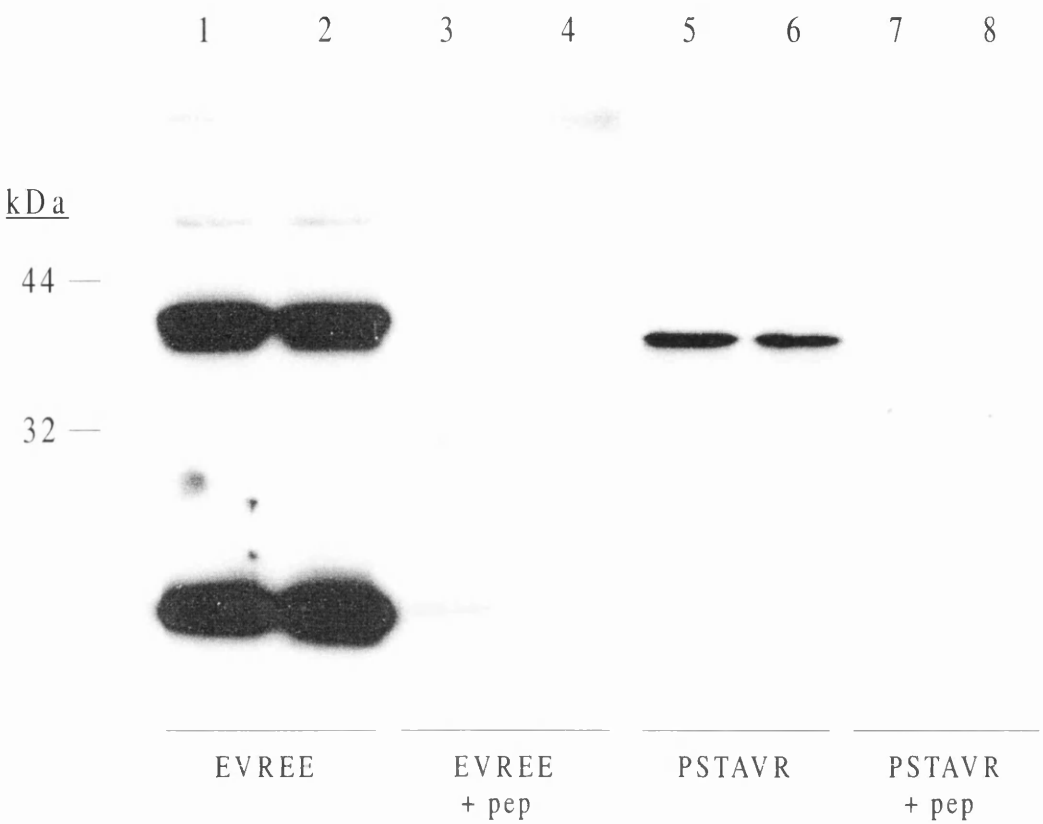


Fig. 5.4: Western blot with 16-mer peptide anti-CRK2 antibodies and M15[pREP4]pGL217

1 ml of a saturated overnight culture of M15[pREP4]pGL217 was used to inoculate each of 2x 10 ml LB. When the OD at 600 nm reached 0.9, one of the cultures was induced by addition of IPTG to a final concentration of 2 mM, and the cultures grown for a further 4 hours. 1 ml of each culture was harvested by centrifugation, whole cell lysates prepared and aliquots of the samples were electrophoresed on a 12.5% Lammeli gel and the protein blotted to PVDF membrane at 250 mA for 2 hours at 4°C. The filter was then probed with purified EVREE and PSTAVR antibodies at a titre of 1:250, detected with secondary anti-rabbit HRP-conjugate at 1:2500 and the blot developed with ECL reagents (Amersham). Markers were the Kalidescope prestained markers (Amersham).

Lanes 1, 3, 5, 7. Uninduced M15[pREP4]pGL217

Lanes 2, 4, 6, 8. Induced M15[pREP4]pGL217

Lanes 1, 2. EVREE antibodies

Lanes 3, 4. EVREE antibodies with 10 μg EVREE peptide

Lanes 5, 6. PSTAVR antibodies

Lanes 7, 8. PSTAVR antibodies with 10 μg PSTAVR peptide

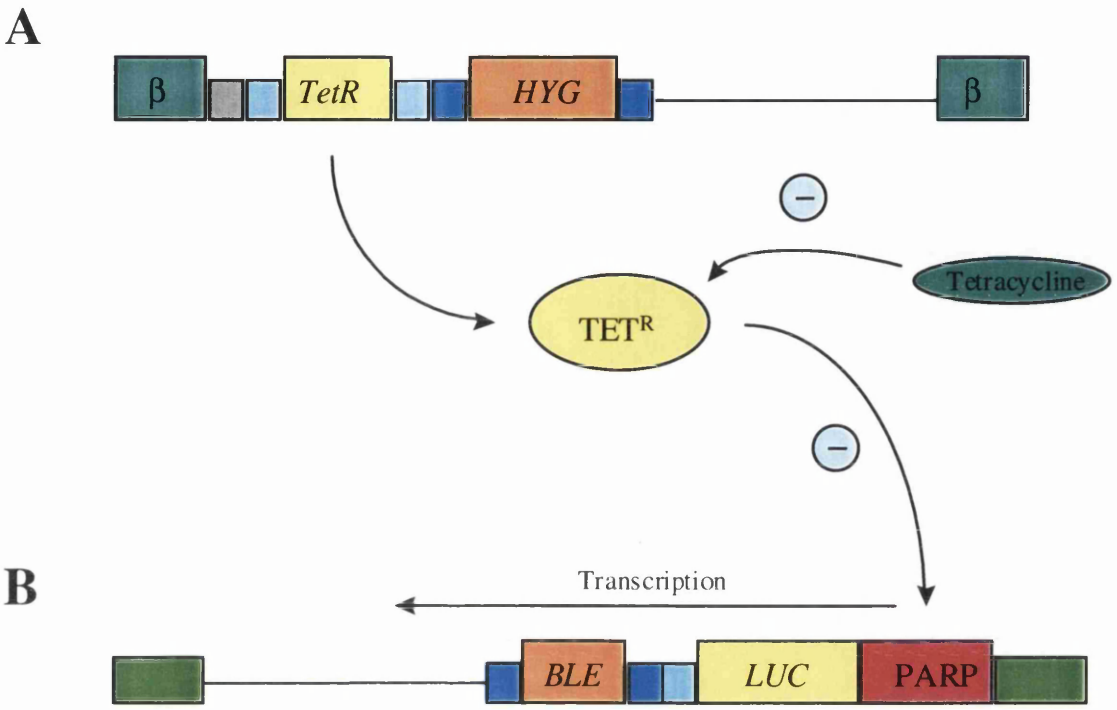


Fig. 5.5: The pHD430 tetracycline induction system

Schematic of the tetracycline-inducible pHD system. A) pHD360 vector is integrated at the repetitive αβ-tubulin locus. The *TET*^{R} gene is constitutively expressed, as is the *HYG* gene due to transcriptional read-through. B) pHD430 vector is integrated at the rRNA intergenic spacer (represented by the dark green boxes) in a reverse orientation to transcription of the 18S rRNA gene. TET^{R} protein binds to the modified PARP promoter (red box), inhibiting transcription of *LUC*. Tetracycline binds to the TET^{R} protein causing conformational change that prevents its binding to the modified PARP promoter, thus allowing transcription of *LUC* and *BLE*. The 5' and 3' processing elements are derived from the aldolase gene (pale blue boxes) and the actin gene array (dark blue boxes).

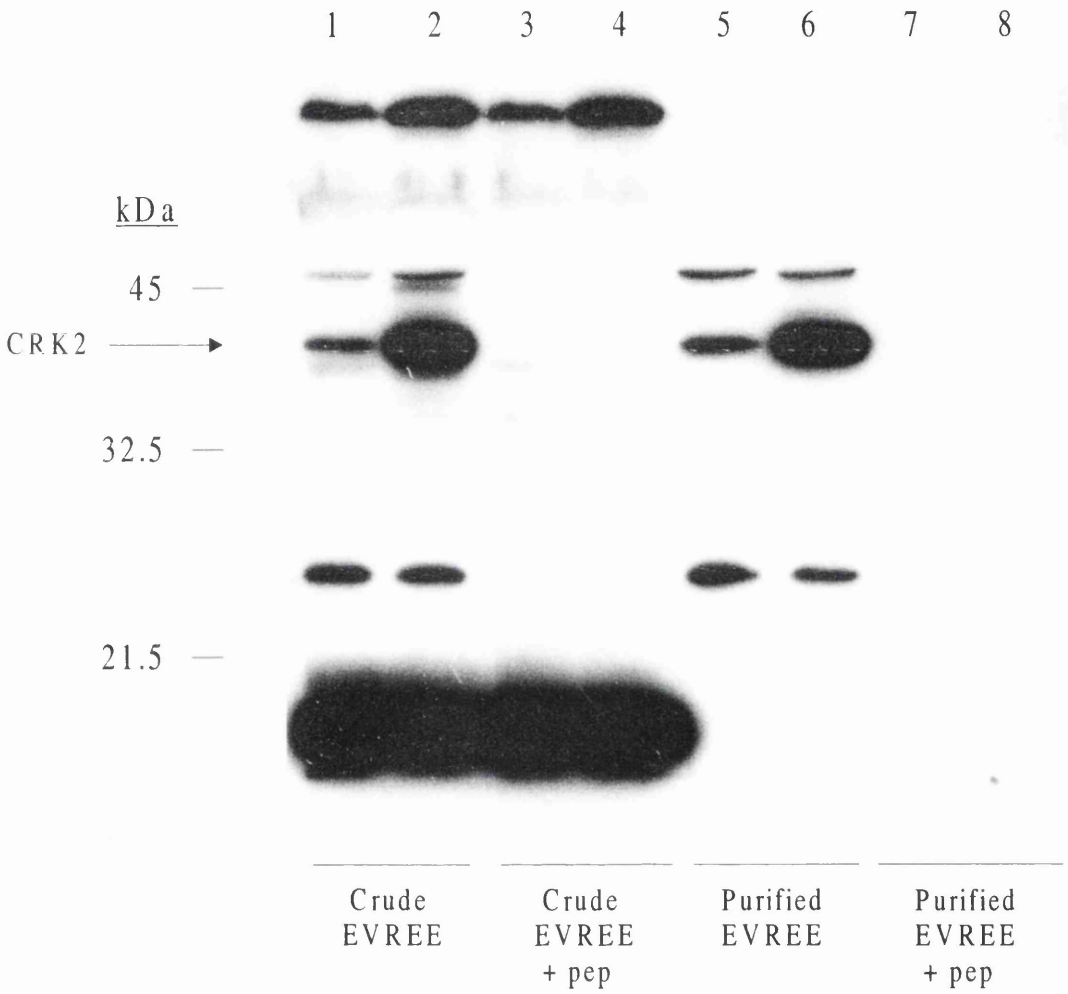

1
2
3
4
5
6
7
8
kDa
45
CRK2
32.5
21.5
Crude EVREE
Crude EVREE + pep
Purified EVREE
Purified EVREE + pep

Fig. 5.6: Western blot comparison of STIB 247 wild-type procyclic and WUMP 827 procyclic total protein lysates with EVREE antibodies

STIB 247 wild-type procyclics were grown to midlog, WUMP 827 procyclics were grown to midlog in the presence of 10 ng ml^{-1} tetracycline. Cells were harvested with thorough washing in PBS, cell pellets resuspended in LSI/TSB, syringed through G25 needles and heated to 100°C for 5 minutes. Approximately 1x 10^7 cells per lane of STIB 247 and WUMP 827 total protein lysates were electrophoresed on 12.5% Lammeli gels at 125 V at room temperature. Protein was transblotted to PVDF membranes at 30 mA overnight at 4°C. The filters were probed with preimmune and crude EVREE antiserum at a titre of 1:500, and purified EVREE antibodies at a titre of 1:50. Secondary anti-rabbit HRP-conjugate was used at 1:2000, and the blots were developed with Supersignal reagents (Pierce). Markers were the Low Range nonprestained markers (BioRad).

Lanes 1, 3, 5, 7. STIB 247 procyclic total protein lysate

Lanes 2, 4, 6, 8. WUMP 827 procyclic total protein lysate

Lanes 1, 2. Crude EVREE antiserum

Lanes 3, 4. Crude EVREE antiserum with 10 μg EVREE peptide

Lanes 5, 6. Purified EVREE antibodies

Lanes 7, 8. Purified EVREE antibodies with 10 μg EVREE peptide

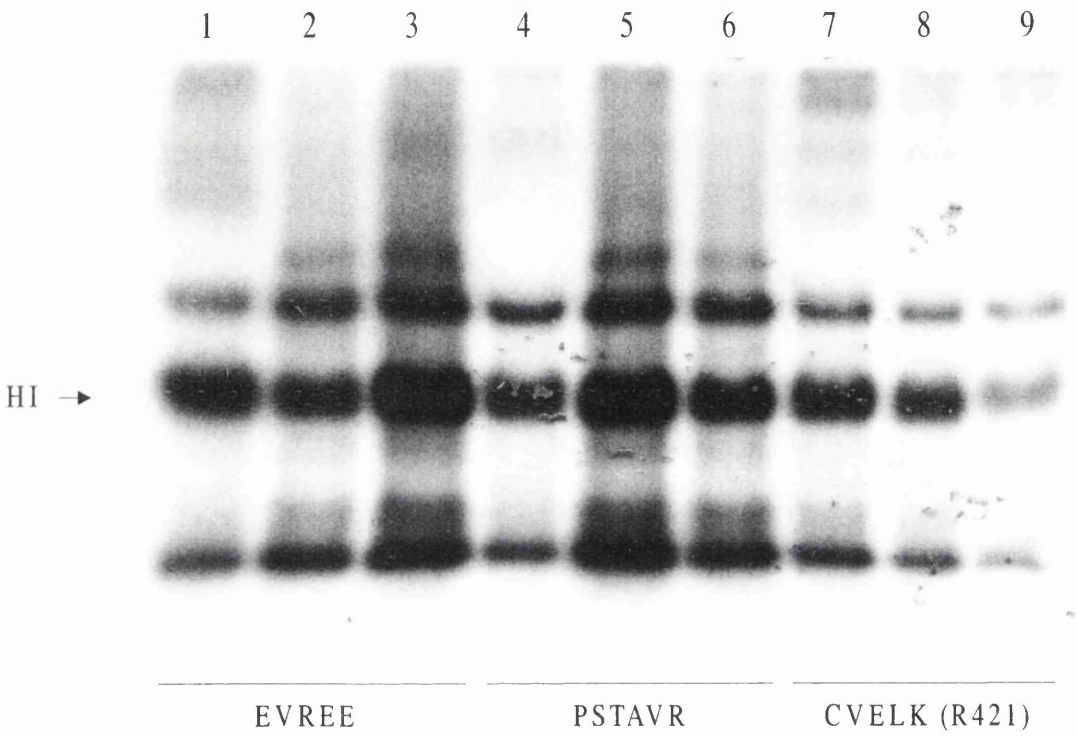

1 2 3 4 5 6 7 8 9
HI
EVREE
PSTAVR
CVELK (R421)

Fig. 5.7: Immunoprecipitation-linked histone HI kinase assay with anti-CRK2 antisera

STIB 247 procyclics were grown to midlog and a total of 4.5x 10^8 cells harvested with extensive washing with PBS, resuspended in 2 ml LSI and an S-100 lysate made. 200 µl of lysate (extract from 4.5x 10^7 cells) was added to 50 µl aliquots of purified EVREE antibodies, purified PSTAVR antibodies, and crude anti-CRK2 (R421) antiserum, and incubated at 4°C for 3 hours. Antisera-lysate mixes were then added to minicolumns containing 80 µl of a 50% slurry of Sepharose-protein A that had been preblocked with a saturating quantity of BSA in LSI, and were left to incubate overnight at 4°C. Columns were drained, washed extensively with LS, HSLS and KAB, the resin aliquots pelleted and resuspended in 20 µl KAM containing 5 µg histone HI as substrate and incubated at 30°C for 20 minutes. Reactions were stopped by addition of 20 µl of TSB and heating to 100°C for 5 minutes, 20 µl samples were run on a 12.5% Lammeli gel, the gel stained with Coomassie blue and dried on an slab dryer (BioRad) and exposed to autoradiography film.

Lane 1. EVREE preimmune serum

Lane 2. Purified EVREE antibodies

Lane 3. Purified EVREE antibodies with 40 mg EVREE peptide

Lane 4. PSTAVR preimmune serum

Lane 5. Purified PSTAVR antibodies

Lane 6. Purified PSTAVR antibodies with 40 mg PSTAVR peptide

Lane 7. R421 preimmune serum

Lane 8. Crude anti-CRK2 (R421) antiserum

Lane 9. Crude anti-CRK2 (R421) antiserum with 50 µg CVELK peptide

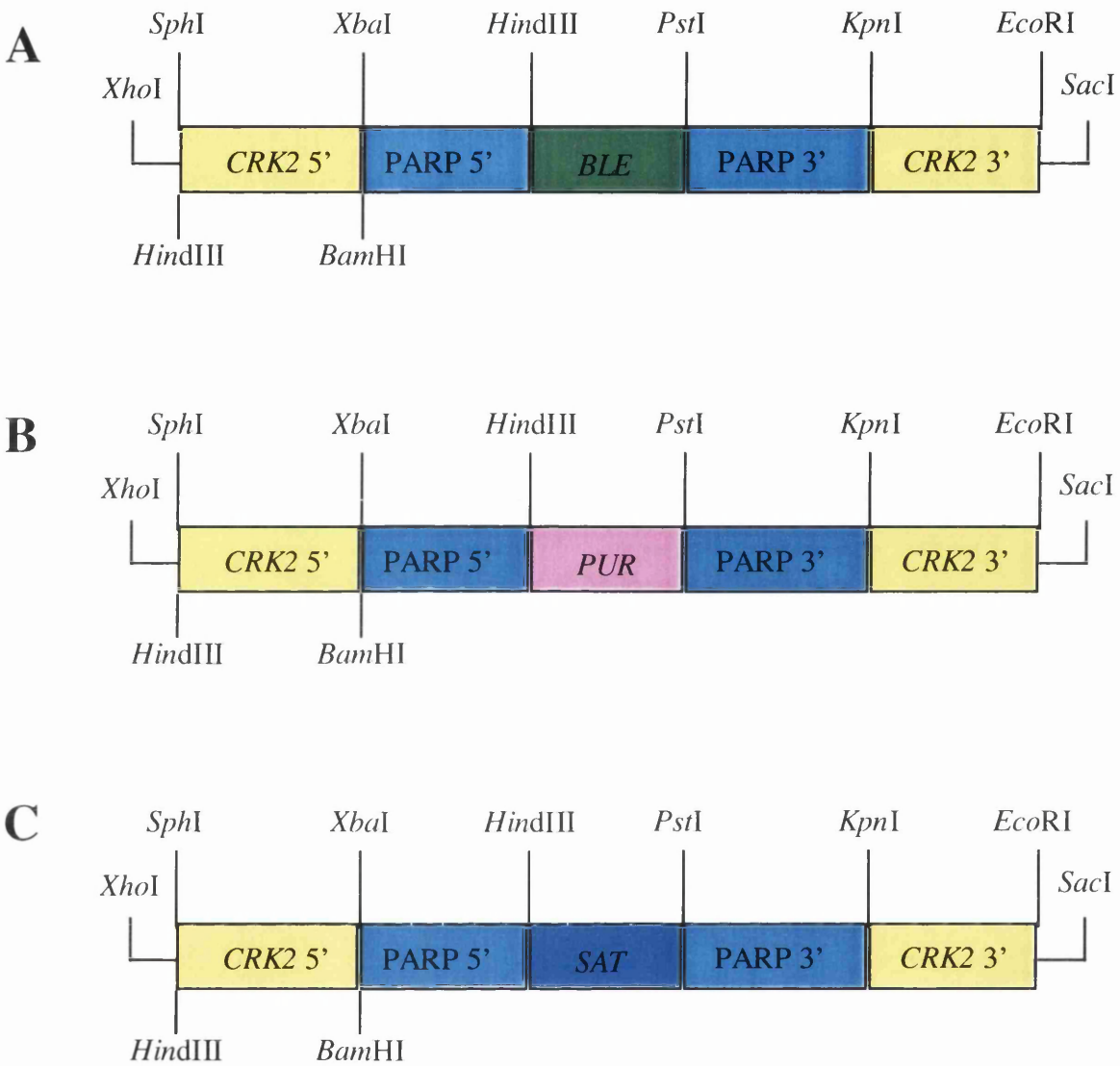
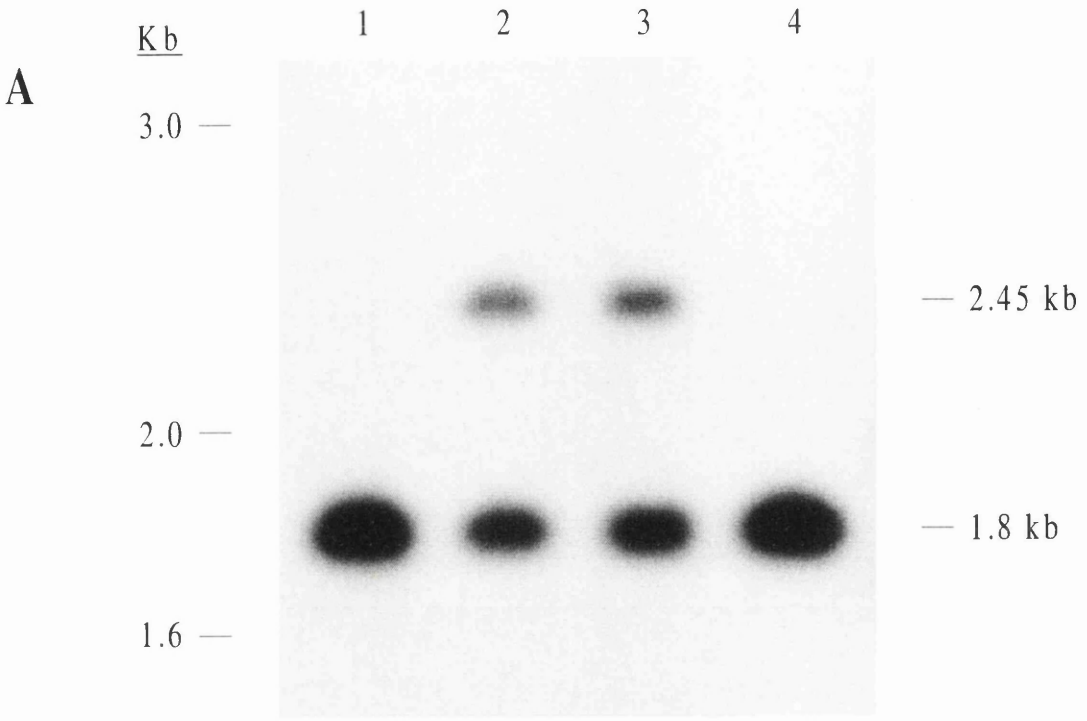


Fig. 5.8: Constructs for targeted disruption of *CRK2*

The three *CRK2* targeted disruption constructs, pGL108 (*BLE*), pGL109 (*SAT*) and pGL110 (*PAC*); the plasmid backbone in each case is pBluescript and is not shown. The constructs were generated as described (chapter 5.2.5). The resistance cassettes were released for transfection by *Sac*I/*Xho*I cleavage and the DNA purified as described (chapter 2.2.3).


A
Kb
1
2
3
4
3.0
2.0
1.6
2.45 kb
1.8 kb

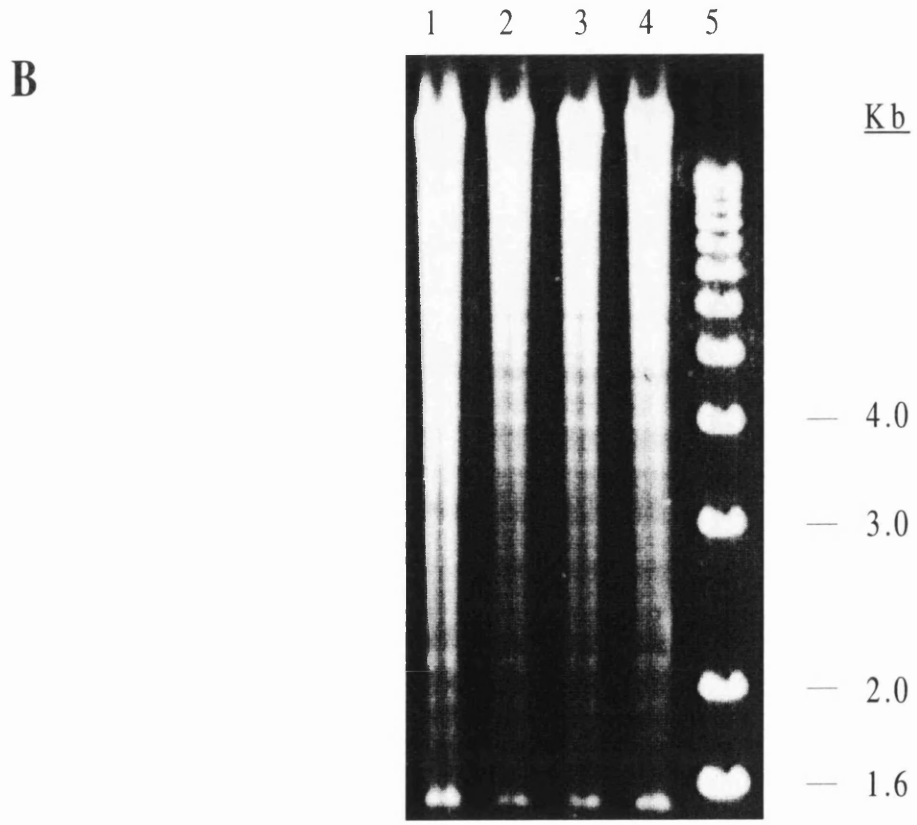

B
1
2
3
4
5
Kb
4.0
3.0
2.0
1.6

Fig. 5.9: Southern blot comparison of WUMP 563 and WUMP 564 with STIB 247 wild-type

Genomic DNA was prepared from procyclic form cells of STIB 247 wild-type, WUMP 563, WUMP 564 and EATRO 795 wild-type were prepared using the Nucleon kit (for details see chapter 2.2.7). 2.5 μg of genomic DNA for each was digested with approximately 20 U of BamHI for 2 hours at 37°C, followed by the addition of a further 10 U of BamHI and incubation at 37°C for a further 2 hours. Samples were electrophoresed on a 0.8% TBE gel at 30 V at room temperature overnight. The gel was stained with 0.5 μg ml^{-1} ethidium bromide and a photograph taken (B), before being processed (see chapter 2.3.7), although here the DNA was transferred to Hybond-N supported nitrocellulose (Amersham) using the Stratagene Posiblot apparatus, with a pressure of 70 mm Hg for 60 minutes. The filter was air-dried and UV fixed, prehybridised at 65°C for 1 hour, and hybridised at 65°C overnight with full-length *CRK2* as the probe, prepared as described in chapter 2.2.3.5. The filter was washed as described in chapter 2.2.3.4 and exposed overnight to autoradiography film (A).

Lane 1. EATRO 795 wild-type

Lane 2. WUMP 564 (*Δcrk2::BLE/CRK2* nonclonal population)

Lane 3. WUMP 563 (*Δcrk2::BLE/CRK2* nonclonal population)

Lane 4. STIB 247 wild-type

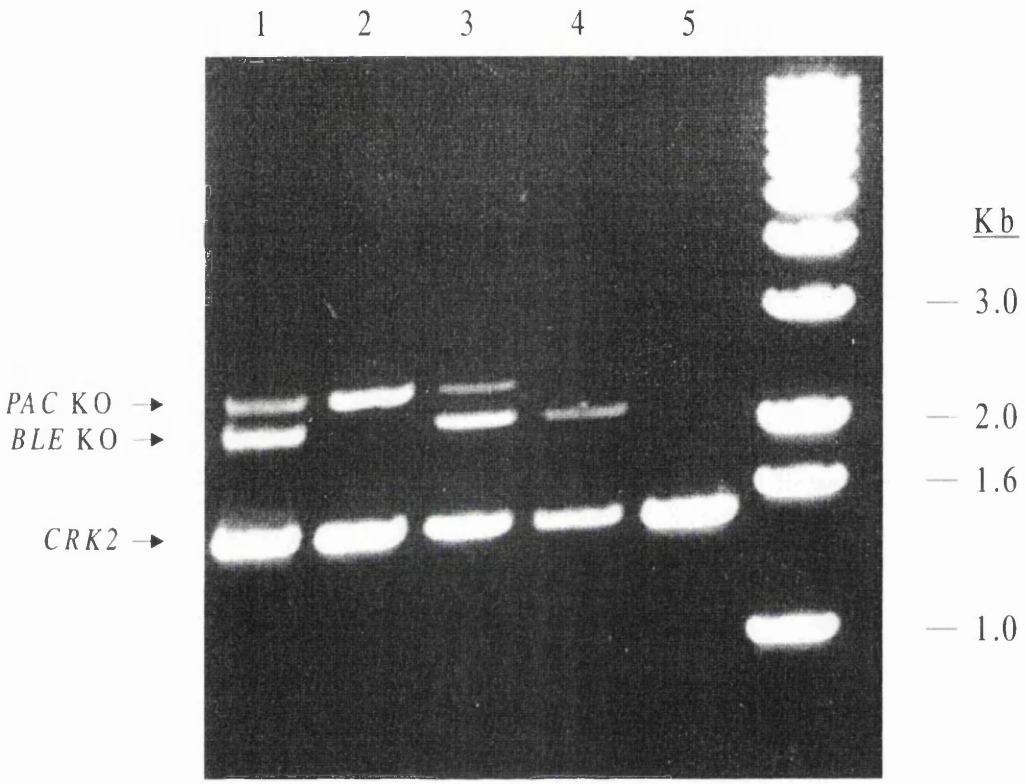

1
2
3
4
5
PAC KO
BLE KO
CRK2
Kb
3.0
2.0
1.6
1.0

Fig. 5.10: PCR from genomic DNA of nonclonal double allele *CRK2* replacement STIB 247 procyclic populations

Genomic DNA was prepared from STIB 247 wild-type procyclics, single allele *CRK2* replacement and double allele *CRK2* replacement nonclonal populations using the Nucleon kit, as described in chapter 2.2.7. 10 µl PCR reactions using 1 µl of resuspended genomic DNA per reaction (representing genomic DNA from 2x 10^6 cells per reaction) were performed using primers CRK2N5' and CRK2C3', with a prehold of 4 minutes at 94°C, 1 minute annealing at 50°C and 1 minute extension at 72°C, for 25 cycles. Reactions were electrophoresed at 70 V on a 1% TAE minigel, and DNA visualised by staining with 1 µg ml^{-1} ethidium bromide. The *Δcrk2::BLE/Δcrk2::PAC* population A arose from transfection of WUMP 563 with the *PAC* cassette, the *Δcrk2::BLE/Δcrk2::PAC* population B from transfection of WUMP 564 with the *PAC* cassette. The reactions in lanes 1, 2 and 3 had DMSO added to a final concentration of 5%, to allow amplification of the GC-rich *PAC* cassette.

Lane 1. WUMP 564 transfected with the cassette from pGL110 and selected with phleomycin and puromycin

Lane 2. WUMP 563 transfected with the cassette from pGL110 and selected with puromycin

Lane 3. WUMP 563 transfected with the cassette from pGL110 and selected with phleomycin and puromycin

Lane 4. WUMP 563

Lane 5. STIB 247 wild-type

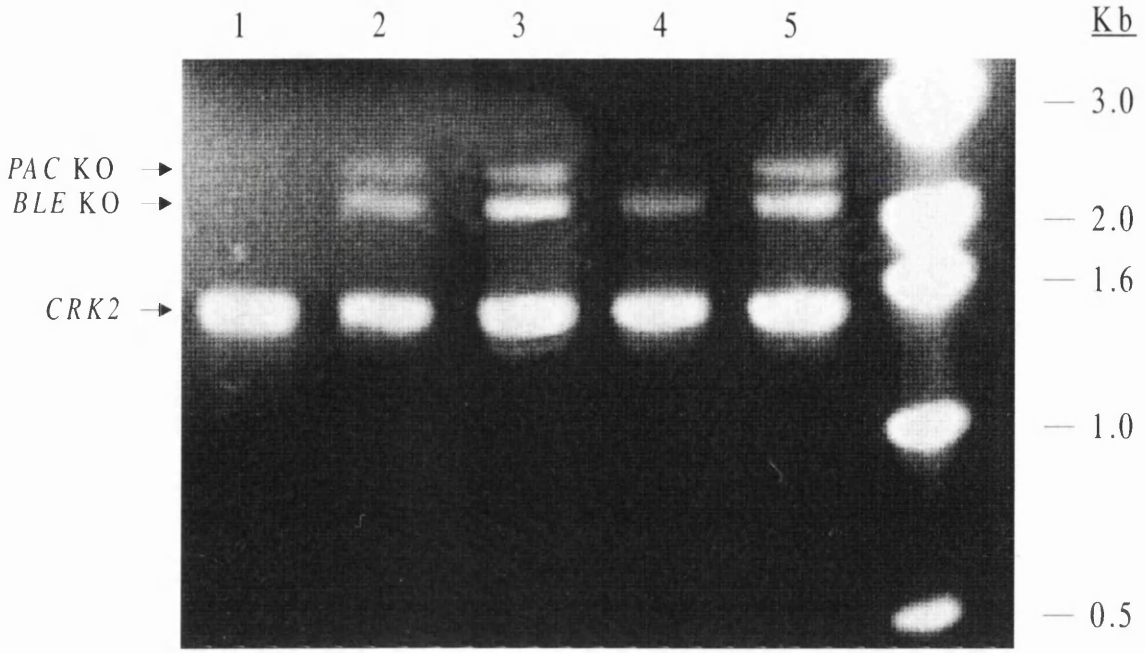


Fig. 5.11: PCR from genomic DNA of *Δcrk2::BLE/Δcrk2::PAC* clones

Genomic DNA was prepared from WUMP 837, WUMP 838, WUMP 839, WUMP 840 and STIB 247 wild-type procyclics with the Nucleon kit, as described in chapter 2.2.7. 10 µl PCR reactions using 1 µl of resuspended genomic DNA per reaction (representing genomic DNA from 2x 10^6 cells per reaction) were performed using primers CRK2N5' and CRK2C3', with a prehold of 4 minutes at 94°C, 1 minute annealing at 50°C and 1 minute extension at 72°C, for 25 cycles. Each reaction also contained 5% DMSO, to allow amplification of the GC-rich *PAC* gene.

Lane 1. STIB 247 wild-type

Lane 2: WUMP 840

Lane 3. WUMP 839

Lane 4. WUMP 838

Lane 5. WUMP 837

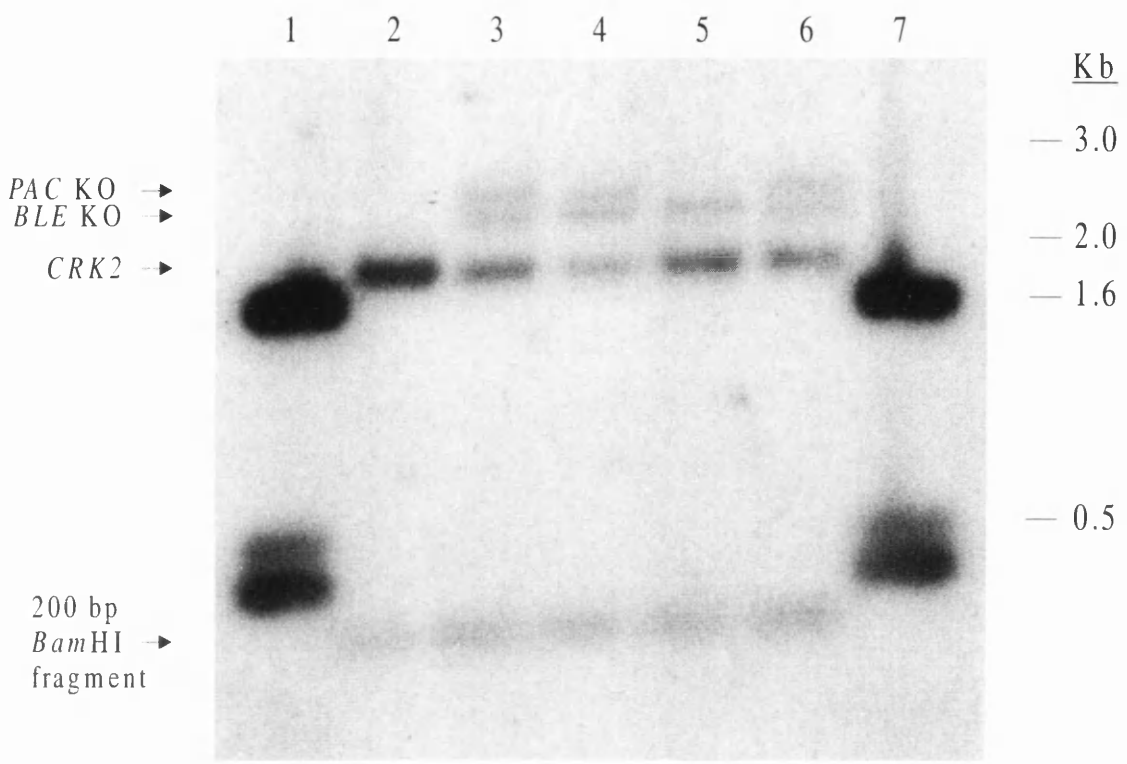


Fig. 5.12: Southern blot of genomic DNA of *Δcrk2::BLE/Δcrk2::PAC* clones

Genomic DNA was prepared from WUMP 837, WUMP 838, WUMP 839, WUMP 840 and STIB 247 wild-type procyclics with the Nucleon kit. 2.5 µg of each were digested with *Bam*HI and electrophoresed on a 0.8% TBE agarose gel at 50 V, and the DNA transferred to Hybond-N supported nitrocellulose (Amersham) by the capillary method. The membrane was probed with full-length *CRK2*, released from pGL130 by *Xba*I/*Xho*I cleavage, and exposed to autoradiography film (DuPont).

Lanes 1, 6. 1 kb ladder markers (GibcoBRL)

Lane 2. STIB 247 wild-type

Lane 3: WUMP 840

Lane 4. WUMP 839

Lane 5. WUMP 838

Lane 6. WUMP 837

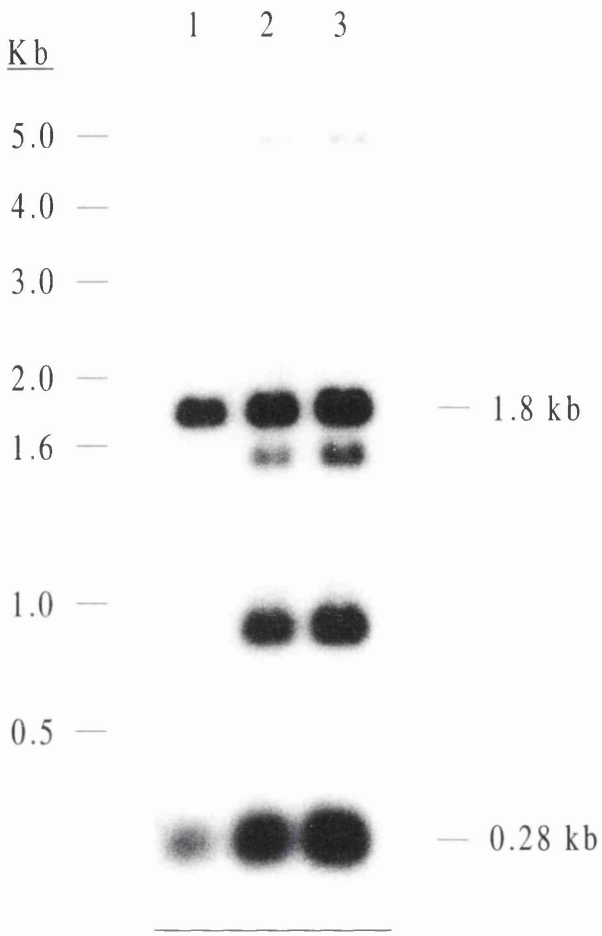


Fig. 5.13: Southern blot of WUMP 835 and WUMP 836

Approximately 2.5 µg of genomic DNA was digested with *Bam*HI and electrophoresed on a 0.8% TBE agarose gel. DNA was blotted to supported nitrocellulose membrane by the capillary method and probed with full-length *CRK2*. Markers were the 1 kb ladder markers (GibcoBRL).

Lane 1. STIB 247 wild-type

Lane 2. WUMP 835

Lane 3. WUMP 836

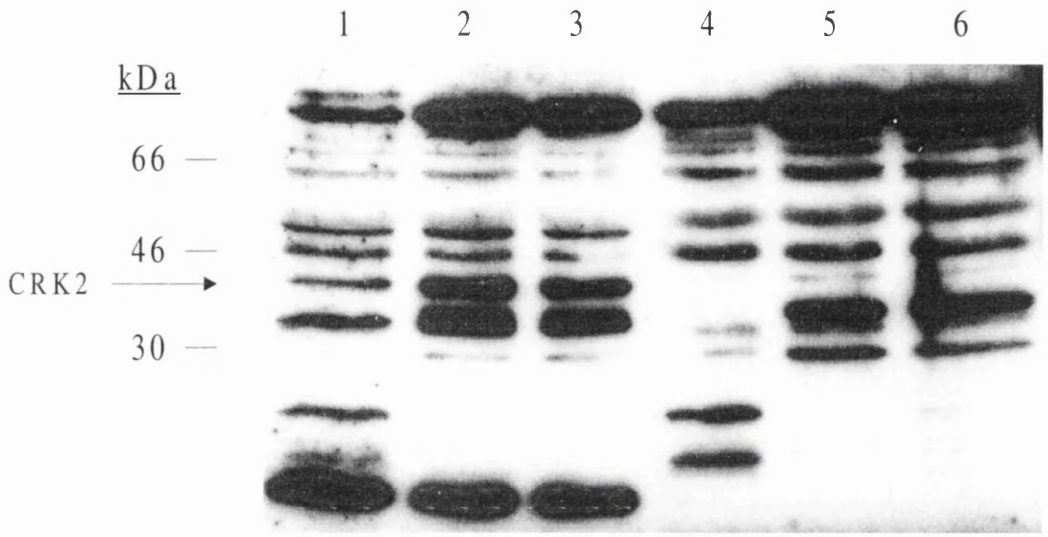


Fig. 5.14: Western blot comparing STIB 247 with WUMP 835 and WUMP 836

For each lane 5x 10^6 *T. b. brucei* STIB 247 wild-type, WUMP 835 and WUMP 836 procyclic lysates were electrophoresed on a 12.5% Lammeli gel at 130 V at room temperature. The protein was transblotted to PVDF membrane by wet-blotting at 250 mA for 3 hours at 4°C. The membranes were probed with partially purified EVREE antibodies (protein A/G purified) at a titre of 1:50, and secondary anti-rabbit HRP-conjugate at 1:2500. The blot was developed with ECL luminol reagents (Amersham). Markers were the Low Range nonprestained markers (BioRad).

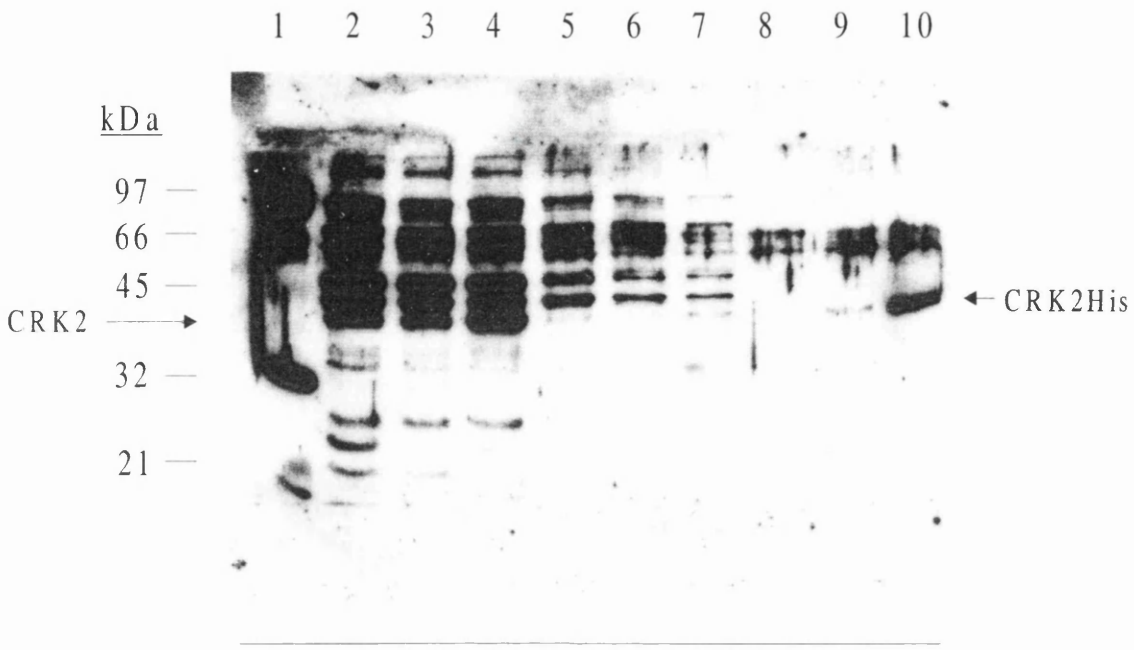
Lanes 1-3. EVREE antibodies

Lanes 4-6. EVREE antibodies with 5 µg EVREE peptide

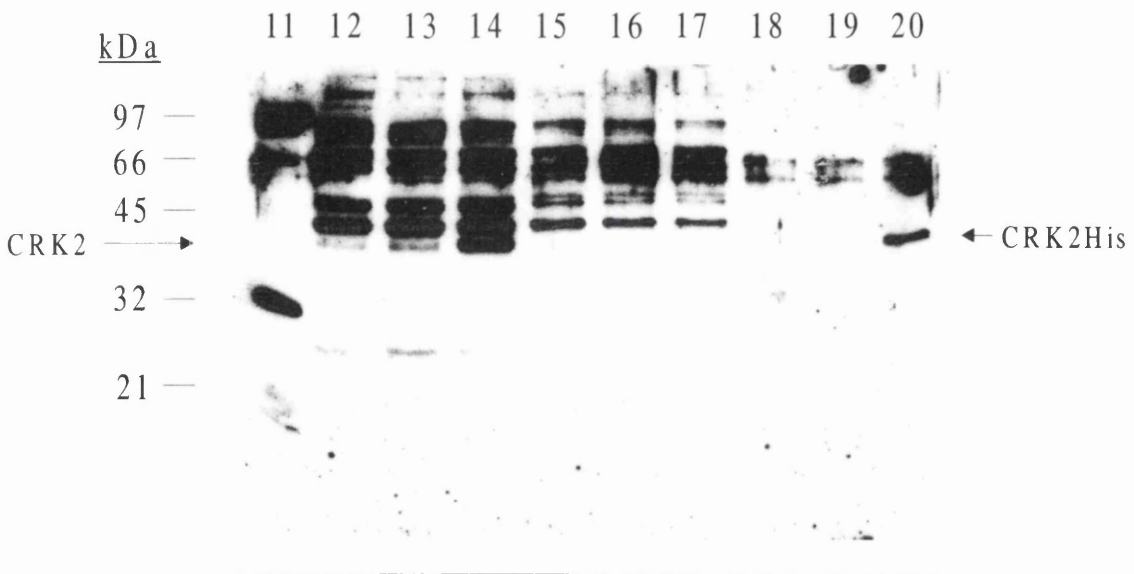
Lane 1, 4. STIB 247 wild-type

Lanes 2, 5. WUMP 835

Lanes 3, 6. WUMP 836



EVREE



EVREE + pep

Fig. 5.15: Western blot of WUMP 1163 with purified EVREE antibodies

10 ml midlog cultures of WUMP 1163 were induced with 0.2, 1.0 and 10.0 ng ml^{-1} tetracycline overnight. Cells were harvested, resuspended in 500 µl LSGI and S-100 lysates made. 200 µl of each lysate was added to minicolumns containing 100 µl Ni-NTA resin (prewashed with LSGI) and incubated with agitation for 60 mins at 4°C. Columns were drained and washed with LSG, HSLS and LS-T, the beads collected and resuspended in 100 µl of a 1:2 mixture of LS-T:TSB to represent eluate (specifically purified) protein. Samples were electrophoresed on a 12.5% Lammeli gel at 120 V at room temperature. Protein was transblotted to PVDF membrane at for at 4°C. The membrane was probed with purified EVREE antibodies at a titre of 1:50, and secondary anti-rabbit HRP-conjugate at 1:2500. The blot was developed with ECL luminol reagents (Amersham). Markers were the Low Range nonprestained markers (BioRad).

Lanes 1-10. EVREE antibodies

Lanes 11-20. EVREE antibodies with 10 µg EVREE peptide

Lanes 1-3, 11-13. Induced S-100 lysate

Lanes 4-6, 14-16. Column flow-through

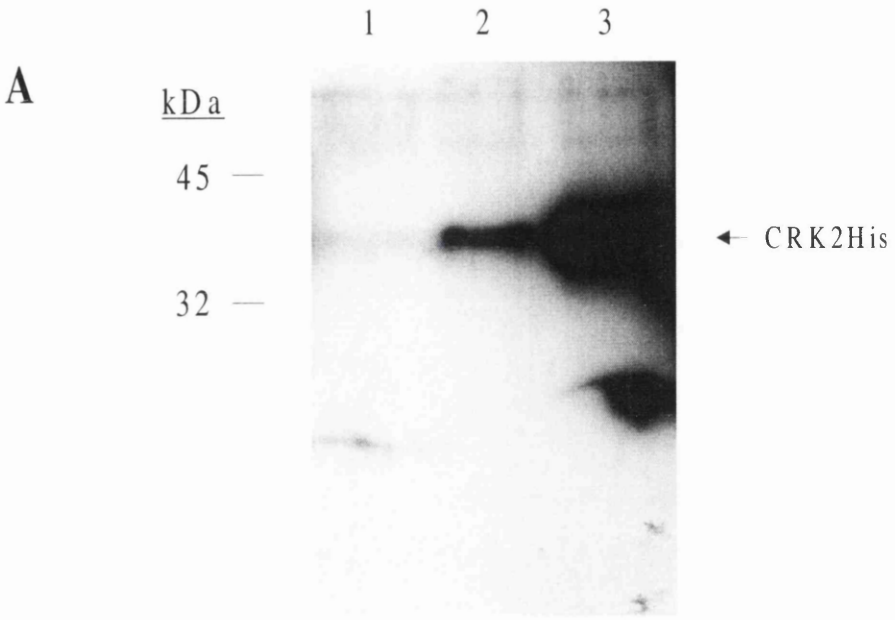
Lanes 7-9, 17-19. Eluate protein from Ni-NTA resin

Lanes 10, 20. Low Range markers.

Lanes 1, 4, 7, 11, 14, 17. WUMP 1163 induced with 0.2 ng ml^{-1} tetracycline

Lanes 2, 5, 8, 12, 15, 18. WUMP 1163 induced with 1.0 ng ml^{-1} tetracycline

Lanes 3, 6, 9, 13, 16, 19. WUMP 1163 induced with 10.0 ng ml^{-1} tetracycline


A
1
2
3
kDa
45
32
CRK2His

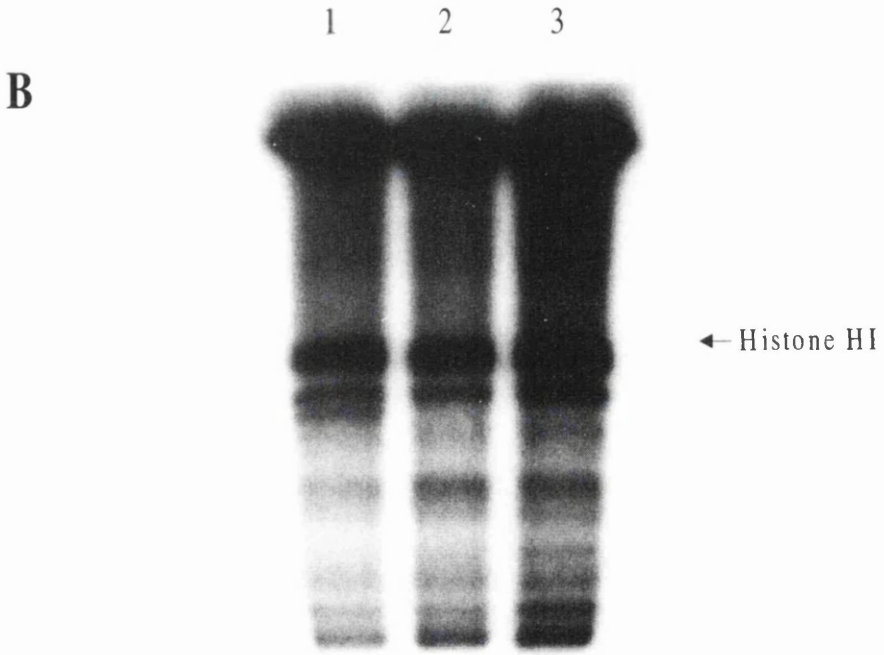

B
1
2
3
Histone HI

Fig. 5.16: Western blot and kinase assay comparing STIB 247 wild-type with WUMP 1163 procyclics

10 ml midlog cultures of STIB 247 wild-type, WUMP 1163 grown without tetracycline and WUMP 1163 grown in the presence of 10.0 ng ml^{-1} tetracycline were harvested, resuspended in 500 µl LSGI and S-100 lysates made. Extracts were added to minicolumns containing 80 µl of Ni-NTA resin (prewashed with LSGI) and incubated with agitation for 60 mins at 4°C. Columns were washed with LSG, HSLS and KAB, beads collected. (A) For the Western blot, beads were resuspended in 100 µl of sample buffer, electrophoresed on a 12.5% Lammeli gel and transblotted to PVDF membrane at 25 mA at 4°C overnight; the membrane was probed with purified EVREE antibodies at a titre of 1:50, with secondary anti-rabbit HRP-conjugate at 1:2500. The blot was developed with ECL luminol reagents (Amersham). Markers were the prestained Rainbow markers (BioRad). (B) For the kinase assay, beads were resuspended in 20 µl of histone HI KAM, incubated at 30°C for 20 mins and reactions stopped by addition of 20 µl sample buffer. Samples were electrophoresed on a 12.5% Lammeli gel and stained with Coomassie blue, dried on a slab dryer (BioRad) and exposed overnight to autoradiography film (DuPont).

A) Western blot: Lane 1. STIB 247 wild-type
Lane 2. WUMP 1163, no tetracycline
Lane 3. WUMP 1163, 10.0 ng ml^{-1} tetracycline

B) Kinase assay: Lane 1. STIB 247 wild-type
Lane 2. WUMP 1163, no tetracycline
Lane 3. WUMP 1163, 10.0 ng ml^{-1} tetracycline

CHAPTER 6: GENERAL DISCUSSION

6.1: Gene disruption of *tbCRK2* and *tbCYC1*

Targeted gene disruption of *tbCRK2* and *tbCYC1* was carried out to determine whether either gene played an essential role in procyclic form *T. b. brucei* STIB 247 *in vitro* growth. The constructs used were based upon expression of *BLE*, *SAT* and *PAC* resistance genes under the constitutive expression of a PARP promoter and PARP 3' elements, flanked by sequences corresponding to the 5' and 3' ends of the *tbCRK2* and *tbCYC1* genes.

In the case of both genes, targeted disruption of one allele was successful. PCR showed one allele of *tbCYC1* replaced by the *tbCYC1-PAC* cassette (Fig. 4.13) although the primers used sat within the targeted disruption construct, and Southern blotting was not performed; Southern blotting and PCR showed one allele of *tbCRK2* replaced by the *tbCRK2-BLE* cassette or the *tbCRK2-PAC* cassette (Fig. 5.9, 5.10). Cell morphology and growth kinetics were comparable with wild-type STIB 247 in both cases, therefore neither single allele disruption was deleterious to the normal cell cycle progression of the procyclic form.

Attempts to generate null mutants for *tbCRK2* and *tbCYC1* were not successful. In both cases, transfection to disrupt the remaining wild-type allele yielded clones that contained wild-type *tbCRK2* or wild-type *tbCYC1* (Fig. 5.11 and 4.14, respectively). For *tbCRK2*, two genotypes were observed (Fig. 5.11, 5.12). Firstly, the correct integration of one resistance marker but the incorrect integration of the second marker (WUMP 838), and secondly, the apparently correct integration of both markers at the *tbCRK2* locus but with the maintenance of a wild-type copy of *tbCRK2* (WUMP 837, 839, 840). For the former, an explanation is that the selection conditions used allowed the isolation of a clone in which the second marker (*PAC*) had integrated at the PARP locus due to nuclease degradation of the *tbCRK2*-specific termini of the *PAC* cassette. For the latter, an explanation is more difficult. FACs analysis demonstrated that WUMP 837, 839 and 840 had the same DNA content as wild-type STIB 247, therefore there had been no change in ploidy of these clones. The possibility that STIB 247 was trisomic at the chromosome containing *tbCRK2* was considered unlikely as genetic analysis has shown STIB 247 to be diploid (Melville *et al.*, 1998), although it was possible that they had become trisomic following transfection. Tertiary transfection of WUMP 837 with the *tbCRK2-SAT* cassette yielded a single clone in which a wild-type allele was present. For

tbCYC1, two genotypes were observed after attempted disruption of the remaining wild-type *tbCYC1* allele. The data is difficult to interpret, with none of the clones apparently showing correct integration of either resistance cassette at the *tbCYC1* locus. However, wild-type *tbCYC1* was clearly present in all clones isolated.

Under the criteria accepted for the definition of essential genes in trypanosomatids - that an essential gene cannot be deleted and attempts to do so lead to genomic changes (Cruz *et al.*, 1993; Mottram *et al.*, 1996; Dumas *et al.*, 1997) - the data may suggest, but does not prove, that *tbCRK2* and *tbCYC1* are essential to the procyclic from of STIB 247. It is important to note that it was not possible to delete *tbCRK2* wild-type alleles with ectopic expression of *tbCRK2*, which is a control required to show that failure to isolate null mutants for the wild-type alleles was specifically due to loss of *tbCRK2*. As the locus around the *tbCYC1* open reading frame was not characterised, it cannot be said that failure to delete wild-type alleles was specific, in that the promoter elements from the disruption constructs may have been interfering with normal expression of nearby genes.

6.2: Biochemical characterisation of tbCRK2 and tbCYC1

It proved difficult to raise sufficiently specific antibodies to tbCRK2 to perform detailed biochemical characterisation. Conformation of the 39 kDa protein recognised by the EVREE antibodies as tbCRK2 by ectopic expression of a cloned copy of *tbCRK2* also confirmed that the 38-39 kDa protein recognised in the procyclic, long slender and short stumpy bloodstream stages is tbCRK2. The gene is therefore expressed and protein present in all three life cycle stages studied. The ectopic expression of *tbCRK2* also showed that abnormally elevated levels of tbCRK2 was not deleterious to the parasite. This would be consistent with activation of tbCRK2 by binding of a cyclin partner, in that the level of tbCRK2-associated activity could be determined by expression and cellular levels of the cyclin partner, not tbCRK2. However, although tbCRK2 was demonstrated to be expressed in the procyclic form, no precipitable histone H1 kinase activity was detected despite the fact that tbCRK2 is shown to bind $p12^{CKS1}$. Leishmanial CRK3 has been shown to bind $p12^{CKS1}$ and to possess a histone H1 kinase activity (Grant *et al.*, 1998). Explanations for this data includes the possibility that histone H1 is not an substrate for tbCRK2, that tbCRK2 is expressed but not active in procyclics, or that it is part of a multimeric complex in procyclics that modify its activity with regards to substrate specificity.

The antibodies raised to purified recombinant tbCYC1His protein showed that tbCYC1 protein is present in the short stumpy bloodstream and procyclic forms of STIB 247, but not in the long slender bloodstream form. This is consistent with the observations that tbCYC1 may be involved in the transformation from the bloodstream to procyclic form (Hua *et al.*, 1997). Proof the tbCYC1 is a functional cyclin homologue requires the demonstration of association with a CDK homologue (CRK protein) but this has not been achieved. Immunoprecipitations using the anti-CYC1 antibodies coupled to Western blots using antibodies to tbCRK1, tbCRK2 and tbCRK3 did not detect any coprecipitated tbCRKs, although the validity of the result is in question, as it appeared that the anti-CYC1 antibodies may not be capable of immunoprecipitating tbCYC1. Selection experiments using recombinant $p13^{SUC1}$ and $p12^{CKS1}$ did not show association of tbCYC1 with either protein. As tbCRK1-3 have been shown to bind $p12^{CKS1}$ (Van Hellemond and Mottram, unpublished) the inference is that tbCYC1 does not functionally associate with tbCRK1-3 in procyclics. Immunoprecipitation-linked kinase assays demonstrated no significant precipitated kinase activity, although these assays may have failed for the reason that the antibodies could not precipitate tbCYC1. Taken together, these data indicate that tbCYC1 either associates with an as yet unidentified tbCRK that does not bind the leishmanial p12 protein, or that it does not associate with tbCRK1-3 in the procyclic form (but possibly does in the short stumpy bloodstream form or other life cycle stages), or that it is not a cyclin homologue and performs a non-cyclin function in the bloodstream form to procyclic transformation.

6.3: Analysis of *tbCRK4*

The full sequence of the *tbCRK4* gene was deduced from fragments derived from genomic DNA, isolated from a λFIXII library screen, and fragments produced by PCR from first-strand cDNA. The two genomic subclone fragments that were isolated represented contiguous sequence meeting at a *Sau*3AI site internal to the gene at position 560 (relative to the A of ATG start), confirmed by the sequence of the 760 bp gene-internal fragment generated by PCR, which overlapped both genomic subclones at this site (Fig. 3.1, 3.9). The full sequence showed *tbCRK4* to have high identity to the other trypanosomatid *CRK*s (Fig. 3.10). The deduced protein sequence showed the presence of all 11 conserved kinase domains common to all protein kinases, as well as the conserved features of members of the CDK family. The PSTAIRE box motif which has been shown to be important for cyclin binding (Jeffrey *et al.*, 1995) is present

though degenerate, with five substitutions (GAPPSTAIREIALLKV) as compared to human CDK1 (EGVPSTAIREISLLKE). Another conserved motif, the DSEI box of CDK1 is present though highly degenerate, with four substitutions (GNTDVDQ) as compared to human CDK1 (GDSEIDQ). Residues important for the regulation of CDK activity are conserved, including the Thr and Tyr in the ATP binding pocket (Thr17 and Tyr18, in tbCRK4), and the conserved Thr161 equivalent (a serine in tbCRK4, at position 231). The tbCRK4 protein is most homologous to the *Crithidia fasciculata* cfCRK protein (Brown *et al.*, 1992), both having large inserts of approximately 70 and 80 amino acids between kinase catalysis domains VIb and VII, and X and XI, respectively. These insert domains showed no homology to each other, and neither of the insert domains of tbCRK4 showed significant homology to either of the insert domains of cfCRK, although hydropathy plotting showed the insert domains of both proteins to be highly hydrophilic. It is therefore possible that there is a homologous structure/function between the insert domains of the two proteins despite a lack of primary sequence homology.

Specific antibodies raised to tbCRK4 demonstrated that the protein is present in the short stumpy bloodstream and procyclic forms, but not in the long slender bloodstream form of *T. b. brucei* STIB 247, confirming the predicted size tbCRK4 of approximately 52kDa. This data, along with the conserved CDK features, suggested that control of tbCRK4 activity is achieved by a combination of stage-specific transcriptional control and post-translational modification. However, no tbCRK4-associated kinase activity could be immunoprecipitated from procyclics using the anti-tbCRK4 antibodies. This could conceivably be due to incorrect folding of the tbCRK4 portion of the fusion protein used to raise the antibodies, so that the resulting antibodies could not recognise native tbCRK4. Alternatively, as the antibodies were raised to the peptide sequence of the first insert domain, it is possible that this insert domain has an *in vivo* function in association with other protein(s), thereby precluding antibody binding to native protein. Other possibilities are that tbCRK4 is present but not active in the procyclic, or that it is not active against any of the substrate proteins assayed or under the conditions used. The failure of tbCRK4 to bind the leishmanial protein $p12^{CKS1}$ suggests that it may have a novel biological function that distinguishes it from tbCRK1-3. Given the lack of success in demonstrating tbCRK4-associated kinase activity or $p12^{CKS1}$ binding, demonstration of cyclin binding would appear to be one prerequisite for classification of tbCRK4 as a functional CDK homologue.

6.4: Future analysis

The tradition reverse genetic approach of targeted gene disruption may be a useful tool for the demonstration of a gene's essential nature in any particular life cycle stage *in vitro*, but due to the genomic plasticity little can be deduced about the gene function in the trypanosome as an observable phenotype often cannot be obtained. For this reason other approaches must be taken. The development of inducible *in vivo* expression systems (Biebinger *et al.*, 1997) should facilitate the examination of CDK and cyclin homologue function in *T. brucei*. To generate phenotypes for *CRK* genes, such systems can be used for the ectopic expression of mutant catalytically inactive forms of CRKs in the bloodstream and procyclic forms, which should allow the investigation of CRK function in the trypanosome cell cycle. In addition, some of the inducible vectors allow tolerable control of expression in both life cycle stages, and therefore may allow the investigation of CRK function in the transformation from bloodstream to procyclic form. For cyclin genes this approach has promise with the expression of forms that are ultrastable due to mutations that disrupt the ubiquitination 'destruction box' sequence, or the expression of truncated forms of the protein. Alternatively, the ectopic expression of antisense RNA to remove cyclin protein could provide insights into cyclin function in *T. brucei*. In particular, given the data that suggests a role for tbCYC1 in the bloodstream to procyclic transformation, these methods could provide a viable approach to determining tbCYC1 function.

From a biochemical perspective, the problems associated with the study of trypanosome CRKs or cyclins has been the difficulty in obtaining sufficiently specific antisera for detailed characterisation. The *in vivo* inducible expression in both bloodstream and procyclic forms of epitope-tagged CRKs and cyclins holds promise for the analysis of the biochemical properties of these proteins, providing that such tags do not destabilise the proteins or interfere with their function. Epitope tagging of CRKs would allow characterisation of kinase activity between life cycle stages, and temporal characterisation of activity in semisynchronous populations of *T. brucei*. Tagging of cyclins would allow characterisation of protein expression between life cycle stages, and possibly protein oscillation through the cell cycle, if trypanosome cyclins behave like mammalian cyclins. The tagging of both CRKs and cyclins should allow the definition of functional CRK-cyclin pairs in both bloodstream and procyclic stages, together with use of yeast two-hybrid systems and coprecipitation experiments. Epitope tagging may also provide a good tool for characterisation of the subcellular localisation

of these proteins by immunofluorescence, and for analysis of tagged CRK or cyclin function in a null mutant background.

In general, trypanosome CRKs and cyclins are an interesting area of research that may have important implications for the development of therapeutics. Now that a number of CRKs have been identified in both *Trypanosoma* and *Leishmania*, and three putative cyclins isolated in *T. brucei*, the challenge is to attempt to define their function as regards cell cycle and life cycle progression in the trypanosomatid parasites.

BIBLIOGRAPHY

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

Affranchino, J. L., Gonzãlez, S. A., and Pays, E. (1993). Isolation of a mitotic-like cyclin homologue from the protozoan *Trypanosoma brucei*. Gene *132*, 75-82.

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

Akimoto, K., Mizuno, K., Osada, S., Hirai, S., Tanuma, S., Suzuki, K., and Ohno, S. (1994). A new member of the third class in the protein kianse C family, PKCλ, expressed dominantly in an undifferentiated mouse embryonal carcinoma cell line and also in many cells and tissues. J. Biol. Chem. *269*, 12677-12683.

Alexandre, S., Paindavoine, P., Tebabi, P., Pays, A., Halleux, S., Steinert, M., and Pays, E. (1990). Differential expression of a family of putative adenylate guanylate-cyclase genes in *Trypanosoma brucei*. Mol. Biochem. Parasitol. *43*, 279-288.

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.

Amon, A., Irniger, S., and Nasmyth, K. (1994). Closing the cell cycle in yeast: G2 cyclin proteolysis in initiated at mitosis presists until the activation of G2 cyclins in the next cell cycle. Cell *77*, 1037-1050.

Amon, A., Surana, U., Muroff, I., and Nasmyth, K. (1992). Regulation of $p34^{CDC28}$ tyrosine phosphorylation is not required for entry into mitosis in *S. cerevisiae*. Nature *355*, 368-371.

Amon, A., Tyres, M., Futcher, B., and Nasmyth, K. (1993). Mechanisms that help the yeast cell cycle clock tick: G2 cyclins transcriptionally activate G2 cyclins and repress G1 cyclins. Cell *74*, 993-1007.

Arion, D., Meijer, L., Brizuela, L., and Beach, D. (1988). Cdc2 is a component of the M-phase specific histone-H1 kinase - evidence for identity with mpf. Cell *55*, 371-378.

Bagella (1998). Cloning of murine CDK9/PITALRE and its tissue-specific expression in development. J. Cell. Physiol. *177*, 206-213.

Bai, C., Richman, R., and Elledge, S. J. (1994). Human cyclin F. EMBO J. *13*, 6087-6098.

Bakalara, N., Seyfang, A., Davis, C., and Baltz, T. (1995a). Characterisation of a life-cycle-stage-regulated membrane protein tyrosine phosphatase in *Trypanosoma brucei*. Eur. J. Biochem. *234*, 871-877.

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

Barrett, M. P., Tetaud, E., Seyfang, A., Bringaud, F., and Baltz, T. (1998). Trypanosome glucose transporters. Mol. Biochem. Parasitol. *91*, 195-205.

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

Bastin, P., Coppens, I., Saint-Remy, J-M., Baudhuin, P., Opperdoes, F. R., and Courtoy, P. J. (1994). Identification of a specific epitope on the extracellular domain of the LDL-receptor of *Trypanosoma brucei brucei*. Mol. Biochem. Parasitol. *63*, 193-202.

Bates, S., Bonetta, L., MacAllan, D., Parry, D., Holder, A., Dickson, C., and Peters, G. (1994). CDK6 (PLSTIRE) and CDK4 (PSK-J3) are a distinct subset of the cyclin-dependent kinases that associate with cyclin D1. Oncogene *9*, 71-79.

Beach, D. H., Durkacz, B., and Nurse, P. (1982). Functionally homologous cell cycle control genes in budding and fission yeast. Nature *300*, 706-709.

Benito, J., Martín-Castellanos, C., and Moreno, S. (1998). Regulation of the G1 phase of the cell cycle by periodic stabilization and degradation of the $p25^{rum1}$ CDK inhibitor. EMBO J. *17*, 482-497.

Biebinger, S., Wirtz, E., Lorenz, P., and Clayton, C. E. (1997). Vectors for inducible expression of toxic gene products in bloodstream and procyclic *Trypanosoma brucei*. Mol. Biochem. Parasitol. *85*, 99-112.

Bitter, W., Gerrits, H., Kieft, R., and Borst, P. (1998). The role of transferrin-receptor variation in the host range of *Trypanosoma brucei*. Nature *391*, 499-502.

Blow, J. J., and Laskey, R. A. (1988). A role for the nuclear envelope in controlling DNA replication within the cell cycle. Nature, Lond. *332*, 546-548.

Blundell, P. A., Rudenko, G., and Borst, P. (1996). Targeting of exogenous DNA into *Trypanosoma brucei* requires a high degree of homology between donor and target DNA. Mol. Biochem. Parasitol. *76*, 215-229.

Booher, R., and Beach, D. (1986). Site-specific mutagenesis of $cdc2^{+}$, a cell-cycle control gene of the fission yeast *Schizosaccharomyces pombe*. Mol. Cell. Biol. *6*, 3523-3530.

Booher, R. N., Alfa, C. E., Hyams, J. S., and Beach, D. H. (1989). The fission yeast *cdc2/cdc13/suc1* protein kinase: regulation of catalytic activity and nuclear localization. Cell *58*, 485-497.

Booher, R. N., Deshaies, R. J., and Kirschner, M. W. (1993). Properties of *Saccharomyces cerevisiae* wee1 and its differential regulation of p34*CDC28* in response to G1 and G2 cyclins. EMBO J. *12*, 3417-3426.

Booher, R. N., Holman, P. S., and Fattaey, A. (1997). Human Myt1 is a cell cycle-regulated kinase that inhibits Cdc2 but not Cdk2 activity. J. Biol. Chem. *272*, 22300-22306.

Boothroyd, J. C., and Cross, G. A. M. (1982). Transcripts coding for variant surface glycoproteins of *Trypanosoma brucei* have a short, identical exon at their 5' end. Gene *20*, 281-289.

Borst, P., Bitter, W., Blundell, P. A., Cross, M., McCulloch, R., Rudenko, G., Taylor, M. C., and Van Leeuwen, F. (1996). The expression sites for variant surface glycoproteins of *Trypanosoma brucei.* Trypanosomiasis and Leishmaniasis: Biology and Control. Ed: Hide, G., Mottram, J. C., Coombes, G. H., and Holmes, P H. British Society for Parasitology/CAB International, Oxford, pp. 109-131.

Boshart, M., and Mottram, J. C. (1997). Protein phosphorylation and protein kinases in trypanosomatids. In Trypanosomiasis and Leishmaniasis: Biology and control, G. Hide, J. C. Mottram, G. H. Coombs and P. H. Holmes, eds. (Oxford, UK: CAB International), pp. 227-244.

Bradbury, E. M. (1992). Reversible histone modifications and the chromosome cell cycle. Bioessays *14*, 9-16.

Brandeis, M., and Hunt, T. (1996). The proteolysis of mitotic cyclins in mammalian cells persists from the end of mitosis until the onset of S-phase. EMBO J. *15*, 5280-5289.

Brandeis, M., Rosewell, I., Carrington, M., Crompton, T., Jacobs, M. A., Kirk, J., Gannon, J., and Hunt, T. (1998). Cyclin B2-null mice develop normally and are fertile whereas cyclin B1-null mice die *in utero.* Proc. Natl. Acad. Sci. USA *95*, 4344-4349.

Braun, K., Hölzl, G., Soucek, T., Geisen, C., Möröy, T., and Hengstschläger, M. (1998). Investigation of the cell cycle regulation of cdk3-associated kinase activity and the role of cdk3 in proliferation and transformation. Oncogene *17*, 2259-2269.

Breeden, L. (1995). Start-specific transcription in yeast. Curr. Top. Microbiol. Immunol. *208*, 95-127.

Brown, L., Hines, J. C., and Ray, D. S. (1992). The *Crithidia fasciculata* CRK gene encodes a novel cdc2-related protein containing large inserts between highly conserved domains. Nucleic Acids Res. *20*, 5451-5456.

Brun, R., and Schonenberger, M. (1979). Cultivation and *in vitro* cloning of procyclic forms of *Trypanosoma brucei* in a semi-defined medium. Acta Trop. *36*, 289-292.

Brun, R. (1980). Hydroxyurea: Effects on growth, structure and [^{3}H] thymidine uptake of *Trypanosoma brucei* procyclic culture forms. J. Protozool. *27*, 122-128.

Carr, A. M., MacNeill, S. A., Hayles, J., and Nurse, P. (1989). Molecular-cloning and sequence-analysis of mutant alleles of the fission yeast cdc2 protein-kinase gene - implications for cdc2+ protein structure and function. Mol. Gen. Genet. *218*, 41-49.

Carruthers, V. B., and Cross, G. A. M. (1992). High-efficiency clonal growth of bloodstream-form and insect- form *Trypanosoma brucei* on agarose plates. Proc. Natl. Acad. Sci. USA *89*, 8818-8821.

Chan, F. K. M., Zhang, J., Chen, L., Shapiro, D. N., and Winoto, A. (1995). Identification of human/mouse p19, a novel CDK4/CDK6 inhibitor with homology to p16^{ink4}. Mol. Cel. Biol. *15*, 2682-2688.

Clayton, C., and Michels, P (1996). Metabolic compartmentation in trypanosomes. Parasitol. Today *12*, 465-471.

Clayton, C., Adams, M., Almeida, R., Baltz, T., Barrett, M., Bastien, P., Belli, S., Beverley, S., Biteau, N., Blackwell, J., Blaineau, C., Boshart, M., Bringaud, F., Cross, G., Cruz, A., Degrave, W., Donelson, J., El-Sayed, N., Fu, G. L., Ersfeld, K., Gibson, W., Gull, K., Ivens, A., and Kelly, J. (1998). Genetic nomenclature for *Trypanosoma* and *Leishmania*. Mol. Biochem. Parasitol. *97*, 221-224.

Cohen, P. (1989). The structure and regulation of protein phosphatases. Annu. Rev. Biochem. *58*, 453-508.

Cohen, P., Holmes, C. F. B., and Tsukitani, Y. (1990). Okadaic acid: a new probe for the study of cellular regulation. Trends Biochem. Sci. *15*, 98-102.

Connel-Crowley, L., Harper, J. W., and Goodrich, D. W. (1997). Cyclin D1/Cdk4 regulates retinoblastoma protein-mediated cell cycle arrest by site-specific phosphorylation. Mol. Biol. Cell *8*, 287-301.

Coombs, G. H., Vickerman, K., Sleigh, M. A., and Warren, A. (1998). Evolutionary relationships among protozoa. Pub: Kluwer Academic Press.

Coppens, I., Baudhuin, P., Opperdoes, F. R., and Courtoy, P. J. (1988). Receptor for the low-density lipoproteins on the haemoflagellate *Trypanosoma brucei*. Purification and involvement in growth of the parasite. Proc. Natl. Acad. Sci. USA *85*, 6753-6757.

Coppens, I., Bastin, P., Courtoy, P. J., Baudhuin, P., and Opperdoes, F. R. (1991). A rapid method purifies a glycoprotein of Mr 145 000 as the LDL receptor of *Trypanosoma brucei brucei*. Biochem. Biophys. Res. Commun. *178*, 185-191.

Correa-Bordes, J., and Nurse, P. (1995). $p25^{rum1}$ orders S phase and mitosis by acting as an inhibitor of the $p34^{cdc2}$ mitotic kinase. Cell *83*, 1001-1009.

Correa-Bordes, J., Gulli, M. P., and Nurse, P. (1997). $p25^{rum1}$ promotes proteolysis of the mitotic B-cyclin $p56^{cdc13}$ during G1 of the fission yeast cell cycle. EMBO J. *16*, 4657-4664.

Coso, O. A., Anel, A. D., Martinetto, H., Muschietti, J. P., Kazanietz, M., Fraidenraich, D., Torres, H. N., and Flawia, M. M. (1992). Characterization of a g(i)-protein from *Trypanosoma cruzi* epimastigote membranes. Biochem. J. *287*, 443-446.

Costello, G., Rodgers, L., and Beach, D. (1986). Fission yeast enters the stationary phase G0 from either mitotic G1 or G2. Curr. Genet. *11*, 119-125.

Coulter, L. J., and Hide, G. (1995). *Trypanosoma brucei*: Characterisation of a life cycle stage-specific G-protein. Exp. Parasitol. *80*, 308-318.

Cox, F. E. G (1992). Systematic of parasitic protozoa. In: Parasitic Protozoa (2nd Ed) Vol 1, 55-80. Ed: Kreier, J. P., and Baker, J. R. San Deigo Academic Press.

Cox, F. E. G. (1993). Modern Parasitology (2nd Ed). Blackwell Academic Press.

Cross, F. (1988). DAF1, a mutant gene affecting size control, pheromone arrest and cell cycle kinetics of *Saccharomyces cerevisiae*. Mol. Cell. Biol. *8*, 4675-4684.

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

Das, A., Gale, M., Jr., Carter, V., and Parsons, M. (1994). The protein phosphatase inhibitor okadaic acid induces defects in cytokinesis and organellar genome segregation in *Trypanosoma brucei*. J. Cell Sci. *107*, 3477-3483.

De Luca, A., MacLachlan, T. K., Bagella, L., Dean, C., Howard, C. M., Claudio, P. P., Baldi, A., Khalili, K., and Giordano, A. (1997). A unique domain of pRb2/p130 acts as an inhibitor of Cdk2 kinase activity. J. Biol. Chem. *272*, 20971-20974.

Devoto, S. H., Mudryj, M., Pines, J., Hunter, T., and Nevins, J. R. (1992). A cyclin A-protein kinase complex possesses sequence-specific DNA binding activity: $p33^{CDK2}$ is a component of the E2F-cyclin A complex. Cell *68*, 167-176.

Draetta, G., and Beach, D. (1988). Activation of cdc2 protein kinase during mitosis in human cells: cell-cycle dependent phosphorylation and subunit rearrangement. Cell *54*, 17-26.

Drewes, G., Lichtenber-Kraag, B., Doring, F., Mandelkow, E.-M., Biernat, J., Goris, J., Doree, M., and Mandelkow, E. (1992). Mitogen activated protein (MAP) kinase transforms Tau protein into an Alzheimer-like state. EMBO J. *11*, 2131-2138.

Ducommun, B., Draetta, G., Young, P., and Beach, D. (1990). Fission yeast cdc25 is a cell-cycle regulated protein. Biochem. Biophys. Res. Commun. *167*, 301-309.

Ducommun, B., Brambilla, P., Felix, M. A., Franza, B. R., Karsenti, E., and Draetta, G. (1991). Cdc2 phosphorylation is required for its interaction with cyclin. EMBO J. *10*, 3311-3319.

Dulic, V., Lees, E., and Reed, S. I. (1992). Association of human cyclin E with a periodic G1-S phase protein kinase. Science *257*, 1958-1961.

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

Dunphy, W. G., Brizuela, L., Beach, D., and Newport, J. (1988). The *Xenopus* cdc2 protein is a component of MPF, a cytoplasmic regulator of mitosis. Cell *54*, 423-431.

Dunphy, W. (1994). The decision to enter mitosis. Trends Cell Biol. *4*, 202-212.

Durieux, P. O., Schütz, P., Brun, R., and Köhler, P. (1991). Alterations in Krebs cycle enzyme activities and carbohydrate catabolism in two strain of *Trypanosoma brucei* during in vitro differentiation of their bloodstream to procyclic stages.

Edwards, M. C., Wong, C., and Elledge, S. J. (1998). Human cyclin K, a novel RNA polymerase II-associated cyclin possessing both carboxy-terminal domain kinase and Cdk-activating kinase activity. Mol. Cell. Biol. *18*, 4291-4300.

Eid, J. E., and Sollner-Webb, B. (1991a). Homologous recombination in the tandem calmodulin genes of *Trypanosoma brucei* yields multiple products: compensation for deleterious deletions by gene amplification. Genes Dev. *5*, 2024-2032

Eid, J., and Sollner-Webb, B. (1991b). Stable integrative transformation of *Trypanosoma brucei* that occurs exclusively by homologous recombination. Proc. Natl. Acad. Sci. USA *88*, 2118-2121.

Eisenthal, R., and Cornish-Bowden, A. (1998). Prospects for antiparasitic drugs - The case of *Trypanosoma brucei*, the causative agent of African sleeping sickness. J. Biol. Chem. *273*, 5500-5505.

Eisenschlos, C., Palandini, A. A., Molina y Vedia, L., Torres, H. N., and Flawia, M. M. (1986). Evidence for the existence of an Ns-type regulatory protein in *Trypanosoma cruzi* membranes. Biochem. J. *237*, 913-917.

Elledge, S. J., and Spottswood, M. R. (1991). A new human p34 protein kinase, cdk2, identified by complementation of a CDC28 mutation in *Saccharomyces cerevisiae*, is a homolog of *Xenopus* Eg1. EMBO J. *10*, 2653-2659.

Elledge, S. J., Richman, R., Hall, F. L., Williams, R. T., Lodgson, N., and Harper, J. W. (1992). CDK2 encodes a 33 kDa cyclin A-associated protein kinase and is expressed before CDC2 in the cell cycle. Proc. Natl. Acad. Sci. USA *89*, 2907-2911.

Endicott, J. A., Nurse, P., and Johnson, L. N. (1994). Mutational analysis supports a structural model for the cell cycle protein p34. Protein Eng. *7*, 243-253.

Epstein, C. B., and Cross, F. R. (1992). CLB5: a novel B cyclin from budding yeast with a role in S phase. Genes Dev. *6*, 1695-1706.

Erondu, N. E., and Donelson, J. E. (1991). Characterization of trypanosome protein phosphatase 1 and 2A catalytic subunits. Mol. Biochem. Parasitol. *49*, 303-314.

Ersfeld, K., Docherty, R., Alsford, S., and Gull, K. (1996). A fluorescence *in situ* hybridization study of the regulation of histone messenger RNA levels during the cell-cycle of *Trypanosoma brucei*. Mol. Biochem. Parasitol. *81*, 201-209.

Espinoza, F. H., Ogas, J., Herskowitz, I., and Morgan, D. O. (1994). Cell cycle control by a complex of the cyclin HCS26 (PCL1) and the kinase PHO85. Science *266*, 1388-1391.

Espinoza, F. H., Farrell, A., Erdjument-Bromage, H., Tempst, P., and Morgan, D. O. (1996). A cyclin-dependent kinase activating kinase (CAK) in budding yeast unrelated to vertebrate CAK. Science *273*, 1714-1717.

Ewen, M. E., Sluss, H. K., Sherr, C. J., Matsushime, H., Kato, J., and Livingston, D. M. (1993). Functional interactions of the retinoblastoma protein with mammalian D-type cyclins. Cell 73, 487-497.

Evans, T., Rosenthal, E. T., Youngblom, J., Distel, D., and Hunt, T. (1983). Cyclin: a protein specified by maternal mRNA in sea urchin eggs that is destroyed at each cleavage division. Cell *33*, 389-396.

Fang, F., and Newport, J. W. (1991). Evidence that the G1-S and G2-M transitions are controlled by different cdc2 proteins in higher eukaryotes. Cell *66*, 731-742.

Fantes, P. (1979).Epistatic gene interactions in the control of division in fission yeast. Nature *279*, 428-430.

Fesquet, D., Labb, J-C., Derancourt, J., Capony, J-P., Galas, S., Girard, F., Lorca, T., Shuttleworth, J., Dor, E. M., and Cavadore, J-C. (1993). The *MO15* gene encodes the catalytic subunit of a protein kinase that activates cdc2 and other cyclin-dependent kinases (CDKs) through phosphorylation of Thr161 and its homologues. EMBO J. *12*, 3111-3121.

Field, M. C., Field, H., and Boothroyd, J. C. (1995). A homologue of the nuclear GTPase Ran/TC4 from *Trypanosoma brucei*. Mol. Biochem. Parasitol. *69*, 131-134.

Field, H., Farjah, M., Pal, A., Cuill, K., and Field, M. C. (1998). Complexity of trypanosomatid endocytosis pathways revealed by Rab4 and Rab5 isoforms in *Trypanosoma brucei*. J. Biol. Chem. *273*, 32102-32110.

Fisher, D., and Nurse, P. (1995). Cyclins of the fission yeast *Schizosaccharomyces pombe*. Semin. Cell Biol. *6*, 73-78.

Fisher, D. L., and Nurse, P. (1996). A single fission yeast mitotic cyclin B $p34^{cdc2}$ kinase promotes both S-phase and mitosis in the absence of G1 cyclins. EMBO J. *15*, 850-860.

Fisher, R. P., and Morgan, D. O. (1994). A novel cyclin associates with MO15/CDK7 to form the CDK-activating kinase. Cell *78*, 713-724.

Fitch, I., Dahmann, C., Surana, U., Amon, A., Nasmyth, K., Groetsch, L., Byers, B., and Futcher, B. (1992). Characterisation of four B-type cyclin genes of the budding yeast *Saccharomyces cerevisiae*. Mol. Biol. Cell *3*, 805-818.

Forsburg, S. L., and Nurse, P. (1991). Cell cycle regulation in the yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. Annu. Rev. Cell Biol. *7*, 227-256.

Forsburg, S. L., and Nurse, P. (1994). Analysis of the *Schizosaccharomyces pombe* cyclin puc1: Evidence for a role in cell cycle exit. J. Cell Sci. *107*, 601-613.

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

Futcher (1996). Cyclins and the wiring of the yeast cell cycle. Yeast *12*, 1635-1646.

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.

Gale, M. Jr., and Parsons, M. (1993). A *Trypanosoma brucei* gene family encoding protein kinases with catalytic domains structurally related to Nek1 and NIMA. Mol. Biochem. Parasitol. *59*, 111-122.

Gallant, P., and Nigg, E. A. (1994). Identification of a novel vertebrate cyclin: Cyclin B3 shares properties with both A- and B-type cyclins. EMBO J. *13*, 595-605.

Gallego, C., Garí, E., Colomina, N., Herrero, E., and Aldea, M. (1997). The CLN3 cyclin is down-regulated by translational repression and degradation during the G1 arrest caused by nitrogen deprivation in budding yeast. EMBO J. *16*, 7196-7206.

Gao, C. Y., and Zelenka, P. S. (1997). Cyclins, cyclin-dependent kinases and differentiation. Bioessays *19*, 307-315.

Gardiner, P. R., Finerty, J. F., and Dwyer, D. M. (1983). Iodination and identification of surface-membrane antigens in procyclic *Trypanosoma rhodesiense*. J. Immunol. *131*, 454-457.

Garriga, J., Mayol, X., and Graña, X. (1996a). The CDC2-related kinase PITALRE is the catalytic subunit of active multimeric protein complexes. Biochem. J. *319*, 293-298.

Garriga, J., Segura, E., Mayol, X., Grubmeyer, C., and Graña, X. (1996b). Phosphorylation site specificity of the CDC2-related kinase PITALRE. Biochem. J. *320*, 983-989.

Gautier, J., Solomon, M. J., Booher, R. N., Bazan, J. F., and Kirschner, M. W. (1991). CDC25 is a specific tyrosine phosphatase that directly activates $p34^{cdc2}$. Cell *67*, 197-211.

Gilmore, E. C., Ohshima, T., Goffinet, A. M., Kulkarni, A. B., and Herrup, K. (1998). Cyclin-dependent kinase 5-deficient mice demonstrate novel developmental arrest in cerebral cortex. J. Neurosci. *18*, 6370-6377.

Girard, F., Stransfeld, U., Fernandez, A., and Lamb, N. (1991). Cyclin A is required for the onset of DNA replication in mammalian fibroblasts. Cell *67*, 1169-1179.

Glasssmith, G. (1997). Characterisation of CDC2-related kinases from *Trypanosoma brucei*. Thesis, University of Glasgow.

Glotzer, M., Murray, A. W., and Kirschner, M. (1991). Cyclin is degraded by the ubiquitin pathway. Nature *348*, 132-138.

Gomez, M. L., Kornblihtt, A. R., and Tellez-Iñón, M. T. (1998). Cloning of a cdc2-related protein kinase from *Trypanosoma cruzi* that interacts with mammalian cyclins. Mol. Biochem. Parasitol. *91*, 337-351.

Gould, K. L., and Nurse, P. (1989). Tyrosine phosphorylation of the fission yeast cdc2+ protein kinase regulates entry into mitosis. Nature *342*, 39-45.

Graham, T. M., Tait, A., and Hide, G. (1997). Characterisation of a polo-like protein kinase gene homologue from an evolutionary divergent eukaryote, *Trypanosoma brucei*. Gene *207*, 71-77.

Graña, X., De Luca, A., Sang, N., Fu, Y., Claudio, P. P., Rosenblatt, J., Morgan, D. O., and Giordano, A. (1994). PITALRE, a nuclear CDC2-related protein kinase that phosphorylates the retinoblastoma protein. Proc. Natl. Acad. Sci. USA *91*, 3834-3838.

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

Guan, K-L., Jenkins, C. W., Li, Y., Nichols, M. A., Wu, X., O'Keefe, C. L., Matera, A. G., and Xiong, Y. (1994). Growth suppression by p18, a $p16^{INK4a/MTS1}$- and $p14^{INK4b/MTS2}$-related CDK6 inhibitor, correlates with wild-type pRb function. Genes Dev. *8*, 2939-2952.

Guidato, S., McLoughlin, D. M., Grierson, A. J., and Miller, C. C. (1998). Cyclin D2 interacts with cdk5 and modulates cellular cdk5/p35 activity. J. Neurochem. *70*, 335-340.

Hadwiger, J. A., Wittenberg, C., Mendenhall, M. D., and Reed, S. I. (1989a). The *Saccharomyces cerevisiae CKS1* gene, a homologue of the *Schizosaccharomyces pombe* $suc1^+$ gene, encodes a subunit of the *CDC28* protein kinase complex. Mol. Cell. Biol. *9*, 2034-2041.

Hadwiger, J. A., Wittenberg, C., Richardson, H. E., de Barros Lopes, M., and Reed, S. (1989b). A family of cyclin homologues that control the G1 phase in yeast. Proc. Natl. Acad. Sci. USA *86*, 6255-6259.

Hagan, I. M., and Hyams, J. S. (1988). The use of cell division cycle mutants to investigate the control of microtubule distribution in the fission yeast *Schizosaccharomyces pombe*. J. Cell Sci. *89*, 343-357.

Hagan, I., Hayles, J., and Nurse, P. (1988). Cloning and sequencing of the cyclin related *cdc13*$^{+}$ gene and a cytological study of its role in fission yeast mitosis. J. Cell Sci. *91*, 587-595.

Hall, D. D., Markwardt, D. D., Parviz, F., and Heideman, W. (1998). Regulation of the Cln3-Cdc28 kinase by cAMP in *Saccharomyces cerevisiae*. EMBO J. *17*, 4370-4378.

Hamaguchi, J. R., Tobey, R. A., Pines, J., Crissman, H. A., Hunter, T., and Bradbury, E. M. (1992). Requirement for p34(cdc2) kinase is restricted to mitosis in the mammalian cdc2 mutant ft210. J. Cell Biol. *117*, 1041-1053.

Hanks, S. K. (1987). Homology probing: Identification of cDNA clones encoding members of the protein-serine kinase family. Proc. Natl. Acad. Sci. USA *84*, 388-392.

Hanks, S. K. (1991). Eukaryotic protein kinases. Curr. Op. Struct. Biol. *1*, 369-383.

Hanks, S. K., and Hunter, T. (1995). The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. FASEB J. *9*, 576-596.

Hannon, G. J., and Beach, D. (1994). p15^{INK4b} is a potential effector of TGFβ-induced cell cycle arrest. Nature *371*, 257-261.

Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K., and Elledge, S. J. (1993). The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell *75*, 805-816.

Harper, J. W., Elledge, S. J., Keyomarsi, K., Dynlacht, B., Tsai, L-H., Zhang, P., Dobrowolski, S., Bai, C., Connell-Crowley, L., Swindell, E., Fox, M. P., and Wei, N. (1995). Inhibition of cyclin-dependent kinases by p21. Mol. Biol. Cell *6*, 387-400.

Harper, J. W., and Elledge, S. J. (1998). The role of Cdk7 in CAK function, a retro-retrospective. Genes Dev. *12*, 285-289.

Hartwell, L. H., Mortimer, R. K., Culotti, J., and Culotti, M. (1973). Genetic control of cell division cycle in yeast: V. Genetic analysis of cdc mutants. Genetics *74*, 267-287.

Hartwell, L. H. (1974). *Saccharomyces cerevisiae* cell cycle. Bacteriol. Rev. *38*, 164-198.

Hartwell, L. H. (1978). Cell division from a genetic perspective. J. Cell Biol. *77*, 627-637.

Hassan, P. (1999). An analysis of the CRK3 kinase of *Leishmania mexicana*. Thesis, University of Glasgow.

Hayles, J., Beach, D., Durkacz, B., and Nurse, P. (1986). The fission yeast cell cycle control gene *cdc2*: isolation of a sequence *suc1* that suppresses *cdc2* mutant function. Mol. Gen. Genet. *202*, 291-293.

Hayles, J., Fisher, D., Woollard, A., and Nurse, P. (1994). Temporal order of S phase and mitosis in fission yeast is determined by the state of the $p34^{cdc2}$-mitotic B cyclin complex. Cell *78*, 813-822.

Heichman, K. A., and Roberts, J. M. (1994). Rules to replicate by. Cell *79*, 557-562.

Henglein, B., Cheuivesse, X., Wang, J., Eick, D., and Brechot, C. (1994). Structure and cell cycle-regulated transcription of the human cyclin A gene. Proc. Natl. Acad. Sci. USA *91*, 5490-5494.

Hide, G., Graham, T., Buchanan, N., Tait, A., and Keith, K. (1994). *Trypanosoma brucei*: Characterization of protein kinases that are capable of autophosphorylation *in vitro*. Parasitology *108*, 161-166.

Hide, G., Gray, A., Harrison, C. M., and Tait, A. (1989). Identification of an epidermal growth factor receptor homologue in trypanosomes. Mol. Biochem. Parasitol. *36*, 51-60.

Hindley, J., Phear, G., Stein, M., and Beach, D. (1987). $suc1^+$ encodes a predicted 13 kDa protein that is essential for cell viability and is directly involved in the division cycle of *Schizosaccharomyces pombe*. Mol. Cell. Biol. *7*, 504-511.

Hirst, K., Fisher, F., McAndrew, P. C., and Goding, C. R. (1994). The transcription factor, the Cdk, its cyclin and their regulator: Directing the transcriptional response to a nutritional signal. EMBO J. *13*, 5410-5420.

Hoffmann, I., Clarke, P. R., Marcote, M. J., Karsenti, E., and Draetta, G. (1993). Phosphorylation and activation of human cdc25C by cdc2- cyclin B and its involvement in the self-amplification of MPF at mitosis. EMBO J. *12*, 53-63.

Holloway, S. L., Glotzer, M., King, R. W., and Murray, A. W. (1993). Anaphase is initiated by proteolysis rather than by the inactivation of maturation-promoting factor. Cell *73*, 1393-1402.

Holmes, J. K., and Solomon, M. J. (1996). A predictive scale for evaluating cyclin-dependent kinase substrates: A comparison of $p34^{cdc2}$ and $p33^{cdk2}$. J. Biol. Chem. *271*, 25240-25246.

Horne, M. C., Goolsby, G. L., Donaldson, K. L., Tran, D., Neubauer, M., and Wahls, A. F. (1996). Cyclin G1 and cyclin G2 comprise a new family of cyclins with contrasting tissue-specific and cell cycle-regulated expression. J. Biol. Chem. *271*, 6050-6061.

Hua, S., and Wang, C. C. (1994). Differential accumulation of a protein kinase homolog in *Trypanosoma brucei*. J. Cell. Biochem. *54*, 20-31.

Hua, S., Mutomba, M. C., and Wang, C. C. (1997). Regulated expression of cyclin-1 during differentiation of *Trypanosoma brucei* from bloodstream to procyclic form. Mol. Biochem. Parasitol. *84*, 255-258.

Huang, D. Q., Farkas, I., and Roach, P. J. (1996). Pho85p, a cyclin-dependent protein kinase, and the Snf1p protein kinase act antagonistically to control glycogen accumulation in *Saccharomyces cerevisiae*. Mol. Cell. Biol. *16*, 4357-4365.

Huang, D. Q., Moffat, J., Wilson, W. A., Moore, L., Cheng, C., Roach, P. J., and Andrews, B. (1998). Cyclin partners determine Pho85 protein kinase substrate specificity *in vitro* and *in vivo*: Control of glycogen biosynthesis by Pcl8 and Pcl10. Mol. Cell. Biol. *18*, 3289-3299.

Hunter, T., and Plowman, G. D. (1997). The protein kinases of budding yeast: six score and more. Trends Biochem. Sci. *22*, 18-26.

Izumi, T., Walker, D. H., and Maller, J. L. (1992). Periodic changes in phosphorylation of the *Xenopus* cdc25 phosphatase regulate its activity. Mol. Biol. Cell *3*, 927-939.

Izumi, T., and Maller, J. L. (1993). Elimination of cdc2 phosphorylation sites in the cdc25 phosphatase blocks initiation of M-phase. Mol. Biol. Cell *4*, 1337-1350.

Jackson, D. G., and Voorheis, H. P. (1990). Changes in the pattern of cell surface proteins during transformation of bloodstream forms of *Trypanosoma brucei* in vitro. Biochem. Soc. Trans. *18*, 1032-1033.

Jackson, D. G., Smith, D. K., Luo, C., and Elliott, J. F. (1993). Cloning of a novel surface antigen from the insect stages of *Trypanosoma brucei* by expression in COS cells. J. Biol. Chem. *268*, 1894-1900.

Jayaraman, P. S., Hirst, K., and Goding, C. R. (1994). The activation domain of a basic helix-loop-helix protein is mediated by repressor interaction with domains distinct from that required for transcriptional regulation. EMBO J. *13*, 2192-2199.

Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massagné, J., and Paveltich, N. P. (1995). Mechanism of CDK activation revealed by the structure of a cyclin A-CDK2 complex. *Nature* 376, 3565-3571.

Johnson, D. G., and Walker, C. L. (1999). Cyclins and cell cycle checkpoints. Annu. Rev. Pharmacol. Toxicol. *39*, 295-312.

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

Julia, V., Rassoulzadegan, M., and Glaichenhaus, N. (1996). Resistance to *Leishmania major* induced by tolerance to a single antigen. Science *274*, 421-423.

Kaldis, P., Sutton, A., and Solomon, M. J. (1996). The Cdk-activating kinase (CAK) from budding yeast. Cell *86*, 553-564.

Kaminsky, R., Nickel, B., and Holy, A. (1998). Arrest of *Trypanosoma brucei rhodesiense* and *Trypanosoma brucei brucei* in the S-phase of the cell cycle by (S)-9-(3-hydroxy-2-phosphonylmethoxypropyl)adenine ((S)-HPMPA). Mol. Biochem. Parasitol. *93*, 91-100.

Kato, J., Matsushime, H., Hiebert, S. W., Ewen, M. E., and Sherr, C. J. (1993). Direct binding of cyclin D to the retinoblastoma gene product (pRb) and pRb phosphorylation by the cyclin D-dependent kinase CDK4. Genes Dev. *7*, 331-342.

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

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.

King, R. W., Jackson, P. K., and Kirschner, M. W. (1994). Mitosis in transition. Cell *79*, 563-571.

Koff, A., Cross, F., Fisher, A., Schumacher, J., Leguellec, K., Philippe, M., and Roberts, J. M. (1991). Human cyclin E, a new cyclin that interacts with 2 members of the cdc2 gene family. Cell *66*, 1217-1228.

Koff, A., Giordano, A., Desai, D., Yamashita, K., Harper, J. W., Elledge, S., Nishimoto, T., Morgan, D. O., Franza, B. R., and Roberts, J. M. (1992). Formation and activation of a cyclin E-cdk2 complex during the G1 phase of the human cell cycle. Science *257*, 1689-1694.

Kuhne, C., and Linder, P. (1993). A new pair of B-type cyclins from *Saccharomyces cerevisiae* that functions early in the cell cycle. EMBO J. *12*, 3437-3447.

Kumagai, A., and Dunphy, W. G. (1991). The cdc25 protein controls tyrosine dephosphorylation of the cdc2 protein in a cell-free system. Cell *64*, 903-914.

Kumagai, A., and Dunphy, W. G. (1992). Regulation of the cdc25 protein during the cell cycle in *Xenopus* extracts. Cell *70*, 139-151.

Kumagai, A., and Dunphy, W. G. (1996). Purification and molecular cloning of Plx1, a Cdc25-regulatory kinase from *Xenopus* egg extracts. Science *273*, 1377-1380.

Labbé, J. C., Capony, J. P., Caput, D., Cavadore, J. C., Derancourt, J., Kaghad, M., Lelias, J. M., Picard, A., and Dorée, M. (1989). MPF from starfish oocytes at 1st meiotic metaphase is a heterodimer containing 1 molecule of cdc2 and 1 molecule of cyclin B. EMBO J. *8*, 3053-3058.

Labib, K., and Moreno, S. (1996). rum1: a CDK inhibitor regulating G1 progression in fission yeast. Trends Cell. Biol. *6*, 62-66.

LaHue, E. E., Smith, A. V., and Orr-Weaver, T. L. (1991). A novel cyclin gene from *Drosophila* complements CLN function in yeast. Genes Dev. *5*, 2166-2175.

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

LeBowitz, J. H., Coburn, C. M., McMahon-Pratt, D., and Beverley, S. M. (1990). Development of a stable *Leishmania* expression vector and application to the study of parasite surface-antigen genes. Proc. Natl. Acad. Sci. USA *87*, 9736-9740.

Lee, M. G., Norbury, C., Spurr, N. K., and Nurse, P. (1988). Regulated expression and phosphorylation of a possible mammalian cell-cycle control protein. Nature *333*, 676-678.

Lee, M. G., and Nurse, P. (1987). Complementation used to clone a human homolog of the fission yeast cell cycle control gene *cdc2*. Nature *327*, 31-35.

Lee, M. G. S., and Van der Ploeg, L. H. T. (1990). Homologous recombination and stable transfection in the parasitic protozoan *Trypanosoma brucei*. Science *250*, 1583-1587.

Lee, M. G. S., and Van der Ploeg, L. H. T. (1991). The hygromycin B resistance-encoding gene as a selectable marker for stable transformation of *Trypanosoma brucei*. Gene *105*, 255-257.

Lee, M.-H., Reynisdɢttir, I., and Massagué, J. (1995). Cloning of $p57^{KIP2}$, a cyclin-dependent kinase inhibitor with unique domain structure and tissue distribution. Genes Dev. *9*, 639-649.

Lee, K-Y., Rosales, J. L., Tang, D., and Wang, J. H. (1996). Interaction of cyclin-dependent kinase 5 (Cdk5) and neuronal Cdk5 activator in bovine brain. J. Biol. Chem. *271*, 1538-1543.

Lees, J. A., Buchkovich, K. J., Marshak, D. R., Anderson, C. W., and Harlow, E. (1991). The retinoblastoma protein is phosphorylated on multiple sites by human cdc2. EMBO J. *10*, 4279-4290.

Lees, E., Faha, B., Dulic, V., Reed, S. I., and Harlow, E. (1992). Cyclin-E/cdk2 and cyclin-A/cdk2 kinases associate with p107 and E2F in a temporally distinct manner. Genes Dev. *6*, 1874-1885.

Leopold, P., and O'Farrell, P. H. (1991). An evolutionarily conserved cyclin homolog from *Drosoplila* rescues Yeast deficient in G1 cyclins. Cell *66*, 1207-1216.

Lew, D. J., Dulic, V., and Reed, S. I. (1991). Isolation of three novel human cyclins by rescue of G1 cyclin (CLN) function in yeast. Cell *66*, 1197-1206.

Lew, J., Qi, Z., Huang, Q.-Q., Paudel, H., Matsuura, I., Matsushita, M., Zhu, X., and Wang, J. H. (1995). Structure, function, and regulation of neuronal cdc2-like protein kinase. Neurobiol. Aging *16*, 263-268.

Li, F., Hua, S. B., Wang, C. C., and Gottesdiener, K. M. (1996). Procyclic *Trypanosoma brucei* cell lines deficient in ornithine decarboxylase activity. Mol. Biochem. Parasitol. *78*, 227-236.

Li, F., Hua, S. B., Wang, C. C., and Gottesdiener, K. M. (1998). *Trypanosoma brucei brucei*: Characterization of an ODC null bloodstream form mutant and the action of alpha-difluoromethylornithine. Exp. Parasitol. *88*, 255-257.

Liao, S-M., Zhang, J., Jeffery, D. A., Koleske, A. J., Thompson, C. M., Chao, D. M., Viljoen, M., Van Vuuren, H. J. J., and Young, R. A. (1995). A kinase-cyclin pair in the RNA polymerase II holoenzyme. Nature *374*, 193-196.

Ligtenberg, M. J. L., Bitter, W., Kieft, R., Steverding, D., Janssen, H., Calafat, J., and Borst, P. (1994). Reconstitution of a surface transferrin binding complex in insect form *Trypanosoma brucei*. EMBO J. *13*, 2565-2573.

Lin, B. T. Y., Gruenwald, S., Morla, A. O., Lee, W. H., and Wang, J. Y. J. (1991). Retinoblastoma cancer suppressor gene-product is a substrate of the cell-cycle regulator cdc2 kinase. EMBO J. *10*, 857-864.

Lindberg, R. A., Quinn, A. M., and Hunter, T. (1992). Dual-specificity protein kinases: will any hydroxyl do? Trends Biochem. Sci. *17*, 114-119.

Litchfield, D. W., Lozeman, F. J., Cicirelli, M. F., Harrylock, M., Ericsson, L. H., Piening, C. J., and Krebs, E. G. (1991). Phosphorylation of the β subunit of casein kinase II in human A431 cells. J. Biol. Chem. *266*, 20380-20389.

Liu, F., Stanton, J. J., Wu, Z. Q., and Piwnica-Worms, H. (1997). The human Myt1 kinase preferentially phosphorylates Cdc2 on threonine-14 and localizes to the endoplasmic reticulum and Golgi complex. Mol. Cell. Biol. *17*, 571-583.

Lucas, R., Magez, S., Songa, B., Darji, A., Hamers, R., and De Baetselier, P. (1993). A role for TNF during African trypanosomiasis: involvement in parasite control, immunosuppression and pathology. Res. Immunol. *144*, 370-376.

Mäkelä, T. P., Tassan, J.-P., Nigg, E. A., Frutiger, S., Hughes, G. J., and Weinberg, R. A. (1994). A cyclin associated with the CDK-activating kinase MO15. Nature *371*, 254-257.

Mancebo, H., Lee, G., Flygare, J., Tomassini, J., Luu, P., Zhu, Y., Blau, C., Hazuda, D., Price, D., and Flores, D. (1997). P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro. Genes Dev. *11*, 2633-2644.

Marshall, N., Peng, J., Xie, Z., and Price, D. (1996). Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J. Biol. Chem. *271*, 27176-27183.

Martinez, A. M., Afshar, M., Martin, F., Cavadore, J. C., Labbé, J. C., and Dorée, M. (1997). Dual phosphorylation of the T-loop in cdk7: Its role in controlling cyclin H binding and CAK activity. EMBO J. *16*, 343-354.

Martín-Castellanos, C., Labib, K., and Moreno, S. (1996). B-type cyclins regulate G1 progression in fission yeast in opposition to the $p25^{rum1}$ CDK inhibitor. EMBO J. *15*, 839-849.

Matsushime, H., Ewen, M. E., Strom, D. K., Kato, J.-Y., Hanks, S. K., Roussel, M. F., and Sherr, C. J. (1992). Identification and properties of an atypical catalytic subunit ($p34^{PSKJ3}$/CDK4) for mammalian D-type G1 cyclins. Cell *71*, 323-334.

Matsushime, H., Quelle, D. E., Shurtleff, S. A., Shibuya, M., Sherr, C. J., and Kato, J.-Y. (1994). D-type cyclin-dependent kinase activity in mammalian cells. Mol. Cell. Biol. *14*, 2066-2076.

Matthews, K. R., and Gull, K. (1994a). Cycles within cycles: The interplay between differentiation and cell division in *Trypanosoma brucei*. Parasitol. Today *10*, 473-476.

Matthews, K. R., and Gull, K. (1994b). Evidence for an interplay between cell cycle progression and the initiation of differentiation between life cycle forms of African trypanosomes. J. Cell Biol. *125*, 1147-1156.

Matthews, K. R., and Gull, K. (1997). Commitment to differentiation and cell cycle re-entry are coincident but separable events in the transformation of African trypanosomes from their bloodstream to their insect form. J. Cell Sci. *110*, 2609-2618.

Matthews, K. R., and Gull, K. (1998). Identification of stage-regulated and differentiation-enriched transcripts during transformation of the African trypanosome from its bloodstream to procyclic form. Mol. Biochem. Parasitol. *95*, 81-95.

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

Maudlin, I., and Ellis, D. S. (1985). Association between rickettsia-like infections of midgut cells and susceptibility to trypanosome infection in *Glossina*. Z. Parasitenkd. *71*, 683-687.

Maudlin, I., and Welburn, S. C. (1987). Lectin mediated establishment of midgut infections of *Trypanosoma congolense* and *Trypanosoma brucei* in *Glossina morsitans*. Trop. Med. Parasitol. *38*, 167-170.

McGowan, C. H., and Russell, P. (1993). Human Wee1 kinase inhibits cell division by phosphorylating p34cdc2 exclusively on Tyr15. EMBO J. *12*, 75-85.

McGowan, C. H., and Russell, P. (1995). Cell cycle regulation of human WEE1. EMBO J. *14*, 2166-2175.

McInerny, C. J. (1997). A novel Mcm1-dependent element in the *SWI4*, *CLN3*, *CDC6* and *CDC47* promoters activates M/G1-specific transcription. Genes Dev. *11*, 1277-1288.

Measday, V., Moore, L., Ogas, J., Tyers, M., and Andrews, B. (1994). The PCL2 (ORFD)-PHO85 cyclin-dependent kinase complex: A cell cycle regulator in yeast. Science *266*, 1391-1395.

Measday, V., Moore, L., Retnakaran, R., Lee, J., Donoviel, M., Neiman, A. M., and Andrews, B. (1997). A family of cyclin-like proteins that interact with the PHO85 cyclin-dependent kinase. Mol. Cell. Biol. *17*, 1212-1223.

Melville, S. E., Leech, V., Gerrard, C. S., Tait, A., and M, B. J. (1998). The molecular karyotype of the megabase chromosomes of *Trypanosoma brucei* and the assignement of chromosome markers. Mol. Biochem. Parasitol. *94*, 155-173.

Meyerson, M., Enders, G. H., Wu, C. L., Su, L. K., Gorka, C., Nelson, C., Harlow, E., and Tsai, L. H. (1992). A family of human cdc2-related protein kinases. EMBO J. *11*, 2909-2917.

Meyerson, M., and Harlow, E. (1994). Identification of G1 kinase activity for cdk6, a novel cyclin D partner. Mol. Cell. Biol. *14*, 2077-2086.

Millar, J. B. A., McGowan, C. H., Lenaers, G., Jones, R., and Russell, P. (1991). $p80^{cdc25}$ mitotic inducer is the tyrosine phosphatase that activates $p34^{cdc2}$ kinase in fission yeast. EMBO J. *10*, 4301-4309.

Minshull, J., Golsteyn, R., Hill, C. S., and Hunt, T. (1990). The A-type and B-type cyclin associated cdc2 kinases in *Xenopus* turn on and off at different times in the cell-cycle. EMBO J. *9*, 2865-2875.

Mondesert, O., McGowan, C. H., and Russell, P. (1996). Cig2, a B-type cyclin, promotes the onset of S in *Schizosaccharomyces pombe*. Mol. Cell. Biol. *16*, 1527-1533.

Moreno, S., Hayles, J., and Nurse, P. (1989). Regulation of $p34^{cdc2}$ protein kinase during mitosis. Cell *58*, 361-372.

Moreno, S., and Nurse, P. (1994). Regulation of progression through the G1 phase of the cell cycle by the *rum1*$^{+}$ gene. Nature *367*, 236-242.

Morgan, D. O. (1995). Principles of CDK regulation. Nature *374*, 131-134.

Morgan, G. A., Laufman, H. B., Otieno-Omondi, F. P., and Black, S. J. (1993). Control of G1 to S cell cycle progression of *Trypanosoma brucei* S427c11 organisms under axenic conditions. Mol. Biochem. Parasitol. *57*, 241-252.

Morgan, G. A., Hamilton, E. A., and Black, S. J. (1996). The requirements for G1 checkpoint progression of *Trypanosoma brucei* S 427 clone 1. Mol. Biochem. Parasitol. *78*, 195-207.

Morla, A. O., Draetta, G., Beach, D., and Wang, J. Y. J. (1989). Reversible tyrosine phosphorylation of cdc2: dephosphorylation accompanies activation during entry into mitosis. Cell *58*, 193-203.

Motokura, T., Keyomarsi, K., Kronenberg, H. M., and Arnold, A. (1992). Cloning and characterization of human cyclin D3, a cDNA closely related in sequence to the PRAD1/cyclin D1 proto-oncogene. J. Biol. Chem. *267*, 20412-20415.

Mottram, J. C., Kinnaird, J., Shiels, B. R., Tait, A., and Barry, J. D. (1993). A novel CDC2-related protein kinase from *Leishmania mexicana*, lmmCRK1, is post-translationally regulated during the life-cycle. J. Biol. Chem. *268*, 21044-21051.

Mottram, J. C. (1994). cdc2-related protein kinases and cell cycle control in trypanosomatids. Parasitol. Today *10*, 253-257.

Mottram, J. C., and Smith, G. (1995). A family of trypanosome cdc2-related protein kinases. Gene *162*, 147-152.

Mottram, J. C., and Grant, K. M. (1996). *Leishmania mexicana* $p12^{cks1}$, a functional homologue of fission yeast $p13^{suc1}$, associates with a stage-regulated histone H1 kinase. Biochem. J. *316*, 833-839.

Mottram, J. C., McCready, B. P., Brown, K. P., and Grant, K. M. (1996). Gene disruptions indicate an essential function for the LmmCRK1 cdc2-related kinase of *Leishmania mexicana*. Mol. Microbiol. *22*, 573-582.

Mougneau, E., Altare, F., Wakil, A. E., Zheng, S., Coppola, T., Wang, Z-E., Waldmann, R., Locksley, R. M., and Glaichenhaus, N. (1995). Expression cloning of a protective *Leishmania* antigen. Science *268*, 563-566.

Mudryj, M., Devoto, S. H., Heibert, S. W., Hunter, T., Pines, J., and Nevins, J. R. (1991). Cell cycle regulation of the E2F transcription factor involves interaction with cyclin A. Cell *65*, 1243-1253.

Mueller, P. R., Coleman, T. R., and Dunphy, W. G. (1995a). Cell cycle regulation of a *Xenopus* Wee1-like kinase. Mol. Biol. Cell *6*, 119-134.

Mueller, P. R., Coleman, T. R., Kumagai, A., and Dunphy, W. G. (1995b). Myt1: A membrane-associated inhibitory kinase that phosphorylates Cdc2 on both threonine-14 and tyrosine-15. Science *270*, 86-90.

Murray, A. W., and Kirschner, M. W. (1989). Cyclin synthesis drives the embryonic cell cycle. Nature *339*, 275-280.

Mutomba, M. C., and Wang, C. C. (1996). Effects of aphidicolin and hydroxyurea on the cell cycle and differentiation of *Trypanosoma brucei brucei* bloodstream form. Mol. Biochem. Parasitol. *80*, 89-102.

Myler, P. J., Andleman, L., DeVos, T., Hixson, G., Kiser, P., Lemley, C., Magness, C., Rickel, E., Sisk, E., Sunkin, S., Swartzell, S., Westlake, T., Bastin, P., Fu, G., Ivens, A., and Stuart, K. (1999). *Leishmania major* Friedlin chromosome 1 has an unusual distribution of protein-coding genes. Proc. Natl. Acad. Sci. USA *96*, 2902-2906.

Nakamura, T., Sanokawa, R., Sasaki, Y. F., Ayusawa, D., Oishi, M., and Mori, N. (1995). Cyclin I: A new cyclin encoded by a gene isolated from human brain. Exp. Cell Res. *221*, 534-542.

Nash, R., Tokina, G., Anaud, S., Erickson, K., and Futcher, A. B. (1988). The *WHI-1* gene of *Saccharomyces cerevisiae* tethers cell division to cell size and is a cyclin homologue. EMBO J. *7*, 4335-4346.

Nasmyth, K. A. (1979). A control acting over the initiation of DNA replication in the yeast *Schizosaccharomyces pombe*. J. Cell. Sci. *36*, 155-168.

Nasmyth, K. A., Nurse, P., and Fraser, R. S. S. (1979). The effect of cell mass on the cell cycle timing and duration of S-phase in fission yeast. J. Cell. Sci. *39*, 215-233.

Nasmyth, K. A., and Nurse, P. (1981). Cell division mutants altered in DNA replication and mitosis in the fission yeast *Schizosaccharomyces pombe*. Mol. Gen. Genet. *182*, 119-124.

Nasmyth, K. (1993). Control of the yeast cell cycle by the CDC28 protein kinase. Curr. Op. Cell Biol. *5*, 166-179.

Nasmyth, K. A. (1996). At the heart of the budding yeast cell cycle. Trends Genet. *12*, 405-412.

Nicolic, M., Dudek, H., Kwon, Y. T., Ramos, Y. F. M., and Tsai, L.-H. (1996). The Cdk5/p35 kinase is essential for neurite outgrowth during neuronal differentiation. Genes Dev. *10*, 816-825.

Nigg, E. A., Gallant, P., and Krek, W. (1992). Regulation of $p34^{cdc2}$ protein kinase activity by phosphorylation and cyclin binding. Reg. Euk. Cell Cycle, 96.

Nigg, E. A. (1993). Targets of cyclin-dependent protein kinases. Curr. Op. Cell Biol. *5*, 187-193.

Nigg, E. A. (1995). Cyclin-dependent protein kinases: Key regulators of the eukaryotic cell cycle. Bioessays *17*, 471-480.

Nishizawa, M., Kawasumi, M., Fujino, M., and Toh (1998). Phosphorylation of Sic1, a cyclin-dependent kinase (Cdk) inhibitor, by Cdk including Pho85 kinase is required for its prompt degradation. Mol. Biol. Cell *9*, 2393-2405.

Nolan, D., Jackson, D. G., Windle, H. J., Pays, A., Genskens, M., Michel, A., Voorheis, H. P., and Pays, E. (1997). Characterisation of a novel, stage-specific invarient surface protein in *Trypanosoma brucei* containing an internal serine-rich repetitive motif. J. Biol. Chem. *272*, 29212-29221.

Nurse, P., Thuriaux, P., and Nasmyth, K. A. (1976). Genetic control of the cell division cycle in the fission yeast *Schizosaccharomyces pombe*. Mol. Gen. Genet. *146*, 167-178.

Nurse, P., and Bissett, Y. (1981). Gene required in G1 for commitment to cell cycle and in G2 for control of mitosis in fission yeast. Nature *292*, 558-560.

Nurse, P. (1985). Cell cycle controls in yeast. Trends Genet. *1*, 51-55.

Nurse, P., and Thuriaux, P. (1980). Regulatory genes controlling mitosis in the fission yeast *Schizosaccharomyces pombe*. Genetics *96*, 627-637.

Obara-Ishihara, T., and Okayama, H. (1994). A B-type cyclin negatively regulates conjugation via interacting with cell cycle 'start' genes in fission yeast. EMBO J. *13*, 1863-1872.

Ohi, R., and Gould, K. L. (1999). Regulating the onset of mitosis. Curr. Op. Cell Biol. *11*, 267-273.

Ohtsubo, M., and Roberts, J. M. (1993). Cyclin-dependent regulation of G1 in mammalian fibroblasts. Science *259*, 1908-1912.

Olssen, T., Bakhiet, M., Hojeberg, B., Ljungdahl, A., Edlund, C., Andersson, G., Ekre, H. P., Fung Leung, W. P., Mak, T., Wigzell, H., Fiszer, U., and Kristensson, K. (1993). CD8 is critically involved in lymphocyte activation by a *T. brucei brucei*-released molecule. Cell *72*, 715-727.

Oshima, T., Ward, J. M., Huh, C-G., Longenecker, G., Veeranna., Pant, H. C., Brady, R. O., Martin, L. J., and Kulkarmi, A. B. (1996). Targeted disruption of the cyclin-dependent kinase 5 gene results in abnormal corticogenesis, neuronal pathology and preinatal death. Proc. Natl. Acad. Sci. USA *93*, 11173-11178.

Oz, H. S., Huang, H., Wittner, M., Tanowitz, H. B., Bilezikian, J. P., and Morris, S. A. (1994). Evidence for guanosine triphosphate-binding proteins in *Trypanosoma cruzi*. Am. J. Trop. Med. Hyg. *50*, 620-631.

Pagano, M., Draetta, G., and Jansen, D. P. (1992). Association of Cdk2 kinase with the transcription factor E2F during S phase. Science *255*, 1144-1147.

Paindavoine, P., Rolin, S., Vanassel, 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.

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.

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. USA *90*, 2656-2660.

Parsons, M., Ledbetter, J. A., Schieven, G. L., Nel, A. E., and Kanner, S. B. (1994). Developmental regulation of pp44/46, tyrosine-phosphorylated proteins associated with tyrosine/serine kinase activity in *Trypanosoma brucei*. Mol. Biochem. Parasitol. *63*, 69-78.

Patnaik, P. K., Kulkarni, S. K., and Cross, G. A. M. (1993). Autonomously replicating single copy episomes in *Trypanosoma brucei* show unusual stability. EMBO J. *12*, 2529-2538.

Patnaik, P. K. (1997). Studies with artificial extrachromosomal elements in trypanosomatids: Could specificity in the initiation of DNA replication be linked to that in transcription? Parasitol. Today *13*, 468-471.

Pays, E., Hanocq, Q-J., Hanocq, F., van Assel, S., Nolan, D., and Rolin, S. (1993). Abrupt RNA changes precede the first cell division during the differentiation of *Trypanosoma brucei* bloodstream forms into procyclic forms *in vitro*. Mol. Biochem. Parasitol. *61*, 107-114.

Pays, E., Rolin, S., and Magez, S. (1997). Cell signalling in trypanosomatids. In Trypanosomiasis and leishmaniasis: Biology and control, G. Hide, J. C. Mottram, G. H. Coombs and P. H. Holmes, eds. (Oxford, UK: CAB International), pp. 199-225.

Pays, E., and Nolan, D. P. (1998). Expression and function of surface proteins in *Trypanosoma brucei*. Mol. Biochem. Parasitol. *91*, 3-36.

Perry, K. and Agabian, N. (1991) mRNA processing in the Trypanosomatidae. *Experientia*, 118-128.

Peter, M., Sanghera, J. S., Pelech, S. L., and Nigg, E. A. (1992). Mitogen-activated protein kinases phosphorylate nuclear lamins and display sequence specificity overlapping that of mitotic protein kinase-$p34^{cdc2}$. Eur. J. Biochem. *205*, 287-294.

Philpott, A., Porro, E. B., Kirschner, M. W., and Tsai, L. H. (1997). The role of cyclin-dependent kinase 5 and a novel regulatory subunit in regulating muscle differentiation and patterning. Genes Dev. *11*, 1409-1421.

Piggot, J. R., Rai, R., and Carter, B. L. A. (1982). A bifunctional gene product involved in two phases of the yeast cell cycle. Nature *298*, 391-393.

Pines, J., and Hunter, T. (1989). Isolation of a human cyclin cDNA: Evidence for cyclin mRNA and protein regulation in the cell cycle and for interaction with $p34^{cdc2}$. Cell *58*, 833-846.
Pines, J., and Hunter, T. (1990). $p34^{cdc2}$: The S and M kinase? New Biol. *2*, 389-401.

Pines, J. (1993). Cyclin-dependent kinases: Clear as crystal? Curr. Biol. *3*, 544-547.

Pines, J. (1995). Cyclins and cyclin-dependent kinases: A biochemical view. Biochem. J. *308*, 697-711.

Polymenis, M., and Schmidt, E. V. (1997). Coupling of cell division to cell growth by translational control of the G1 cyclin *CLN3* in yeast. Genes Dev. *11*, 2522-2531.

Poon, R. Y. C., Yamashita, K., Adamczewski, J. P., Hunt, T., and Shuttleworth, J. (1993). The cdc2-related protein $p40^{MO15}$ is the catalytic subunit of a protein kinase that can activate $p33^{cdk2}$ and $p34^{cdc2}$. EMBO J. *12*, 3123-3132.

Prescott, D. M. (1976). Reproduction of Eukaryotic Cells. New York Academic Press.

Priest, J. W., and Hadjuk, S. L. (1994). Developmental regulation of mitochondrial biogenesis in *Trypanosoma brucei*. J. Bioenerg. Biomembr. *26*, 179-191.

Pringle, J. (1978). The use of conditional lethal cell cycle mutants for temporal and functional sequence mapping of cell cycle events. J. Cell. Physiol. *93*, 393-406.

Pringle, J. R., and Hartwell, L. H. (1981). The *Saccharomyces cerevisiae* cell cycle. In: Strathern, J. N., Jones, E. W., and Broach, J. R. (1981) *The Molecular Biology of the Yeast Saccharomyces, I. Life cycle and inheritance*. New York: Cold Spring Harbor Lab.

Reed, S. I. (1980). Selection of *S. cerevisiae* mutants defective in the start event of cell division. Genetics *95*, 561-577.

Reed, S. I., Hadwiger, J. A., and Lorincz, A. T. (1985). Protein kinase activity associated with the product of the yeast cell division cycle gene *CDC28*. Proc. Natl. Acad. Sci. USA *82*, 4055-4059.

Reed, S. I., and Wittenberg, C. (1990). Mitotic role for the CDC28 protein kinase of *Saccharomyces cerevisiae*. Proc. Natl. Acad. Sci. USA *87*, 5697-5701.

Reiner, C. L., Borras, A. M., Kurdistani, S. K., Garrean, J. R., Chung, M., Aaronson, S. A., and Lee, S. W. (1999). Altered regulation of cyclin G in human breast cancer and its specific localisation at replication foci in response to DNA damage in $p53^{+/+}$ cells. J. Biol. Chem. *274*, 11022-11029.

Richardson, H. E., Lew, D. J., Henze, M., Sugimoto, K., and Reed, S. I. (1992). Cyclin B homologs in *Saccharomyces cerevisiae* function in S phase and in G2. Genes Dev. *6*, 2021-2034.

Rickert, P., Seghezzi, W., Shanahan, F., Cho, H., and Lees, E. (1996). Cyclin C/CDK8 is a novel CTD kinase associated with RNA polymerase II. Oncogene *12*, 2631-2640.

Robinson, D. R., and Gull, K. (1991). Basal body movements as a mechanism for mitochondrial genome segregation in the trypanosome cell cycle. Nature *352*, 731-733.

Robinson, D. R., Sherwin, T., Ploubidou, A., Byard, E. H., and Gull, K. (1995). Microtubule polarity and dynamics in the control of organelle positioning, segregation, and cytokinesis in the trypanosome cell cycle. J. Cell Biol. *128*, 1163-1172.

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 Overath, P. (1989). Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of *Trypanosoma brucei*. J. Cell Biol. *108*, 737-746.

Rolin, S., Halleux, S., Vansande, J., Dumont, J., Pays, E., and Steinert, M. (1990). Stage-specific adenylate cyclase activity in *Trypanosoma brucei*. Exp. Parasitol. *71*, 350-352.

Rolin, S., Hanocq-Quertier, J., and Paturiaux-Hanocq, F. (1996). Simultaneous but independent activation of adenylate cyclase and glycosylphosphatidylinositol-phospholipase C under stress conditions in *Trypanosoma brucei*. J. Biol. Chem. *271*, 10844-10852.

Rosenberg, A. R., Zindy, F., Le Deist, F., Mouly, H., Métézeau, P., Bréchot, C., and Lamas, E. (1995). Overexpression of human cyclin A advances entry into S phase. Oncogene *10*, 1501-1509.

Rosenblatt, J., Gu, Y., and Morgan, D. O. (1992). Human cyclin-dependent kinase 2 is activated during the S and G2 phases of the cell cycle and associates with cyclin A. Proc. Natl. Acad. Sci. USA *89*, 2824-2828.

Rossignol, M., Kolb-Cheynel, I., and Egly, J. M. (1997). Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIIH. EMBO J. *16*, 1628-1637.

Roy, R., Adamczewski, J. P., Seroz, T., Vermeulen, W., Tassan, J-P., Schaeffer, L., Nigg, E. A., Hoeijmakers, J. H. J., and Egly, J-M. (1994). The MO15 cell cycle kinase is associated with the TFIIH transcription-DNA repair factor. Cell *79*, 1093-1101.

Russell, P., and Nurse, P. (1986). *cdc25*$^{+}$ functions as an inducer in the mitotic control of fission yeast. Cell *45*, 145-153.

Russell, P., and Nurse, P. (1987a). Negative regulation of mitosis by *wee1*$^{+}$, a gene encoding a protein kinase homologue. Cell *49*, 559-567.

Russell, P., and Nurse, P. (1987b). The mitotic inducer *nim1*$^{+}$ functions in a regulatory network of protein kinase homologues controlling the initiation of mitosis. Cell *49*, 569-596.

Russell, P., Moreno, S., and Reed, S. I. (1989). Conservation of mitotic controls in budding and fission yeasts. Cell *57*, 295-303.

Salmon, D., Genskens, M., Hanocq, F., Hanocq-Quertier, J., Nolan, D., Ruben, L., and Pays, E. (1994). A novel heterodimeric transferrin receptor encoded by a pair of VSG expression site-associated genes in *Trypanosoma brucei*. Cell *78*, 75-86.

Salmon, D., Hanocq-Quertier, J., Paturiaux-Hanocq, F., Pays, A., Tebabi, P., Nolan, D. P., Michel, A., and Pays, E. (1997). Characterization of the ligand-binding site of the transferrin receptor in *Trypanosoma brucei* demonstrates a structural relationship with the N-terminal domain of the variant surface glycoprotein. EMBO J. *16*, 7272-7278.

Sambrook, J., Fritsch, E. J., and Maniatis, T. (1989). A Laboratory Manual of Molecular Cloning (2nd Ed). New York, Cold Spring Harbour.

Sanchez-Diaz, A., Gonzalez, I., Arellano, M., and Moreno, S. (1998). The Cdk inhibitors $p25^{rum1}$ and $p40^{sic1}$ are functional homologues that play similar roles in the regulation of the cell cycle in fission and budding yeast. J. Cell Sci. *111*, 843-851.

Schell, D., Evers, R., Preis, D., Ziegelbauer, K., Kiefer, H., Lottspeich, F., Cornelissen, A. W. C. A., and Overath, P. (1991). A transferrin-binding protein of *Trypanosoma brucei* is encoded by one of the genes in the variant surface glycoprotein gene- expression site. EMBO J. *10*, 1061-1066.

Schneider, K. R., Smith, R. L., and O'Shea, E. K. (1994). Phosphate-regulated inactivation of the kinase PHO80-PHO85 by the CDK inhibitor PHO81. Science *266*, 122-126.

Schneider, B. L., Yang, Q. H., and Futcher, A. B. (1996). Linkage of replication to start by the Cdk inhibitor Sic1. Science *272*, 560-562.

Schulman, B. A., Lindstrom, D. L., and Harlow, E. (1998). Substrate recruitment to cyclin-dependent kinase 2 by a multipurpose docking site on cyclin A. Proc. Natl. Acad. Sci. USA *95*, 10453-10458.

Schwob, E., and Nasmyth, K. (1993). *CLB5* and *CLB6*, a new pair of B cyclins involved in DNA replication in *Saccharomyces cerevisiae*. Genes Dev. *7*, 1160-1175.

Schwob, E., Böhm, T., Mendenhall, M. D., and Nasmyth, K. (1994). The B-type cyclin kinase inhibitor $p40^{SIC1}$ controls the G1 to S transition in *S. cerevisiae*. Cell *79*, 233-244.

Selbie, L. A., Schmitz-Pfeiffer, C., Sheng, Y., and Biden, T. J. (1993). Molecular cloning and characterisation of PKCι, an atypical isoform of the protein kinase C derived from insulin-secreting cells. J. Biol. Chem. *268*, 24296-24302.

Serrano, M., Hannon, G. J., and Beach, D. (1993). A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4. Nature *366*, 704-707.

Sherman, D. R., Janz, L., Hug, M., and Clayton, C. (1991). Anatomy of the parp gene promoter of *Trypanosoma brucei*. EMBO J. *10*, 3379-3386.

Sherr, C. J. (1993). Mammalian G1 cyclins. Cell *73*, 1059-1065.

Sherr, C. J. (1994). G1 phase progression: cycling on cue. Cell *79*, 551-555.

Sherr, C. J., and Roberts, J. M. (1999). CDK inhibitors: positive and negative regulators of G1-phase progression. Genes Dev. *13*, 1501-1512.

Sherwin, T., and Gull, K. (1989). The cell-division cycle of *Trypanosoma brucei brucei*: timing of event markers and cytoskeletal modulations. Philos. Trans. R. Soc. Lond. Ser. B *323*, 573-588.

Sicinski, P., Donaher, J. L., Parker, S. B., Li, T., and Fazeli, A. (1995). Cyclin D1 provides a link between development and oncogenesis in the retina and breast. Cell *82*, 621-630.

Sleigh, M A. (1989). Protozoa and other protists. London, Edward Arnold.

Solomon, M. J., Glotzer, M., Lee, T. H., Philippe, M., and Kirschner, M. (1990). Cyclin activation of $p34^{cdc2}$. Cell *63*, 1013-1024.

Solomon, M. J. (1993). Activation of the various cyclin/cdc2 protein kinases. Curr. Op. Cell. Biol. *5*, 180-186.

Sorger, P. K., and Murray, A. W. (1992). S-phase feedback control in budding yeast independent of tyrosine phosphorylation of $p34^{CDC28}$. Nature *355*, 365-368.

Souza, A. E., Bates, P. A., Coombs, G. H., and Mottram, J. C. (1994). Null mutants for the *lmcpa* cysteine proteinase gene in *Leishmania mexicana*. Mol. Biochem. Parasitol. *63*, 213-220.

Sterner, D. E., Lee, J. M., Hardin, S. E., and Greenleaf, A. L. (1995). The yeast carboxyl-terminal repeat domain kinase CTDK-I is a divergent cyclin-cyclin-dependent kinase complex. Mol. Cell. Biol. *15*, 5716-5724.

Steverding, D., Stierhof, Y. D., Fuchs, H., Tauber, R., and Overath, P. (1995). Transferrin-binding protein complex is the receptor for transferrin uptake in *Trypanosoma brucei*. J. Cell Biol. *131*, 1173-1182.

Steverding, D., Stierhof, Y-D., Chaudhri, M., Ligtenberg, M., Schell, D., Beck-Sickinger, A. G., and Overath, P. (1994). ESAG 6 and 7 products of *Trypanosoma brucei* form a transferrin binding protein complex. Eur. J. Cell Biol. *64*, 78-87.

Surana, U., Robitsch, H., Price, C., Schuster, T., Fitch, I., Futcher, A. B., and Nasmyth, K. (1991). The role of CDC28 and cyclins during mitosis in the budding yeast *Saccharomyces cerevisiae*. Cell *65*, 145-161.

Swindle, J., and Tait, A. (1996). Trypanosomatid genetics. In: Molecular biology of parasitic protozoa, D. F. Smith and M. Parsons, eds. (Oxford: IRL Press), pp. 6-34.

Tachado, S. D., and Schofield, L. (1994). Glycosylphosphatidylinositol toxin of *Trypanosoma brucei* regulates IL-1 and TNF-α expression in macrophages by protein tyrosine kinase mediated signal transduction. Biochem. Biophys. Res. Commun. *205*, 984-991.

Tait, A., and Turner, C. M. R. (1990). Genetic exchange in *Trypanosoma brucei*. Parasitol. Today *6*, 70-75.

Tamura, K., Kanaoka, Y., Jinno, S., Nagata, A., Ogiso, Y., Shimizu, K., Hayakawa, T., Nojima, H., and Okayama, H. (1993). Cyclin G: A new mammalian cyclin with homology to fission yeast Cig1. Oncogene *8*, 2113-2118.

Tassan, J. P., Jaquenoud, M., Fry, A. M., Frutiger, S., Hughes, G. J., and Nigg, E. A. (1995a). *In vitro* assembly of a functional human CDK7-cyclin H complex requires MAT1, a novel 36 kDa RING finger protein. EMBO J. *14*, 5608-5617.

Tassan, J. P., Jaquenoud, M., Léopold, P., Schultz, S. J., and Nigg, E. A. (1995b). Identification of human cyclin-dependent kinase 8, a putative protein kinase partner for cyclin C. Proc. Natl. Acad. Sci. USA *92*, 8871-8875.

Ten Asbroek, A. L. M. A., Ouellette, M., and Borst, P. (1990). Targeted insertion of the neomycin phosphotransferase gene into the tubulin gene cluster of *Trypanosoma brucei*. Nature *348*, 174-175.

Ten Asbroek, A. L. M. A., Mol, C. A. A. M., Kieft, R., and Borst, P. (1993). Stable transformation of *Trypanosoma brucei*. Mol. Biochem. Parasitol. *59*, 133-142.

Tennyson, C. N., Lee, J., and Andrews, B. J. (1998). A role for the Pcl9-Pho85 cyclin-cdk complex at the M/G1 boundary in *Saccharomyces cerevisiae*. Mol. Microbiol. *28*, 69-79.

Tetley, L., Turner, C. M. R., Barry, J. D., Crowe, J. S., and Vickerman, K. (1987). Onset of expression of the variant surface glycoproteins of *Trypanosoma brucei* in the tsetse fly studied using immunoelectron microscopy. J. Cell Sci. *87*, 363-372.

Thorner, J. (1981). Pheromonal regulation of development in *Saccharomyces cerevisiae*. In: Strathern, J. N., Jones, E. W., and Broach, J. R. (1981) *The Molecular Biology of the Yeast Saccharomyces, I. Life cycle and inheritance*. New York: Cold Spring Harbor Lab.

Thuret, J, Y., Valay, J. G., Faye, G., and Mann, C. (1996). Civ1 (CAK *in vivo*), a novel Cdk-activating kinase. Cell *86*, 565-576.

Timblin, B. K., Tatchell, K., and Bergman, L. W. (1996). Deletion of the gene encoding the cyclin-dependent protein kinase Pho85 alters glycogen metabolism in *Saccharomyces cerevisiae*. Genetics *143*, 57-66.

Torruella, M., Flawia, M. M., Eisenschlos, C., Vedia, L. M. Y., Rubinstein, C. P., and Torres, H. N. (1986). *Trypanosoma cruzi* adenylate-cyclase activity: purification and characterization. Biochem. J. *234*, 145-150.

Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. Proc. Natl. Acad. Sci. USA *76*, 4340-4354.

Toyoshima, H., and Hunter, T. (1994). p27, a novel inhibitor of G1 cyclin-Cdk protein kinase activity, is related to p21. Cell *78*, 67-74.

Tsai, L. H., Harlow, E., and Meyerson, M. (1991). Isolation of the human cdk2 gene that encodes the cyclin A- and adenovirus E1A-associated p33 kinase. Nature *353*, 174-177.

Turner, C. M. R. (1992). Cell-cycle coordination in trypanosomes. Parasitol. Today *8*, 3-4.

Tyres, M., Tokina, G., and Futcher, B. (1993). Comparison of the *Saccharomyces cerevisiae* G1 cyclins: Cln3 may be an upstream activator of Cln1, Cln2 and other cyclins. EMBO J. *12*, 1955-1968.

Valay, J. G., Simon, M., Dubois, M. F., Bensaude, O., Facca, C., and Faye, G. (1995). The *KIN28* gene is required both for RNA polymerase II mediated transcription and phosphorylation of the Rbp1p CTD. J. Mol. Biol. *249*, 535-544.

Van Assel, S., and Steinert, M. (1971). Nuclear and kinetoplastic DNA replication cycles in normal and synchronously dividing *Crithidia luciliae*. Exp. Cell. Res. *65*, 353-358.

Van den Abbeele, J., Rolin, S., Claes, Y., Le Ray, D., Pays, E., and Coosemans, M. (1995). *Trypanosoma brucei*: Stimulation of adenylate cyclase by proventriculus and esophagus tissue of the tsetse fly, *Glossina morsitans morsitans*. Exp. Parasitol. *81*, 618-620.

Van den Heuvel, S., and Harlow, E. (1993). Distinct roles for cyclin-dependent kinases in cell cycle control. Science *262*, 2050-2054.

Van Hellemond, J. J., Neuville, P., Schwartz, R. J., Matthews, K., and Mottram, J. C. (1999). Isolation of *Trypanosoma brucei* CYC2 and CYC3 cyclin genes by rescue of a yeast G1 cyclin mutant. Functional characterisation of CYC2. In prep.

Vassella, E., and Boshart, M. (1996). High-molecular mass agarose matrix supports growth of bloodstream forms of pleomorphic *Trypanosoma brucei* strains in axenic culture. Mol. Biochem. Parasitol. *82*, 91-105.

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.

Vickerman, K. (1985). Developmental cycles and biology of pathogenic trypanosomes. Br. Med. Bull. *41*, 105-114.

Vickerman, K., Tetley, L., Hendry, K. A. K., and Turner, C. M. R. (1988). Biology of African trypanosomes in the tsetse fly. Biol. Cell *64*, 109-119.

Voorheis, H. P., and Martin, B. R. (1981). Characterisation of the calcium-mediated mechanism activating adenylate cyclase in *T. brucei*. Eur. J. Biochem. *116*, 471-477.

Voorheis, H. P., and Martin, B. R. (1982). Local anaesthetics including benzyl alcohol activte adenylate cyclase in *T. brucei* by a calcium-dependent machanism. Eur. J. Biochem. *123*, 371-376.

Walter, R. D., and Opperdoes, F. R. (1982). Subcellular distribution of adenylate cyclase, cAMP phosphodiesterase, protein kinases and phosphoprotein phosphatase in *Trypanosoma brucei*. Mol. Biochem. Parasitol. *6*, 287-295.

Wei, P., Garber, M. E., Fang, S. M., Fischer, W. H., and Jones, K. A. (1998). A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell *92*, 451-462.

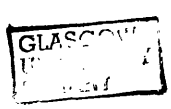
Weinberg, R. A. (1995). The retinoblastoma protein and cell cycle control. Cell *81*, 323-330.

Wiese, M. (1998). A mitogen-activated protein (MAP) kinase homologue of *Leishmania mexicana* is essential for parasite survival in the infected host. EMBO J. *17*, 2619-2628.

Wirtz, E., and Clayton, C. E. (1995). Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. Science *268*, 1179-1183.

Wittenberg, C., and Reed, S. I. (1988). Control of the yeast-cell cycle is associated with assembly and disassembly of the CDC28 protein kinase complex. Cell *54*, 1061-1072.

Woodward, R., and Gull, K. (1990). Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*. J. Cell Sci. *95*, 49-57.



Xiong, Y., and Beach, D. (1991). Human D-type cyclin. Curr. Biol. *1*, 362-364.

Xiong, Y., Zhang, H., and Beach, D. (1992). D type cyclins associate with multiple protein kinases and the DNA replication and repair factor PCNA. Cell *71*, 505-514.

Xiong, Y., Hannon, G. J., Zhang, H., Casso, D., Kobayashi, R., and Beach, D. (1993). p21 is a universal inhibitor of cyclin kinases. Nature *366*, 701-704.

Yang, X., Gold, M. O., Tang, D. N., Lweis, D. E., Aguilar-Cordova, E., Rice, A. P., and Herrmann, C. H. (1997). TAK, an HIV Tat-associated kinase, is a member of the cyclin-dependent family of protein kinases and is induced by activation of peripheral blood lymphocytes and differentiation of promonocytic cell lines. Proc. Natl. Acad. Sci. USA *94*, 12331-12336.

Yang, X., Herrmann, C., and Rice, A. (1996). The human immunodeficiency virus Tat proteins specifically associate with TAK in vivo and require the carboxyl-terminal domain of RNA polymerase II for function. J. Virol. *70*, 4576-4584.

Yankulov, K. Y., and Bentley, D. L. (1997). Regulation of CDK7 substrate specificity by MAT1 and TFIIH. EMBO J. *16*, 1638-1646.

Ziegelbauer, K., Stahl, B., Karas, M., Stierhof, Y.-D., and Overath, P. (1993). Proteolytic release of cell surface proteins during differentiation of *Trypanosoma brucei*. Biochemistry *32*, 3737-3742.