# A study of two protein kinases from *Trypanosoma brucei*

Timothy Marcus Graham

PhD

Wellcome Unit of Molecular Parasitology
Anderson College
University of Glasgow

May 1996



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

**Summary**

The protozoan parasite *Trypanosoma brucei* undergoes major differentiation events during its complex life cycle which involves the tsetse fly and mammal as obligate hosts. At the same time it alternates between proliferative and non-proliferative forms. In higher eukaryotes differentiation and the cell cycle are controlled by complex signalling networks many of which involve protein kinases as components and, by analogy, this would be expected to be the situation in *T. brucei*. However, very little is known about cellular signalling in this parasite and the work presented in this thesis is a study of two *T. brucei* protein kinases as an approach to the identification of pathways regulating differentiation and the cell cycle.

Two approaches were followed: firstly, the characterisation and purification of a 60 kDa autophosphorylating protein kinase from *T. brucei* that was found to be expressed in a stage-specific manner (Hide *et al.*, 1994). This protein kinase activity was found to localise in the 100 000 g supernatant of total extracts of bloodstream forms and this supernatant was used as the starting material for purification. Using anion-exchange chromatography as the first purification step, the autophosphorylating protein kinase was detected in two peaks of activity eluting at slightly different salt concentrations. A number of different chromatography matrices were then tested for their suitability in the further purification of the protein kinase and a further level of purification was achieved using a Sepharose to which the protein kinase inhibitor H-9 had been immobilised. In addition new assays for measuring activity were developed, one of which, based on autophosphorylation in solution was found to be robust and informative.

The second approach taken was to isolate and sequence a full length cDNA corresponding to an amplified cDNA fragment (Hua and Wang, 1994) for which the predicted peptide sequence showed homology to catalytic domain sequences of a rat protein kinase C family member and the *Drosophila* protein kinase polo. Isolation and sequencing of genes encoding *T. brucei* protein kinase C family members would provide valuable evidence for cell signalling in this parasite and provide tools to complement earlier studies on protein kinase C like activities in *T. brucei* (Keith *et al.*, 1990). A clone was isolated in a screen of a λgt11 cDNA library using a probe homologous to the original amplified cDNA fragment. Partial sequencing of the insert has shown that it is a chimaera of cDNAs for a ribosomal protein S4 homologue and for a protein kinase (this

part of the chimaera includes a section identical to the probe). Hybridisation of Southern blots of *T. brucei* genomic DNA to probes from the ribosomal protein and protein kinase coding regions, under high stringency conditions, shows that both sequences are of *T. brucei* origin but that the two sequences are not co-linear in the genome. Comparisons of the two predicted peptide sequences with protein sequences in the databases show that the open reading frames for both proteins are incomplete. The open reading frame for the protein kinase homologue is probably almost complete but there is no 3′ stop codon. The 5′ coding sequence could be complete but there could be an upstream start codon not contained within the clone isolated. The putative protein kinase has all the defining features of a member of the polo-like kinase family and is clearly not a protein kinase C. Polo-like kinases are implicated in regulation of mitotic spindle formation and therefore mitosis, and the *T. brucei* enzyme would be worthy of further study given that mitosis and spindle formation are very different in *T. brucei* compared to higher eukaryotes.

**Declaration**

This thesis and the results presented in it are entirely my own work except where otherwise stated.

T. Graham

19/9/96

**Acknowledgements**

I would firstly like to thank my supervisors Dr G. Hide and Professor A. Tait for advice and encouragement throughout this project and for speedy correction of drafts of this thesis. I would also like to thank all those in WUMP, Vet. Parasitology and elsewhere who helped and encouraged me during this work. This work was supported by the Medical Research Council.

CONTENTS

## List of Figures :

## List of tables :

## List of Abbreviations :

| | |
|---|---|
| 2-D: | 2-dimensional |
| ADP: | Adenosine diphosphate |
| AGC: | AGC group of protein kinase families including the cyclic nucleotide-dependent protein kinase family (PKA and PKG) and the protein kinase C (PKC) family |
| AMP: | Adenosine monophosphate |
| AMPK: | AMP-dependent protein kinase |
| ATP: | Adenosine triphosphate |
| AUFS: | Absorbance units full scale |
| CaMK: | a) $Ca^{2+}$/calmodulin-dependent protein kinase<br>b) CaMK group of protein kinase families including the family of protein kinases regulated by $Ca^{2+}$/calmodulin, the Snf1/AMPK family, and other close relatives |
| cAMP: | cyclic AMP |
| CDK: | Cyclin dependent kinase |
| cDNA: | complementary DNA |
| CDPK: | Calcium-dependent protein kinase |
| cGMP: | cyclic GMP |
| CK2: | Casein kinase 2 |
| CMGC: | CMGC group of protein kinase families including the cyclin-dependent protein kinase family, the mitogen-activated protein kinase (MAP kinase) family, the glycogen synthase kinase 3 (GSK3) family and the Cdk-like (Clk) protein kinase family |
| CRK: | cdc2 related kinase |
| DAG: | diacylglycerol |
| DEAE: | Diethylaminoethyl |
| DFMO: | difluoromethylornithine |
| DNA: | Deoxyribonucleic acid |
| dNTP: | deoxynucleotide triphosphate |
| DTT: | Dithiothreitol |

| | |
|---|---|
| EATRO: | East African Trypanosomiasis Research Organisation |
| EDTA: | Ethylenediamine tetra-acetic acid (disodium salt) |
| EGF: | Epidermal growth factor |
| EGTA: | Ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid |
| ESAG: | Expression site associated gene |
| FGF: | Fibroblast growth factor |
| G+C: | guanosine + cytosine |
| GPI: | glycosyl-phosphatidylinositol |
| gRNA: | guide RNA |
| GSK-3: | Glycogen synthase kinase 3 |
| GTP: | Guanosine triphosphate |
| HIC: | Hydrophobic interaction chromatography |
| HPLC: | High performance liquid chromatography |
| IEF: | Isoelectric focusing |
| IGF: | Insulin-like growth factor |
| IL-1: | Interleukin 1 |
| ISG: | Invariant surface glycoprotein |
| KRP: | kinesin related protein |
| LDL: | Low-density lipoprotein |
| LMP: | Low melting point |
| MAPK: | Mitogen-activated protein kinase |
| MAPKK: | Mitogen-activated protein kinase kinase |
| MLCK: | Myosin light chain kinase |
| mRNA: | messenger RNA |
| NAD: | nicotinamide adenine dinucleotide |
| NADH: | nicotinamide adenine dinucleotide (reduced form) |
| NP-40: | Nonidet P-40 |
| PARP: | Procyclic acidic repetitive protein |
| PBSG: | Phosphate buffered saline containing 1% glucose |
| PCR: | Polymerase chain reaction |
| PDGF: | Platelet derived growth factor |
| PH2: | Polo-homology 2 |

| | |
|---|---|
| pI: | Isoelectric point |
| PKA: | Protein kinase A |
| PKC: | Protein kinase C |
| PKG: | Protein kinase G |
| polI: | RNA polymerase I |
| poly(A)$^+$: | polyadenylated |
| pp: | Phosphoprotein |
| PP: | Protein phosphatase |
| PS: | phosphatidylserine |
| PTK: | Protein tyrosine kinase |
| PTP: | Protein tyrosine phosphatase |
| PVDF: | Polyvinylidene difluoride |
| RACE: | Rapid amplification of cDNA ends |
| RNA: | Ribonucleic acid |
| RPC: | Reversed-phase chromatography |
| RPS4: | Ribosomal Protein S4 |
| SDS-PAGE: | Sodium dodecyl sulphate polyacrylamide gel electrophoresis |
| SDS: | Sodium dodecyl sulphate |
| SL RNA: | Spliced leader RNA |
| TbPLK: | *T. brucei* polo-like kinase |
| TbS4lp: | *T. brucei* S4-like protein |
| TEMED: | Tetra-methyl-1,2-diaminoethane |
| TetR: | tetracycline repressor |
| TGF: | Transforming growth factor |
| TLC: | Thin layer chromatography |
| TREU: | Trypanosomiasis Research Edinburgh University |
| Tris: | Tris (hydroxymethyl) amino methane |
| TX-100: | Triton X-100 |
| UV: | Ultraviolet |
| VSG: | Variant surface glycoprotein |

# Chapter One

# Introduction

## 1.1 *Trypanosoma brucei*

*Trypanosoma brucei* is a single-celled protozoan belonging to the Order Kinetoplastida, the defining feature of this Order being the kinetoplast, a specialised structure containing the mitochondrial DNA. *T. brucei* belongs to the Salivaria which are digenetic parasites transmitted by feeding tsetse flies. The other Section of *Trypanosoma*, the Stercoraria, which includes *Trypanosoma cruzi* (the agent of South American trypanosomiasis), are only very distantly related to *T. brucei.* The early evolutionary history of *T. brucei* and other trypanosomatids is currently controversial and it is not clear whether the ancestors of *T. brucei* first became parasitic in an invertebrate or vertebrate host (reviewed by Vickerman, 1994; Maslov and Simpson, 1995).

The classification of the Salivarian trypanosomes is summarised in Fig. 1.1. The *T. brucei* subspecies *T. b. brucei*, *T. b. gambiense* and *T. b. rhodesiense* belong to the subgenus Trypanozoon and are morphologically indistinguishable. They have traditionally been identified on the basis of their host range, the severity of the disease they cause in man, their infectivity to laboratory rodents and their geographical distribution. *T. b. brucei* which is found throughout the tsetse belt of Africa causes the disease nagana in cattle and also infects wild animals, but does not infect man. *T. b. gambiense* infects man causing chronic disease, infects laboratory rodents poorly and is found in West and Central Africa. *T. b. rhodesiense* infects man causing acute disease, but also infects wild and domestic animals as well as laboratory rodents and is found in East Africa (Hoare, 1972). However, many examples are known where these criteria have proved unsatisfactory and more recently techniques such as isoenzyme and DNA analysis have been used to try to define these subspecies (e.g. Hide *et al.*, 1990; reviewed by Gibson, 1986). Among the other subgenera of Salivaria, *T. congolense* is particularly important as a pathogen of cattle. Human African trypanosomiasis occurs in 36 African countries in which a total of over 50 million people are at risk and about 25,000 new cases are reported annually, although this is probably an underestimate (Kuzoe, 1993). As well as causing disease in both humans and livestock, the threat from trypanosome-infected tsetse flies severely restricts the use of land that would otherwise be suitable for agriculture.

*Trypanosoma brucei* is among the best characterised protozoan parasites. As well as its importance as a pathogen, *T. brucei* is of interest as an organism that diverged

very early from the main eukaryotic lineage (Sogin *et al.*, 1989) and as a single cell that undergoes major differentiation events in alternating between mammalian and insect hosts. The life cycle of *T. brucei* species is outlined briefly in Section 1.1.1. In subsequent sections some of the unusual structures and processes of the trypanosome cell are described in more detail.

### 1.1.1 Life cycle

Fig. 1.2 (from Vickerman, 1985) shows the developmental cycle of *Trypanosoma brucei* species. Following injection of non-dividing metacyclic forms into the dermal tissue of the mammalian host by the feeding tsetse fly, the trypanosomes enter the draining lymphatics and bloodstream. The trypanosomes differentiate into the slender bloodstream forms which divide with a doubling time of about 6 hours. The slender bloodstream forms are dependent on uptake of glucose from the host which they convert to pyruvate by glycolysis. The mitochondrion in these forms is small and lacks a functional tricarboxylic acid cycle. The bloodstream forms are vulnerable to the host immune system and antibody recognition results in parasite killing. However, the surface of the cell in the bloodstream stages is completely covered by a coat of variant surface glycoprotein (VSG) that in dividing bloodstream forms can be replaced by an antigenically different variant. In the simplest possible situation, during the time that parasitaemia is increasing, the majority of the population may express a single type of VSG (the "homotype") to which the host mounts an immune response. As the antibody response becomes effective against the homotype the parasite population falls dramatically but one of the unrecognised "heterotypes", previously expressed by only a few cells, outgrows the others to give rise to a new homotype. Therefore a chronic infection can develop with regular waves of parasitaemia. However, in most cases the situation is probably considerably more complex (reviewed by Barry and Turner, 1991).

The continuance of infection in the mammal depends on the cycle described above, but towards the end of each peak of parasitaemia a significant proportion of the population will consist of intermediate and short stumpy cells that are non-dividing as well as the dividing population of slender forms. These have a less elongate flagellum than the slender forms, and the mitochondrion is larger and has enzymes that will be needed in the insect procyclic stage where proline is the parasite's major energy source.

The differentiation process that produces these forms is not well understood but appears to be irreversible (Matthews and Gull, 1994a; Hamm *et al.*, 1990). Cell division is probably required for VSG switching (Matthews and Gull, 1994a) so stumpy cells that remain within the mammalian host will eventually be cleared by the immune system due to their inability to change their surface coat. It is generally accepted that the stumpy forms represent a pre-adaptation for transmission to the tsetse fly, but the production of non-dividing stumpy forms may have a further advantage to the parasite as a way of limiting the population within the host, thereby reducing the danger of early host death. Stumpy forms can be produced in experimental infections in immunosuppressed animals (Balber, 1972), suggesting that VSG-specific antibodies cannot be the trigger for their production, although the production of stumpy forms may well involve an exogenous factor (Seed and Sechelski, 1989).

Blood ingested by the feeding tsetse passes through to the midgut where stumpy bloodstream forms transform into the procyclic form. The bloodmeal in the insect midgut is separated from the midgut epithelium by the chitinous peritrophic membrane (reviewed by Miller and Lehane, 1993) and differentiation of the trypanosomes occurs in the space between the membrane and the midgut wall. Whether the trypanosomes pass into the space at the posterior (open) end of the membrane or pass directly through the anterior membrane is currently controversial (Maudlin and Welburn, 1994). Death is likely to be the normal fate for slender bloodstream forms taken up in the bloodmeal although, *in vitro*, slender forms can differentiate through to established procyclics (Bass and Wang, 1991; Matthews and Gull, 1994b). The differentiation events that occur include the loss of the VSG coat (Overath *et al.*, 1983; Ziegelbauer *et al.*, 1990; Ziegelbauer *et al.*, 1993) and the appearance of a new invariant surface glycoprotein known as procyclin or PARP (Roditi *et al.*, 1989; Roditi and Pearson, 1990). Body length also increases and the mitochondrion, partly activated in the stumpy form, undergoes further changes in structure and enzyme complement. The kinetoplast also changes its position from the extreme posterior of the cell to lie close to the nucleus. Concomitant with these differentiation events is re-entry into a dividing cell cycle (Matthews and Gull, 1994a+b) with the kinetoplast repositioning apparently occurring during S phase (Matthews *et al.*, 1995) but not requiring DNA synthesis (Matthews and Gull, 1994b).

After a few days procyclic cells begin to move to the ectoperitrophic space, the

route taken in returning to the midgut again being controversial. This space becomes packed with parasites which continue to migrate towards the proventriculus. Proventricular forms are even more elongate than the early procyclics and have a smaller mitochondrial volume. How these cells then reach the salivary gland is not well understood (see review by Ellis, 1986), but proliferative epimastigote cells in the salivary glands attach themselves through the flagellum to the microvilli of the cells lining the lumen (Tetley and Vickerman, 1985). Genetic exchange occurs in the tsetse but is not an obligatory part of the life cycle (reviewed by Sternberg and Tait, 1990; Tait and Turner, 1990). Further differentiation processes eventually produce the mature metacyclic forms which are non-dividing and have a VSG coat, although these VSGs are distinct from those in the bloodstream forms (Graham *et al.*, 1990). This is the form that is infective to the mammal.

For experimental purposes bloodstream form *T. brucei* can be obtained from infected laboratory rodents. Stocks described as pleomorphic produce infections apparently similar to those of trypanosomes in their natural hosts. Many stocks that have been highly passaged have lost the ability to produce short stumpy forms and are described as monomorphic; these stocks rapidly kill their hosts due to very high parasitaemias of dividing slender forms. It is possible to induce some of these stocks to produce morphologically stumpy forms by the administration of the trypanocide DFMO to the rodent host. The differentiation of slender bloodstream or short stumpy bloodstream forms through to procyclic forms can be induced by incubation at 27 °C in medium containing citrate and cis-aconitate (Brun and Schönenberger, 1981; Bass and Wang, 1991). Cells morphologically and biochemically similar to dividing procyclic forms can be satisfactorily grown in culture.

### 1.1.2 Genome and gene expression

*T. brucei* is diploid with a haploid genome size estimated at 3-4 x $10^7$ bp (Tait, 1980; Borst *et al.*, 1980; Borst *et al.*, 1982). The chromosomes do not undergo condensation at any life cycle stage but can be analysed by pulsed field gel electrophoresis and three size classes of chromosomes can be recognised. The large chromosomes range in size from 800 kb to 6 Mb (Van der Ploeg *et al.*, 1989), the intermediate chromosomes from 150 kb to 700 kb and minichromosomes from 50 kb to 150 kb (Van der Ploeg *et al.*,

1984a). The diploid genome contains at least eighteen of these large chromosomes and probable homologous pairs have been identified, although these are not all of the same size (Gottesdiener *et al.*, 1990; Bastian *et al.*, 1995). The number of minichromosomes in *T. brucei* was estimated at about 100 (Van der Ploeg *et al.*, 1984b), while *T. vivax* has very small numbers (Dickin and Gibson, 1989). The function of minichromosomes seems to be as a reservoir of VSG genes (Van der Ploeg *et al.*, 1984b). The number of intermediate sized chromosomes (which also contain VSG genes) is small, but variable, and rearrangements in these chromosomes have been detected (Van der Ploeg *et al.*, 1984c). Intermediate and minichromosomes do not show Mendelian inheritance (see Sternberg and Tait, 1990).

The mitochondrial (kinetoplast) genome of *T. brucei*, as of other kinetoplastids, is highly unusual and consists of a concatenated network of DNA circles (Simpson, 1987; Chen *et al.*, 1995). In *T. brucei* there are 25-50 maxicircles and 5000-10000 minicircles. The 22 kb maxicircle encodes at least 13 proteins and the 9S and 12S mitochondrial RNAs, while the minicircles are about 1 kb in size and probably do not encode proteins. The production of a number of the mature mitochondrial mRNAs involves a process known as RNA editing in which maxicircle transcripts (precursor mRNAs) are modified by the insertion or deletion of uridines. Transcripts from both maxicircles and minicircles, known as guide RNAs (gRNAs), act as templates directing these reactions (Seiwert and Stuart, 1994). Differential RNA editing between life cycle stages could contribute to the developmental regulation of the mitochondrion (Priest and Hajduk, 1994).

Nuclear gene structure and expression in *T. brucei* is markedly different from that of higher eukaryotes (see reviews by Pays *et al.*, 1994; Graham, 1995; Vanhamme and Pays, 1995). The great majority of *T. brucei* genes for which genomic organisation and production of primary transcripts has been studied in detail are clustered in the genome in polycistronic transcription units with each unit being under the control of a 5′ promoter. The only known exceptions to this rule are the transcription units for metacyclic VSGs which appear to be either monocistronic or contain a single gene in addition to the metacyclic VSG gene, and are probably transcriptionally regulated (Graham and Barry, 1991; Graham *et al.*, 1993). The genes of each polycistronic transcription unit are separated by intergenic regions, but *T. brucei* protein coding genes

do not contain introns. The processing of polycistronic precursor RNAs into monocistronic mRNAs involves the trans-splicing of a leader sequence (which caps the mRNA) at the 5′ end of each message (Parsons *et al.*, 1984; DeLange *et al.*, 1984) and the addition of a 3′ poly(A) tail. The spliced leader sequence added to the 5′ end of each precursor RNA is 39 bp long but originates from a 141 bp SL RNA (Milhausen *et al.*, 1984). Trans-splicing is not unique to kinetoplastids but they are unusual in processing all pre-mRNAs in this way. Polyadenylation in *T. brucei* is directed by a very different mechanism to that of higher eukaryotes. Polyadenylation of one gene appears to be linked to the addition of the spliced leader to the next gene downstream in the transcription unit, and is at least partly directed by polypyrimidine tracts in the intergenic region (Hug *et al.*, 1994). Gene expression in *T. brucei* has been most intensively studied for the bloodstream VSG genes and for the procyclic PARP genes, and these are the only protein-coding genes for which promoters have been identified. However, it is not clear to what extent the control mechanisms identified for these transcription units are typical, especially since the polymerases that transcribe the surface antigen genes are α-amanitin resistant (like pol1), while most *T. brucei* genes are transcribed by α-amanitin sensitive polymerases. An important consequence of the organisation of *T. brucei* genes into polycistronic transcription units is that the rate of transcriptional initiation at the promoter cannot result in differential expression of genes within a transcription unit. However, many examples of differential expression of genes from within an individual transcription unit are known. It seems likely that in many cases the regulation is at the level of mRNA maturation or stability, although evidence has also been presented for the regulation of gene expression at the levels of translation (Gale *et al.*, 1994b) and protein stability (Torri *et al.*, 1993).

In summary both genome structure and the regulation of gene expression are unusual in *T. brucei* compared to higher eukaryotes. The great majority of *T. brucei* genes appear to be clustered in polycistronic transcription units but differential expression of the genes within a transcription unit is common and *T. brucei* has to radically change the expression of a large number of genes in undergoing differentiation between life cycle stages. The expression of a number of genes, and particularly those encoding the bloodstream and procyclic surface antigens, has been studied in some detail but little is known about the DNA and RNA binding factors that must be involved in

regulating gene expression.

### 1.1.3 Energy metabolism

*T. brucei* bloodstream forms have access to an abundant supply of glucose and convert this into pyruvate at an extremely high rate. A high glycolytic rate is essential because there is no Krebs cycle or mitochondrial ATP-synthesising respiratory chain (at least in slender forms) and only 2 moles of ATP are produced per mole of glucose consumed. The pyruvate produced is excreted into the host bloodstream. In the procyclic forms proline is the preferred substrate (Evans and Brown, 1972) and the mitochondrion is active. Acetate and succinate are the metabolic products excreted (reviewed by Turrens, 1991). The glycolytic enzymes of higher eukaryotes are cytoplasmic but, in all kinetoplastids that have been examined, the majority of the glycolytic enzymes are contained in a unique organelle, the glycosome. The glycosomes of *T. brucei* contain nine enzymes responsible for the conversion of glucose to 3-phosphoglycerate and glycerol, and in the bloodstream stage these account for over 90% of the organelle's protein content (reviewed by Hannaert and Michels, 1994). The glycolytic reactions occurring in the glycosome result in no net ATP gain or loss but net ATP production occurs in the cytoplasm in the pyruvate kinase reaction. Under aerobic conditions the NADH produced is reoxidised via the glycerol-3-phosphate/di-hydroxyacetone phosphate shuttle linking glycolysis to a mitochondrial alternative oxidase without ATP synthesis, while under anaerobic conditions glycerol is excreted as well as pyruvate. The compartmentalisation of glycolytic enzymes in glycosomes, which occupy less than 5% of cellular volume, probably partly explains how bloodstream *T. brucei* sustain such high glycolytic rates. Glycosomes have a number of features in common with peroxisomes and glyoxysomes found in other eukaryotes (Sommer and Wang, 1994) and Michels and Hannaert (1994) have proposed that these microbodies may have a common endosymbiont ancestor.

In the procyclic insect stage a number of the glycosomal glycolytic enzymes are repressed and reoxidation of NADH produced by glycolysis uses a different pathway involving the fixation of $CO_2$ in a reaction catalysed by phosphoenolpyruvate carboxykinase. The mitochondrion in the procyclic insect stage is cristate and has both cytochrome-mediated electron transport using a conventional cytochrome c oxidase and

a complete citric acid cycle (see reviews by Turrens, 1991; Priest and Hajduk, 1994). Energy metabolism in the salivary gland forms is not well understood, but it is thought that mitochondrial activity is down-regulated in the mature metacyclics. The short stumpy bloodstream forms metabolise glucose to pyruvate and reoxidise NADH using the same pathway as long slender cells, but pyruvate dehydrogenase and α-ketoglutarate dehydrogenase are found while the mitochondrion is larger than in the slender forms and is cristate. Evidence has been presented that stumpy forms are able to generate an electromotive force across the inner mitochondrial membrane using NADH dehydrogenase which might be used for ATP production (Bienen *et al.*, 1991). These changes are consistent with the short stumpy forms representing a pre-adaptation for transmission to the insect.

### 1.1.4 Cytoskeleton and cell cycle, cell surface, and flagellar pocket

The morphology of the procyclic form of *T. brucei* has been examined in some detail (see Sherwin and Gull, 1989a+b; Robinson and Gull, 1991; Robinson *et al.*, 1995). *T. brucei* has a well defined shape which is determined by a subpellicular "corset" of microtubules running along the length of the cell. This cytoskeleton is very stable compared to that of many eukaryotes and the microtubules are cross-linked both to each other and to the cell membrane. It does not undergo extensive disassembly during the cell cycle and cell division in *T. brucei* is not symmetrical. The *T. brucei* cell cycle is also unusual in that as well as the nucleus the flagellum, basal body and mitochondrion are present in single copies and therefore have to be segregated accurately at cell division.

Fig. 1.3 panel A (from Sherwin and Gull, 1989a) summarises schematically the morphological changes occurring during the procyclic cell cycle as determined by electron and fluorescent microscopy (Sherwin and Gull, 1989a). Cells at the earliest stage of the cell cycle (I) have a single flagellum emerging from the cell body at a posterior position, with a pro-basal body lying next to the mature flagellar basal body. The nucleus is round in shape and is in the centre of the cell, while the kinetoplast is adjacent to the basal body. The next stage (II) is marked by the development of the pro-basal body which lies adjacent and posterior to the mature basal body. This is followed by the appearance of two new pro-basal bodies and the partial elongation of the daughter flagellum from the posterior basal body (stage III). The daughter flagellum continues to

elongate and, in the stage IV cell, its anterior end extends out of the flagellar pocket and possesses a paraflagellar rod (which runs alongside the flagellar axoneme) but has not reached its final length. In stage V the two basal bodies begin to move apart. Mitosis then begins while the nucleus still lies anterior to both basal bodies (VI), but the elongate nucleus later comes to lie across the axis of the parent flagellum (VII). Stage VIII is marked by the completion of mitosis with the separated nuclei occupying very specific positions within the cell. A cleavage furrow then begins to form running towards the posterior of the dividing cell (IX) which elongates until the daughter cells are connected only at their posterior ends (X) before their final separation. It has been shown that the kinetoplast is physically linked to the basal body and that the movement apart of the old and new basal bodies is responsible for kinetoplast segregation (Robinson and Gull, 1991). In a study of various morphological parameters of the procyclic cell cycle Robinson *et al.* (1995) found that all of the increase in cell length occurring during cell cycle progression could be accounted for by an increase in the inter basal body distance (see Fig. 1.3 panel B). A number of other aspects of the cellular structure and cell cycle of *T. brucei* have also been examined including the elongation, nucleation and polarity of microtubules within the subpellicular array and flagellum (Sherwin and Gull, 1989b; Robinson *et al.*, 1995), the repositioning the kinetoplast during *in vitro* differentiation of stumpy forms to procyclic forms (Matthews *et al.*, 1995) and the timing of cell cycle events (Woodward and Gull, 1990). In summary a number of the basic features of the morphology of the procyclic form of *T. brucei* have been quite well studied and many of the findings described above are also likely to apply to other life cycle stages. The cell surface and flagellar pocket of the bloodstream forms of *T. brucei* are described in the next few paragraphs, in relation to the selective pressures imposed by the mammalian immune system.

The cell surface of *T. brucei* is completely covered in a dense coat of VSG glycoproteins in the bloodstream and metacyclic stages. The VSG proteins consist of about 500 amino acids of which the first 20-40 amino acids are a signal sequence, the next 360 amino acids are the variable region, while the C-terminus of 110-120 amino acids is less variable (see Blum *et al.*, 1993). The signal peptide and the C-terminal 17-23 amino acids are cleaved off in the mature protein and a glycoconjugate, the GPI anchor, is added at the C-terminus (GPI anchors are reviewed by McConville and Ferguson,

1993 and Ferguson, 1994). In the membrane the VSG molecules associate as homodimers with the N-terminal end facing outwards and the GPI anchor inserted in the membrane bilayer. It is thought that the glycan part of the anchor lies across the membrane to form a "glycocalyx" which may play a part in determining the barrier properties of the surface coat. While the surface of the bloodstream *T. brucei* cell is covered by the VSG coat, bloodstream *T. brucei* are highly dependent on uptake of nutrients such as glucose (ter Kuile, 1993) and low-density lipoprotein (Coppens *et al.*, 1988), and these functions are known to involve specific cell surface receptors of which at least part of the extracellular domains must have invariant structure essential to ligand binding. Because of this, these receptors have to be "hidden" from the host immune system. There is evidence that at least some invariant surface proteins fail to provoke a powerful immune response by virtue of being hidden beneath the protruding and tightly packed VSG proteins. Two proteins (ISG 65 and ISG 75) which are probably the most abundant surface proteins after VSG, are dispersed over the cell surface and, in spite of large predicted extracellular domains, are not readily accessible to antibodies (Overath *et al.*, 1994). The idea that the lack of immune reactivity is due to the packing of the surface coat is consistent with the finding that the accessibility of these proteins to the small protease trypsin varied between clones expressing different VSGs. In spite of the shielding of invariant surface proteins beneath the VSG coat, it is possible that in a chronic infection an immune response to invariant surface proteins can be mounted. One strain of *T. brucei* has been found to have seven genes for ISG 75 with two isoforms being co-expressed. Overath *et al.* (1994) have suggested that a possible reason for this might be to keep the concentration of identical ISG 75 molecules at a low level thereby reducing the chance of bivalent antibody binding. Slightly different variants of surface molecules such as nutrient receptors might also be expressed at different times and most of the genes that are co-transcribed with VSG genes (expression-site associated genes - ESAGs) encode known or putative surface proteins with the corresponding ESAGs of different expression sites being slightly variant. Borst (1991) proposed that the differences between corresponding ESAGs in different expression sites represents a mechanism for varying the cell surface in addition to and alongside VSG switching.

The greatest variety and concentration of invariant surface molecules are located in a particular cellular structure known as the flagellar pocket. This is the sole site of

endocytosis in *T. brucei* (reviewed by Webster and Russell, 1993) and is buried in the cell body at the base of the flagellum. Individual proteins and protein complexes from the host have access to the walls of the flagellar pocket, but its structure ensures that while antibodies have access to proteins located in the pocket, antibodies bound at this site will not be able to stimulate macrophage killing of the parasite. In addition any antibodies that do bind will be rapidly endocytosed and so only a very strong antibody response is likely to activate host effector mechanisms (such as complement) or kill the parasite by preventing nutrient uptake. Bastin *et al.* (1994) found that one monoclonal antibody raised against a fragment of the *T. brucei* LDL-receptor could cause trypanosome lysis through activation of complement and its binding was inhibited by LDL. The flagellar pocket also appears to be the site targeted by a component of human serum that kills *T. b. brucei* species (Hajduk *et al.*, 1994; Smith *et al.*, 1995).

### 1.1.5 Chemotherapy of African trypanosomiasis

The current strategies for prevention and treatment of African trypanosomiasis are unsatisfactory. No vaccines have been developed that can defeat the ability of *T. brucei* to vary its surface coat, although as described above, some experiments in mice suggest that a vaccine against flagellar pocket components may not be completely out of the question (Bastin *et al.*, 1994). It might also be possible to develop a vaccine blocking transmission from the mammalian host to the tsetse since *T. brucei* ingested by the insect will differentiate to procyclic forms that express a number of invariant surface antigens while still exposed to the blood meal. A number of drugs are available for the treatment of African trypanosomiasis but these all have significant side effects.

The treatment of African trypanosomiasis has been recently reviewed (Pépin and Milord, 1994; Jernigan and Pearson, 1993). The aromatic diamidine pentamidine has been used in treating human trypanosomiasis for over 50 years and is still an important trypanocide. Its mode of action is unknown although a number of enzymes have been suggested as targets (see Berger *et al.*, 1995). In *in vitro* experiments, *T. brucei* was found to rapidly accumulate the drug to a level much higher than the surrounding medium but the parasites were only killed relatively slowly (Berger *et al.*, 1995). Pentamidine has been found to be effective in treating early-stage *T. b. gambiense* disease, especially in combination with suramin, but it is less effective in treating *T. b.*

*rhodesiense*. A number of side-effects of pentamidine treatment can occur including hypotension, vomiting, rashes, and hypoglycaemia and when it was used on a large scale as a chemoprophylactic the estimated death rate due to the drug was 1 per 300,000. Another aromatic diamidine, diminazene, may be an effective treatment for disease caused by both *T. b. gambiense* and *T. b. rhodesiense* if used in the early stages of disease (see references in Pépin and Milord, 1994), but very little data is available on its effectiveness in humans.

Suramin is a sulphated naphthylamine which, like pentamidine, was developed more than half a century ago. Its mechanism of action is poorly understood since it inhibits a large number of enzymes (see Pépin and Milord, 1994) and it is unclear which of these are relevant to its action as a trypanocide. Suramin has two clusters of negative charges about 4 nm apart and it has been proposed that its inhibition of the glycosomal phosphoglycerate kinase is due to its interaction with two positively charged clusters in separate domains of the enzyme (Hart *et al.*, 1989). Similar electrostatic interactions might explain its effects on other enzymes. It binds plasma proteins and persists in the body for some time. Suramin is used for early stage rhodesiense disease for which it is more effective than pentamidine but is also used in combination with pentamidine for treating early stage gambiense disease. Like pentamidine, it is not effective in treating late-stage disease when the parasites have entered the central nervous system, although the combination of suramin and eflornithine has been tested for the treatment of late-stage rhodesiense disease (Kuzoe, 1993). The possible side effects of suramin include fever, renal toxicity, and very occasionally anaphylactic shock.

Melarsoprol (Mel B) is a trivalent arsenical and is the most powerful trypanocide available. The predominant trypanocidal action of melarsoprol seems to be through its strong binding to trypanothione, a low molecular weight thiol compound unique to the kinetoplastids, which is involved in regulating intracellular thiol/disulphide redox balance (Henderson and Fairlamb, 1987; Fairlamb *et al.*, 1989 Cunningham *et al.*, 1994). However, exactly how interference with trypanothione metabolism kills the trypanosome is not fully understood. The importance of melarsoprol is in treating the later stages of disease where it has the advantage of being able to clear parasites from the central nervous system. This may be due more to its strongly trypanocidal activity than to efficient crossing of the blood-brain barrier. Before the availability of melarsoprol there

was no drug effective against the late stages of either form of human trypanosomiasis and the great majority of patients with infection in the central nervous system died. Unfortunately, melarsoprol has severe side effects, the most serious of which is an invariably fatal reactive encephalopathy occurring in 5-10% of cases. It is possible that some side effects of melarsoprol are due to immune responses to antigens released when the parasites die and it may be possible to reduce the likelihood of encephalopathy by co-administration of other drugs (see Pépin and Milord, 1994), but this increases the complexity and cost of treatment. Because of its severe side effects the use of melarsoprol is avoided if possible, although for late-stage rhodesiense disease there is currently no other drug available.

Eflornithine (DFMO) reduces trypanosome polyamine synthesis by irreversibly inhibiting ornithine decarboxylase (see Schechter and Sjoerdsma, 1986; Wang, 1991). Administration of putrescine, spermidine, and spermine with the drug to experimentally infected animals prevents its curative effect. The drug crosses the blood-brain barrier and is effective against early and late-stage gambiense trypanosomiasis, often proving successful in patients infected with parasites resistant to arsenicals. Limited trials of eflornithine against late stage *T. b. rhodesiense* were not encouraging but its use in combination with suramin for treating this form of the disease has been investigated (see Kuzoe, 1993). The drug has unpleasant side-effects but these are generally much less severe than those caused by arsenicals. A serious barrier to its widespread use is its high cost (~ $500 per patient) and melarsoprol is still the drug most commonly used to treat the later stages of disease.

In summary, the currently available drugs for the treatment of African trypanosomiasis are unsatisfactory. When trypanosomiasis is diagnosed early the combination of suramin and pentamidine for treating *T. b. gambiense* disease or suramin monotherapy for treating *T. rhodesiense* can be used (Pépin and Milord, 1994) but are quite toxic. However, the situation is much worse when the disease has spread to the central nervous system since one of the two drugs available, melarsoprol, is extremely toxic while the other, eflornithine, is too expensive for most African countries to be able to afford and is less effective against *T. b. rhodesiense*. Unfortunately, the diagnosis of African trypanosomiasis in its early stages is difficult and therefore there is a great need for new drugs effective against the later stages of disease.

### 1.1.6 Summary

*T. brucei* has been the subject of intensive research and many of the basic features of its biology have been elucidated. Unusual features of its biology which are now reasonably well understood include polycistronic transcription and the processing of precursor RNAs by trans-splicing, the genetic control of VSG switching (see Pays *et al.*, 1994), the structural organisation of the cell surface and underlying cytoskeleton, the compartmentalisation of glycolysis in glycosomes, the presence of the mitochondrial DNA as a concatenate of maxicircles and minicircles and the process of RNA editing. The differentiation events that *T. brucei* undergoes in adapting to different host environments have also received considerable research attention. This is particularly true for the transition from stumpy or slender bloodstream forms to the procyclic form for which both differentiation protocols and markers for following the progress of differentiation have been developed. However, very little is known about what signals initiate differentiation *in vivo* and how these are processed by the cell. By analogy with higher eukaryotes it would be expected that *T. brucei* might utilise signalling pathways involving protein phosphorylation in controlling such events. In higher eukaryotes the cell cycle is also controlled by regulatory networks involving protein phosphorylation and again by analogy the same might be true in *T. brucei*. As in many higher eukaryotes cellular differentiation and cell cycle progression can be controlled coordinately in *T. brucei* and both of the transmission competent forms (stumpies and metacyclics) are cell cycle arrested (see Mottram, 1994). Protein kinases in *T. brucei* are the general theme of the present work. However, since very little is known about protein kinases and protein phosphatases in *T. brucei*, while protein phosphorylation has been intensively researched in other biological systems, the next section is a general introduction to the field of protein phosphorylation in eukaryotes.

## 1.2 Protein phosphorylation in eukaryotes

Protein phosphorylation was first recognised as a mechanism for regulating enzyme activity in the 1950's in work on the control of glycogen synthesis and breakdown. In characterising an enzyme that converts muscle glycogen phosphorylase from an inactive *b* form to an active *a* form, Krebs and Fischer (1956) found that its activity was

dependent on ATP and $Mg^{2+}$ or $Mn^{2+}$ ions. When they used radiolabelled ATP in the reaction mixture, activation of glycogen phosphorylase was accompanied by its concomitant phosphate labelling. They also found that another enzyme, that inactivated glycogen phosphorylase, removed the radiolabel incorporated during activation. Since these early experiments, protein phosphorylation has been found to be one of the most common post-translational protein modifications that occurs in eukaryotes. Enzymes that catalyse the transfer of phosphate groups from a nucleotide donor to amino acids in proteins are known as protein kinases while enzymes catalysing the removal of such phosphate groups are known as protein phosphatases. The effects of phosphorylation on target proteins varies and multi-site phosphorylation can lead to combinatorial effects. In the case of a bacterial isocitrate dehydrogenase, which is inhibited by phosphorylation at a single serine, it has been shown that the phosphorylated residue is a substrate binding residue and that enzyme inhibition is due only to local effects (Hurley *et al.*, 1990). However, for glycogen phosphorylase, the activating phosphorylation occurs away from the active site and causes significant conformational shifts in the enzyme dimers (Barford and Johnson, 1989). In higher eukaryotes, at least, protein phosphorylation is involved in regulating all kinds of cellular activities and many protein kinases and phosphatases are linked into networks. This control mechanism, involving distinct forward and backward reactions, has particular advantages where a protein's activity has to be switched independent of large changes in the concentration of an effector molecule binding to that protein (see review by Hardie, 1995). The three major roles for protein phosphorylation pathways in higher eukaryotes are in the processing of extracellular signals such as hormone binding, the control of the cell cycle and the coordination of cellular responses to environmental stress.

Cloning of the genes for a large number of protein kinases and comparisons between the predicted peptide sequences has shown that nearly all known eukaryotic protein kinases belong to a single superfamily in which members are known to phosphorylate proteins on either serine, threonine, or tyrosine residues, or a combination of these. There are eukaryotic protein kinases that are not members of this main superfamily, but the term "protein kinase" will be used to mean a member of this superfamily unless otherwise stated. In mammals it has been estimated that the protein kinase genes collectively make up at least 1% of the coding capacity of the genome

(Hunter, 1987; Hanks and Hunter, 1995a+b) and that there may also be several hundred protein phosphatase genes (Hunter, 1995).

Protein phosphatases are not considered below except in Section 1.3 (Protein phosphorylation in *T. brucei*). Briefly, the original classification of protein serine/threonine phosphatases was on the basis of biochemical properties (reviewed by Cohen, 1989), but it is now clear that protein phosphatases of classes 1, 2A, and 2B belong to the same superfamily while protein phosphatase 2C is either unrelated or very distantly related to PP1, PP2A, and PP2B (see review by Wera and Hemmings, 1995). Protein tyrosine phosphatases (PTPs) were reviewed by Stone and Dixon (1994), and dual-specificity protein phosphatases by Keyse (1995). Recently, crystal structures were published for mammalian PP1 (Goldberg *et al.*, 1995), PP2B (Griffith *et al.*, 1995) and PTP1B (Barford *et al.*, 1994).

### 1.2.1 Protein kinase classification

Protein kinases cannot satisfactorily be classified on the basis of their substrate specificity since most protein kinases will phosphorylate a number of substrates which may or may not be physiologically relevant. Protein kinases have been grouped in a number of different ways, but now that many protein kinase sequences are available a phylogeny based on sequence comparisons can be used (see below). A more general system that includes all protein kinases of both eukaryotic and prokaryotic origin uses specificity for the acceptor amino acid as the basis for classification (see Hunter, 1991). The separate categorisation of protein kinases into those that phosphorylate on serine/threonine (E.C. 2.7.10), and those that phosphorylate on tyrosine (E.C. 2.7.11) is now out of date as a number of protein kinases are known that phosphorylate on serine/threonine and tyrosine (dual-specificity protein kinases). However, the conventional protein tyrosine kinases remain a valid protein kinase family, the members of which phosphorylate exclusively tyrosine residues, possess particular catalytic domain sequence motifs (see Section 1.2.2), and cluster as a group in catalytic domain phylogenetic analyses (Hanks and Hunter, 1995a+b). In addition to the protein kinases that phosphorylate on serine, threonine or tyrosine residues, a eukaryotic example of another class of protein kinases, defined originally in prokaryotes and known as "two-component systems", has been identified in *S. cerevisiae* (Maeda *et al.*, 1994; Chang *et al.*, 1993; reviewed by Swanson

*et al.*, 1994). In these enzymes, an environmental stimulus causes a histidine kinase domain to autophosphorylate. The phosphate is then transferred to an aspartate in another domain, resulting in the activation of downstream events. While these kinases are unrelated to the main family of eukaryotic protein kinases, the *S. cerevisiae* two-component system is linked to a signalling cascade involving conventional protein kinases (Maeda *et al.*, 1994).

Hanks *et al.* (1988) produced a phylogenetic tree of the protein kinase superfamily based on catalytic domain peptide sequence comparisons and this has since been updated to include new protein kinase sequences (Hanks and Quinn, 1991; Hanks and Hunter, 1995a+b). In their latest analysis, which was performed using a parsimony method (see Hanks and Hunter, 1995a+b for more details), 99 protein kinase catalytic domain peptide sequences were used in the construction of a "skeleton" tree in which only representative members were included from each of four major groups of protein kinases. Separate trees were then constructed for these four major groups that included a larger number of sequences. The skeleton tree is reproduced in Fig. 1.4 with the four major groups of protein kinases that were subsequently analysed in more detail indicated by dashed boxes. The numbers indicate the number of amino acid substitutions required to reach hypothetical common ancestors while the asterisks indicate recognised protein kinase families. Original references for each of the sequences used in the phylogenetic analysis can be found in Hanks and Hunter (1995a). The AGC, CaMK, CMGC, and PTK groups of protein kinases which account for a large proportion of the known protein kinases and the more closely related families of protein kinases within them are described briefly below.

The first group of protein kinases (labelled as the AGC group) includes the cyclic-nucleotide-dependent protein kinases (PKA and PKG) and the protein kinase C (PKC) family. The cyclic-nucleotide binding domains of PKA (cAMP-dependent) and PKG (cGMP-dependent) are related to each other (Shabb and Corbin, 1992) but in PKA the cAMP-binding domains and the protein kinase catalytic domains are located in different polypeptides (Corbin *et al.*, 1978) while in PKG, two cGMP-binding domains are contained within the same polypeptide as the protein kinase catalytic domain (Takio *et al.*, 1984). The protein kinase C enzymes are best known as protein kinases activated by diacylglycerol and calcium released after agonist-induced stimulation of phospholipase

C (Nishizuka, 1984). The calcium and diacylglycerol binding domains are both within the same polypeptide chain as the protein kinase catalytic domain. Since the discovery of the first protein kinase C enzymes, new members of the family with different modes of activation have been identified (Nishizuka, 1992; Dekker and Parker, 1994). In addition to the PKA, PKG and PKC families there are a number of other families of protein kinases belonging to the AGC group which have different modes of regulation. A feature common to several of the protein kinase families in the AGC group is a preference for basic residues surrounding the phosphate acceptor amino acids (Pearson and Kemp, 1991).

The second group (labelled as the CaMK group) includes the protein kinases regulated by interaction with $Ca^{2+}$/calmodulin. Many members of this group, including CaMKI, CaMKII, CaMKIV, PhK-γ (phosphorylase kinase γ) and skMLCK (skeletal muscle myosin light chain kinase) are known to be activated by the binding of $Ca^{2+}$/calmodulin at a domain slightly C-terminal to the catalytic domain. As well as the protein kinases regulated by $Ca^{2+}$/calmodulin, there is a family of protein kinases isolated from plants (calcium-dependent protein kinases, CDPKs) that are regulated by calcium binding to an intrinsic calmodulin-like domain (Harper *et al.*, 1991). The $Ca^{2+}$/calmodulin dependent protein kinases and the CDPKs form a major branch of the CaMK group. The other major branch of the CaMK group is the Snf1/AMPK family of protein kinases. Protein kinases of this family are involved in regulating cellular responses to stress (see review by Hardie, 1994). They are not regulated by calmodulin but have similar substrate recognition motifs to CaMKI (Dale *et al.*, 1995).

The third group of protein kinases (designated the CMGC group) includes the cyclin-dependent protein kinases, the MAP kinases, the glycogen-synthase kinase 3 family and the family of protein kinases related to the human protein kinase Clk ("Cdc-like kinase"). The cyclin-dependent protein kinases are protein kinases whose activity depends on binding regulatory subunits known as cyclins. The protein kinase cdc2 of *S. pombe* and functional homologues in other organisms play a critical role in controlling cell cycle progression. The MAP kinases are protein kinases activated by phosphorylation on tyrosine and threonine and are involved in signalling cascades (see review by Neiman, 1993). The glycogen-synthase kinase 3 (GSK3) family are protein kinases regulated by phosphorylation on tyrosine and serine (Wang *et al.*, 1994). In

mammals the targets for GSK3 include the transcription factor Jun and GSK3 homologues are implicated in control of development in *Drosophila* (Siegfried *et al.*, 1992) and *Dictyostelium* (Harwood *et al.*, 1995). The casein kinase II (CK2) family of protein kinases are most closely related to the GSK-3 family (casein kinase I is not in the same family as casein kinase II) and members of this family are found in highly divergent eukaryotes. They are protein kinases with multiple substrates in both the nucleus and cytoplasm and may be constitutively active (Allende and Allende, 1995). The majority of members of the CMGC group recognise proline-rich phosphorylation site motifs.

The fourth major group of protein kinases (labelled as the PTK group) are the conventional protein tyrosine kinases which are known only in metazoans and exclusively phosphorylate tyrosine residues. The members of this group have distinctive peptide sequence motifs in catalytic subdomains VIb and VIII (see Section 1.2.2). A number of protein tyrosine kinases are transmembrane receptors with intracellular protein kinase domains, such as the EGF, PDGF and insulin receptor protein tyrosine kinase families. Several other families in this group are protein kinases with an intracellular location that lack membrane-spanning domains, an example being the Src family of protein tyrosine kinases. A number of protein tyrosine kinases phosphorylate sites with at least one acidic residue adjacent to the phosphorylated tyrosine (Pearson and Kemp, 1991).

The protein kinase families identified by Hanks and Hunter (1995a+b) that fall outside the four largest groups are not described here for reasons of space, although some are considered in Section 1.3 in relation to protein kinase genes identified in *T. brucei*.

Many of the protein kinases grouped together by catalytic domain sequence comparisons are similar in other ways such as their modes of regulation and substrate specificity, indicating that the catalytic domain sequence comparisons are valid in indicating evolutionary relatedness in a large number of cases. As a whole, the evolutionary origin of the superfamily is uncertain. No gene encoding a eukaryotic-like protein kinase has been discovered in *E. coli* despite the sequencing of a significant proportion of its genome (Hanks and Hunter, 1995a) and a eukaryotic protein kinase homologue from the bacterium *Yersinia pseudotuberculosis* (Galyov *et al.*, 1993) is carried on a plasmid and may well have been transferred from a eukaryote. However, the myxobacterium, *M. xanthus,* which undergoes complex differentiation events has a

number of eukaryotic type protein kinase genes, and gene disruption experiments have shown that one is essential for normal development while another affects spore formation (Zhang *et al.*, 1992; Udo *et al.*, 1995). Nucleotide sequences that predict (very divergent) eukaryotic type protein kinases have also been identified in archaebacteria (Smith and King, 1995). On this basis the protein kinases may have a very ancient evolutionary origin.

### 1.2.2 Catalytic domains

The catalytic domain of protein kinases is highly conserved and the key residues were originally identified on the basis of multiple sequence alignments and by mutational analysis (Hanks *et al.*, 1988 and references therein). Eleven subdomains were defined by Hanks *et al.* (1988) as regions containing conserved residues that are not interrupted by large insertions although, more recently, subdomain VI was redefined as VIa and VIb (Hanks and Quinn, 1991). The publication of crystal structures for a number of protein kinases has shown a conserved bilobar catalytic domain structure in which the smaller N-terminal lobe is mainly responsible for binding the nucleotide substrate and the cleft between the two lobes is the catalytic site. The great majority of protein kinase catalytic domains contain between 250 and 300 amino acids (most are about 260 amino acids), although that of CDC7 contains 437 amino acids. However, the catalytic domain is found in different positions in the primary structure in different protein kinases. Some protein kinase catalytic subunits (e.g. that of cdc2) have very little peptide outside the boundaries of the catalytic domain and are regulated by association with separate polypeptides. Other protein kinases have large N- and/or C-terminal non-catalytic domains with regulatory and targeting function which may include membrane spanning portions.

The first protein kinase catalytic subunit structure to be solved by X-ray crystallography was that of PKA (Knighton *et al.*, 1991a+b; Zheng *et al.*, 1993; Bossemeyer *et al.*, 1993). Crystal structures have also been published for the catalytic subunits of human cyclin-dependent kinase CDK2 (De Bondt *et al.*, 1993), the rat MAP kinase ERK2 (Zhang *et al.*, 1994) and a fission yeast casein kinase I (Xu *et al.*, 1995). Crystal structures that include the catalytic domain have been published for recombinant fragments of the *C. elegans* twitchin kinase (Hu *et al.*, 1994) and the human insulin

receptor protein tyrosine kinase (Hubbard *et al.*, 1994). Protein kinase structure was reviewed by Taylor *et al.* (1992), Taylor and Radzio-Andzelm (1994), Wei *et al.* (1994) and Morgan and De Bondt (1994).

Fig. 1.5 (slightly modified from Taylor and Radzio-Andzelm, 1994) shows an alignment of the sequences of three protein kinase catalytic domains that have been analysed by X-ray crystallography (PKA, ERK2 (MAPK) and CDK2). The twelve catalytic subdomains are indicated by roman numerals with very highly conserved residues marked by asterisks; the peptide sequences that form α-helices and β-strands are shaded and labelled. Fig. 1.6 (from Taylor and Radzio-Andzelm, 1994) shows a ribbon diagram of the structure of PKA with bound ATP and a naturally occurring pseudosubstrate inhibitor (PKI). The β-strands are labelled 1-9, of which strands 1-8 superimpose on β-strands in CDK2 and ERK2 while β-strand 9 is not present in CDK2. The α-helices that superimpose between all three structures are D, E, F, and H, while α-helices C, G, and I are conserved but displaced and α-helix B is not conserved. The ATP and the side-chains of some of the essential catalytic residues are shown in black and the magnesium ion is represented by a black dot. The inhibitor peptide PKI is darkly shaded while a threonine residue that is phosphorylated shortly after protein synthesis (Shoji *et al.*, 1979) is indicated by an arrow.

The function of some of the most highly conserved residues is summarised in Table 1.1. The dominant feature of the minor lobe (uppermost in Fig. 1.6) is a five-stranded β-sheet structure that is largely conserved between all of the protein kinases for which crystal structures have been determined. The glycine rich loop of subdomain I that is involved in anchoring the β-phosphate of ATP (reviewed by Bossemeyer, 1994) is found between β-strands 1 and 2, while the conserved lysine of subdomain II is in the third β-strand, and the conserved glutamate of subdomain III is in a conserved α-helix between β-strands 3 and 4. The residues listed for subdomains I-III are highly conserved among all protein kinases although the third glycine and the valine of subdomain I are somewhat less highly conserved than the other residues. Subdomains IV and V do not contain any invariant or nearly invariant residues but correspond to β-strand 4 (subdomain IV) and β-strand V and α-helix D (subdomain V), the latter two structural elements forming the link between the small and large lobes.

In the larger lobe, subdomain VIa forms a long α-helix that appears to have a mainly structural role while subdomain VIb forms two short β-strands with an intervening loop (the catalytic loop) that contains an invariant aspartate and asparagine. The aspartate is a catalytic base while the asparagine stabilises the loop by hydrogen bonding to the backbone carbonyl of the aspartate. The consensus sequence for the catalytic loop among serine/threonine kinases is H/Y-R-**D**-L-K-X-X-**N** (where X is any amino acid) while that for the protein tyrosine kinases is H-R-**D**-L-R-A-A-**N** (Src family) or H-R-**D**-L-A-A-R-**N** (other protein tyrosine kinases). The lysine in serine/threonine kinases helps to neutralise charge and makes a hydrogen bond to a serine or threonine located seven residues N-terminal of the subdomain VIII glutamate. In protein tyrosine kinases the arginine appears to play a similar role in charge neutralisation but also hydrogen bonds to the catalytic base (Hubbard *et al.*, 1994). The serine/threonine residue seven amino acids N-terminal of the subdomain VIII glutamate is replaced in protein tyrosine kinases by a proline. The aspartate of the "DFG" motif of subdomain VII chelates the $Mg^{2+}$ ion that aids phosphate transfer and this motif is highly conserved among all protein kinases, however, there are differences in secondary structure between CDK2, PKA and ERK2 in this region. In many protein kinases, including PKA, ERK2 and CDK2, there are phosphorylation sites between the conserved motifs of subdomains VII and VIII, the phosphorylation state of which affects enzyme activity. The "APE" motif of subdomain VIII is highly conserved among both serine/threonine and tyrosine kinases, the glutamate being most highly conserved. Ion-pairing between the glutamate and the conserved arginine of subdomain XI helps to stabilise the large lobe. Casein kinase I lacks both the "APE" and the arginine in subdomain XI but the triplet "SIN" at the equivalent position to "APE" appears to adopt the same secondary structure but makes a different set of interactions that stabilise the large lobe (Xu *et al.*, 1995). Subdomain IX, which includes several highly conserved residues and an almost invariant aspartate, forms an α-helix in which the aspartate contacts the highly conserved arginine that lies immediately N-terminal to the catalytic base in subdomain VIb. As can be seen in the sequence alignments, subdomain X includes a small conserved helix but is also a site at which there are significant differences between the three aligned sequences; large insertions at similar positions to those in CDK2 and ERK2 are found in many protein

kinases of the CMGC group. Subdomain XI includes a nearly invariant arginine that lies between two α-helices and which contacts the conserved glutamate in subdomain VIII. Many other features of the function of the protein kinases, for which structures are known, can be at least partly explained in structural terms and are dealt with in the original references and reviews cited. To summarise, the essential catalytic domain core structure is common to all protein kinases but this large structural element can accommodate significant insertions at particular positions.

## 1.3 Protein phosphorylation in *T. brucei*

### 1.3.1 Biochemical studies

The study of protein phosphorylation in *T. brucei* and other kinetoplastids is at an early stage and it is only recently that the first sequences for *T. brucei* protein kinases have been reported (Gale and Parsons, 1993; Hua and Wang, 1994; Mottram and Smith, 1995). The predicted catalytic domains of these protein kinases have a subdomain structure similar to that defined by Hanks *et al.* (1988). *T. brucei* genes encoding protein phosphatases with strong catalytic domain homology to mammalian protein phosphatases of classes 1 and 2A have also been isolated (Evers and Cornelissen, 1990; Erondu and Donelson, 1991). In addition to these protein kinase and protein phosphatase coding sequences, a number of protein kinase activities have also been defined biochemically and show similarities to particular classes of mammalian protein kinases. The available data on protein phosphorylation in *T. brucei* is reviewed below. Biochemical studies aimed at characterising protein kinase activities are described first while studies based on the cloning of protein kinase or protein phosphatase genes are described in Section 1.3.2. Protein phosphorylation in protists has been reviewed by Fazio *et al.* (1994).

Early evidence for the presence of protein serine/threonine kinases in *T. brucei* came from studies of protein kinase activity in extracts of bloodstream *T. b. gambiense* (Walter, 1976; Walter, 1978; Walter, 1980) and *T. b. brucei* (Walter and Opperdoes, 1982). In the first of these studies a protein kinase activity was partially purified from the soluble fraction of trypanosome extracts by sequential ammonium sulphate precipitation, size exclusion chromatography and isoelectric focusing. The molecular weight of the kinase activity was estimated at 95 kDa by size-exclusion chromatography and at 88 kDa by sucrose gradient sedimentation while its isoelectric point was estimated at 4.85.

Protein kinase activity was greatest with basic histone and protamine substrates with little activity detected with acidic casein and phosvitin. Phosphorylation of histones was predominantly on serine but phosphothreonine was also detected. Magnesium and manganese were both found to be acceptable as divalent cation co-factors. The protein kinase activity was unaffected by the cyclic nucleotides cAMP and cGMP but the nucleosides adenosine, guanosine, inosine, uridine and cytidine each stimulated the reaction rate about two-fold at a concentration of 0.1 mM. Two further protein kinase activities from the soluble fraction of bloodstream *T. b. gambiense* extracts were detected in a later study (Walter, 1978). These had apparent molecular weights of 200 kDa and 37 kDa in size-exclusion chromatography. One (PKI) preferred the acidic substrate phosvitin but phosphorylated lysine-rich histones and protamine (basic) as effectively as acidic casein, while the other (PKIII) used only the basic substrates tested. Phosphorylation of phosvitin (PKI) or protamine (PKIII) was predominantly on serine (≥ 95%). Cyclic nucleotides had no effect on PKI and PKIII as had been shown for the previously identified protein kinase (PKII). In a comparison of the effects of the trypanocide suramin on the three protein kinase activities, it was shown that PKII and PKIII were only slightly inhibited at a drug concentration of 0.1 mM but PKI activity could be inhibited by 1 μM suramin (Walter, 1980). Activities apparently similar to PKI and PKII and a protein phosphatase activity were also identified in *T. brucei* (Walter and Opperdoes, 1982).

Keith *et al.* (1990) performed experiments to determine whether enzyme activities similar to those of mammalian protein kinase C could be detected in *T. brucei*. Five protein kinase activities from bloodstream form *T. brucei* were identified in assays of anion-exchange fractions in which histone IIIS was the substrate and $Ca^{2+}$, diacylglycerol (DAG) and phosphatidylserine (PS) were added to each assay. Protein kinase activities PKIII, PKIV and PKV were strongly stimulated by $Ca^{2+}$, DAG and PS together, but not by $Ca^{2+}$ or DAG/PS alone. Phosvitin was demonstrated to be the preferred substrate for PKIV and a $Ca^{2+}$/DAG stimulated protein kinase from the procyclic stage suggesting that these kinases were not protein kinase C like activities as mammalian protein kinase C does not use phosvitin. In addition PKIV was only slightly inhibited by the protein kinase C inhibitor H7 at 1 μM, while PKV was unaffected and

PKIII was inhibited. PKV was found to phosphorylate histone IIIS, phosvitin, and protamine to equal extents. Overall, of the protein kinase activities identified, PKIII was found to be most similar to mammalian protein kinase C in its stimulation by the combination of $Ca^{2+}$, DAG, and phospholipid, its preference for the basic substrates histone and protamine and its sensitivity to H7. All of the other activities had properties distinct from those of mammalian protein kinase C. Protein kinase activities stimulated by $Ca^{2+}$ and DAG were also detected from crude extracts using an IEF gel *in situ* assay and these activities were inhibited by H7. However, it was not determined whether these were the same as any of the activities characterised in the assays of anion-exchange fractions. An antibody raised against an epitope from a mammalian PKC was also shown to recognise a 95 kDa protein from *T. brucei* bloodstream forms by Western blotting.

An alternative approach to studying protein phosphorylation in *T. brucei* was adopted by Aboagye-Kwarteng *et al.* (1991) who compared SDS-PAGE phosphoprotein profiles from the high speed supernatants of total extracts of slender and stumpy bloodstream *T. brucei* labelled with [$\gamma$-$^{32}$P]ATP *in vitro*, in the absence of exogenous substrate. Trypanosomes were sampled over a time course of infection at the start of which all cells were slender form while at the end 95% of cells were short stumpy form. Among the phosphoproteins detected it was observed that an 80 kDa phosphoprotein (pp80) became progressively less prominent over this time, while two phosphoproteins, pp42 and pp37, became more prominent. The same phosphoproteins were also found to be more prominent in *T. brucei* that was induced to become morphologically stumpy by administration of the trypanocide DFMO to the rodent hosts. In order to characterise the protein kinases phosphorylating pp42 and pp37, labelling reactions were performed *in vitro* under different conditions. It was found that both ATP and GTP were acceptable phosphoryl donors for the labelling of pp42 and pp37 but the phosphorylation of these proteins was differentially affected by the concentration of KCl in the labelling buffer. A number of activators of some of the main families of protein kinases in higher eukaryotes (cAMP, cGMP, $Ca^{2+}$/calmodulin, $Ca^{2+}$/phospholipid) showed no effect on activity but the negatively charged compounds heparin and suramin inhibited phosphorylation of all the phosphoproteins detected. Positively charged polylysine and polyarginine inhibited phosphorylation of pp42 and pp37 while having little effect on the phosphorylation of other proteins. Since the properties of the protein kinase activities that phosphorylate

pp42 and pp37 were in some ways similar to those of mammalian casein kinase II (ability to utilise GTP and inhibition by heparin) extracts of the slender and stumpy cells were tested for their ability to phosphorylate exogenous casein and phosvitin. In spite of the increased phosphorylation of pp42 and pp37 in extracts of stumpy bloodstream forms, the extracts of slender bloodstream forms were found to be more active in phosphorylating these substrates. However, the results of experiments in which heat-denatured extracts were used as substrates suggested that pp42 was more abundant in stumpy cells while pp80 was less abundant, and although phosphorylation of pp37 from heat-treated extracts could not be shown, the results of a Western blot analysis with an antiserum against pp37 indicated that pp37, like pp42, is also present at a higher level in the stumpy cells.

Parsons *et al.* (1993) used a protein kinase renaturation assay (Ferrell and Martin, 1989) to examine protein kinase activities from different *T. brucei* life cycle stages. Multiple protein serine/threonine kinase activities were detected from slender, stumpy and procyclic forms with apparent molecular weights ranging from 43 kDa to 181 kDa. A protein kinase with an apparent molecular weight of 71 kDa (PK71) showed highest activity in slender bloodstream forms, protein kinases PK181, PK123 and PK43 showed highest activity in stumpy forms while PK52 showed highest activity in procyclic forms. Only one of the eight activities analysed showed no stage specific differences in the level of activity (PK97). Stumpy bloodstream forms are cell cycle arrested while the slender bloodstream and procyclic forms are proliferative. Therefore, differences in protein kinase activities between the stumpy forms and the slender and procyclic forms might be due to protein kinases involved in cell cycle control. In order to address this question the comparison of the protein kinase activities was extended to include samples from stationary phase procyclic cells. None of the activities that were upregulated in stumpy forms were also upregulated in stationary phase procyclics, but a number of differences were found between the two procyclic samples. In particular PK89, which had much lower activity in stumpy forms than in slender and proliferating procyclic, was found to have very low activity in stationary phase procyclic forms (see below).

These experiments demonstrated the existence of a number of renaturable protein kinases in *T. brucei* with the activities of most being stage specific. While only trace amounts of phosphotyrosine were detected in the kinase renaturation assay (on PK97),

tyrosine phosphorylation was demonstrated in *T. brucei* (and in *Crithidia fasciculata*, *Trichomonas vaginalis*, and *Giardia lamblia*) by Western blotting with a previously characterised anti-phosphotyrosine serum (Parsons *et al.*, 1991). A reason suggested for the lack of protein tyrosine kinase activity in the kinase renaturation assay is that these protein kinases may not renature well. The kinase renaturation assay was also used to investigate protein kinase activities from different life cycle stages of *T. congolense* grown in culture or in rodents (Parsons *et al.*, 1995). As observed previously for *T. brucei*, there were significant differences in the protein kinase activities detected between the life cycle stages. Several of the activities were of similar size and showed similar stage regulation to protein kinases detected in the earlier study of protein kinase activities in *T. brucei*, but there were also a number of activities detected in one species and not in the other. Phosphoamino acid analysis showed that the proteins labelled in the kinase denaturation assay were only labelled on serine and threonine, although multiple tyrosine-phosphorylated proteins were demonstrated in extracts of all stages by Western blotting. Protein kinase activities were also assayed over a time course from cells stimulated to differentiate from bloodstream to procyclic forms by the addition of citrate and cis-aconitate to their culture medium and a temperature shift to 24 °C. The profiles of protein kinase activities were different within 24 hours between the stimulated and unstimulated cells and further differences were found at later time points.

The 89 kDa protein kinase activity from *T. brucei*, identified in the study of Parsons *et al.* (1993) as being potentially specific to proliferating forms, was investigated further in studies with cultured procyclic cells (Gale *et al.*, 1994a). It was first shown that stationary phase procyclics stimulated to proliferate by fresh culture medium underwent one cell cycle in a semi-synchronous manner. Induction of the 89 kDa protein kinase (named SPK89 in this study) in the first cell cycle occurred at a time corresponding to S phase (5 hours); levels remained high at the time when the proportion of cells in $G_2$/M was high (8 hours), but was lower at the 12 hour time point by which time the parasites had undergone one division cycle. There was rapid loss of synchrony after this time, but in accordance with the results of the earlier study, SPK89 activity in stationary phase procyclics was very low. To test whether the pattern of expression of SPK89 activity observed was specific to previously quiescent cells re-entering the first dividing cell cycle, or whether the same pattern would be repeated in subsequent cycles,

SPK89 activity was measured in cells from an asychronous culture sorted into populations at different cell cycle stages by flow cytometry. SPK89 activity was highest in the S phase population, lower in the $G_2$/M population and lowest in the G1 population, supporting the idea that SPK89 activity is S-phase specific. The effects of protein kinase inhibitors and cycloheximide on induction of SPK89 in stationary phase procyclic cultures placed in fresh medium were also investigated in this study. The protein kinase inhibitors genistein and staurosporine added at the same time as fresh medium blocked induction of both S phase and SPK89 activity, but had no direct effect on the protein kinase activity when added to the kinase renaturation assay. Cycloheximide also blocked induction of S phase and the appearance of SPK89 activity if added within 3 hours of the addition of fresh medium, but not at subsequent times. The gene encoding SPK89 has not yet been identified but some similarities of the protein's properties and pattern of expression to those of the *S. cerevisiae* protein kinase Spk1 (Zheng, P. *et al.*, 1993) were noted.

Wheeler-Alm and Shapiro (1992) investigated protein tyrosine kinase activity in *T. brucei*. An activity that could phosphorylate a synthetic glutamate/tyrosine polymer was detected in extracts of bloodstream forms and could be heat-inactivated. A number of endogenous proteins also became phosphorylated when these extracts were incubated with [$\gamma$-$^{32}$P]ATP in a buffer containing $Mn^{2+}$ ions without exogenous substrate. A comparison of the autoradiograph of an SDS-PAGE gel lane containing the labelled extract that had been treated with alkali, with that of an equivalent untreated lane showed that a number of bands were more prominent after alkali treatment. As phosphotyrosine is more resistant to alkaline hydrolysis than phosphoserine and phosphothreonine this result suggests that the more prominent bands (resistant to alkali treatment) were labelled on tyrosine. The compound genistein which inhibits a number of mammalian protein tyrosine kinases was found to inhibit both phosphorylation of the glutamate/tyrosine polymer and to inhibit multiplication of *T. brucei* grown in culture. However, a later study (Gale *et al.*, 1994a) showed marked effects of cell treatment with genistein on protein synthesis and on protein serine/threonine phosphorylation, and while these could have been downstream effects of inhibition of protein tyrosine kinases it is not clear that the effects of genistein on the proliferation of *T. brucei* actually result from inhibition of protein tyrosine kinase activity.

As well as the studies described above in which protein kinase activity was assayed *in vitro* in whole cell extracts or after partial protein kinase purification and addition of exogenous substrates, a number of studies have been performed to detect *T. brucei* proteins phosphorylated in living cells. In some of these studies cultured procyclic forms or bloodstream forms isolated from laboratory rodents were metabolically labelled in culture, while in other experiments to examine tyrosine phosphorylation phosphotyrosine antisera were used to detect tyrosine-phosphorylated proteins without the requirement for prior $^{32}P$ labelling. In a further experiment described below, the phosphorylation state of *T. brucei* RNA polymerase II was examined after photoaffinity labelling of the active form of the enzyme in isolated nuclei.

Clayton and Fox (1989) examined the phosphorylation of the glycosomal enzyme fructose bisphosphate aldolase in *T. brucei* bloodstream cells isolated from rats and then maintained in culture for metabolic labelling and pulse-chase experiments. Phosphate labelling of immunoprecipitated aldolase could usually be detected from cells lysed after a 30 minute labelling period, and was reliably detected after a 90 minute labelling period. The kinetics of phosphate labelling of VSG (not due to protein kinase activity) which was to be used as a control proved to be poorly reproducible. Pulse-chase experiments in which parasites were labelled with both [$^{3}H$]leucine and [$^{32}P$]orthophosphate were performed to assess the stability of the phosphate labelling of aldolase. Tritium labelling of both VSG and aldolase did not change significantly after chase periods of up to 19 hours but, while the ratio of tritium to phosphate label incorporated into aldolase did not vary during this time, the amount of phosphate label incorporated into VSG increased dramatically. Two hypotheses were proposed to explain the results, one being that phosphate labelling of aldolase is stable and that the cytoplasmic or glycosomal phosphate pool equilibrates rapidly but the endoplasmic reticulum and golgi phosphate pool (responsible for phosphate labelling of VSG) equilibrates much more slowly. Alternatively, the aldolase phosphate groups might be continually turned over with the constant level of aldolase phosphorylation being explained by very slow removal of phosphate from the cells during the chase period. The phosphorylation of aldolase was found to be solely on serine but the phosphorylation site was not determined. No studies were performed to determine whether the phosphorylation of aldolase in *T. brucei* has functional significance.

Parsons *et al.* (1991) investigated tyrosine phosphorylation in different *T. brucei* life cycle stages using a polyclonal anti-phosphotyrosine serum. Proteins were identified in all life cycle stages that reacted with the serum, and specificity for phosphotyrosine was indicated by the fact that this could be blocked by phenyl phosphate. Phosphoamino acid analysis of immunoprecipitates from procyclic cells metabolically labelled with [$^{32}$P]orthophosphate showed that immunoprecipitated proteins contained phosphoserine and phosphothreonine as well as phosphotyrosine. In an examination of stage specific tyrosine phosphorylation by Western blot analysis of total cell extracts, it was found that the profiles of reacting proteins were fairly similar between life cycle stages, except that a set of strongly-labelled bands of about 40 kDa and 42 kDa in procyclic, intermediate and stumpy forms were much more weakly labelled in the slender bloodstream forms. Similar results were obtained for the slender and procyclic forms of a monomorphic stock. However, induction of stumpy forms by the trypanocide DFMO did not induce a stronger signal from the 40/42 kDa proteins compared to that from slender forms. Monoclonal antibodies were raised against the 40/42 kDa tyrosine phosphorylated proteins (Parsons *et al.*, 1994) and used for further investigations. Two monoclonal antibodies were each found to recognise both of the proteins at 40/42 kDa (although the size estimates were changed to 44/46 kDa) suggesting that 44/46 kDa phosphoproteins (pp44/46) are related. The possibility that the common epitope was phosphotyrosine was ruled out by the fact that recognition of pp44/46 in Western blots by the antibodies was not blocked by phosphotyrosine. The two antibodies were used in addition to the anti-phosphotyrosine serum to determine the levels of pp44/46 in different stages and their tyrosine phosphorylation levels. For a naturally differentiating stock, the abundance of pp44 was low in slender bloodstream forms while almost no signal was observed for pp46. The protein level of pp44 was 3-4 fold higher in stumpy and procyclic forms than in slender forms, while the stoichiometry of phosphate labelling was 9-fold higher in procyclic forms compared to slender forms and higher still in stumpy forms. The 46 kDa protein was partially induced in stumpy forms but was most abundant in procyclic forms and there was no detectable tyrosine phosphorylation of this protein in slender forms. The two immunoprecipitated proteins were found to have associated protein serine and tyrosine kinase activity, but no stoichiometric association with other proteins was observed. Reactivity with antisera to MAP kinases could not be demonstrated. To

summarise, these two studies provided firm evidence that protein tyrosine phosphorylation does occur in *T. brucei* and that, as in other eukaryotes, at least some tyrosine phosphorylated proteins are also phosphorylated on serine and threonine. Additionally, clear stage-specific differences in the abundance and extent of tyrosine phosphorylation of two related proteins was shown, and less dramatic stage-specific differences in reactivity of other proteins with an anti-phosphotyrosine serum were also found.

In order to investigate microtubule-associated proteins *in T. brucei*, Woods *et al.* (1989) raised monoclonal antibodies against the whole trypanosome cytoskeleton. Characterisation of a protein recognised by one of these antibodies showed it to be distributed at low abundance (relative to tubulin) over the entire surface of the cytoskeleton, but not the flagellum (Woods *et al.*, 1992). Immunogold microscopy showed the protein localised mainly on the exterior face of the subpellicular cytoskeleton and to the side of the microtubules. However, experiments performed to test for microtubule binding gave inconclusive results. The apparent molecular weight of the protein in SDS-PAGE was 210 kDa, while sucrose gradient density sedimentation suggested that it was globular, unlike many other cytoskeletal proteins. Since a low-abundance globular microtubule-associated protein might play a regulatory role in cytoskeletal function, an analysis was performed to test whether the protein was phosphorylated in cultured procyclic forms metabolically labelled with $^{32}P$ orthophosphate. A number of proteins were labelled from different cytoskeletal fractions, and in 2-D gel analysis it was shown that one of these was the 210 kDa antigen. Therefore, while the role of the 210 kDa antigen has not been determined, there is at least preliminary evidence that in *T. brucei*, as in many other eukaryotes, the function of cytoskeletal proteins may be regulated by phosphorylation.

As discussed in Section 1.1.2, transcription and other aspects of gene expression in *T. brucei* have been found to have a number of unusual features. One unusual feature relates to the large subunit of *T. brucei* RNA polymerase II. In organisms ranging from *S. cerevisiae* to mammals the C-terminus of the large subunit of RNA polymerase II is made up of tandemly repeated heptapeptide sequences (e.g. 25 repeat units in *S. cerevisiae*, 52 in mouse) that are phosphorylated as the polymerase activity changes from initiation to elongation (Cadena and Dahmus, 1987). In *T. brucei* there are two genes

encoding the large subunit of RNA polymerase II and both of the predicted proteins lack C-terminal heptapeptide repeats (Evers *et al.*, 1989; Smith *et al.*, 1989). In order to investigate whether *T. brucei* RNA polymerase II large subunits were nevertheless phosphorylated, Chapman and Agabian (1994) used antisera to identify RNA polymerase II large subunits that had been labelled by an affinity technique specific to elongating complexes. It was found that treatment of the immunoprecipitated RNA polymerase II with calf intestinal phosphatase led to its faster migration in SDS-PAGE (consistent with dephosphorylation), and this did not occur in a control experiment in which a phosphatase inhibitor was included in the incubation buffer. Experiments in which the polymerase large subunits were partially protease digested before or after phosphatase treatment and then detected with a C-terminus specific antiserum indicated that despite the lack of heptapeptide repeats the C-termini were extensively phosphorylated. In addition, when the phosphatase treatment was applied to total RNA polymerase II large subunits rather than those specifically engaged in transcriptional elongation, the C-termini of higher mobility forms were unaffected by phosphatase treatment. Thus the results of Chapman and Agabian (1994) indicate that despite the lack of heptapeptide repeats, phosphorylation of the C-termini of RNA polymerase II large subunits is associated with the change from a transcription initiation complex to an elongation complex in *T. brucei*, as in higher eukaryotes.

### 1.3.2 Protein kinase and protein phosphatase genes

Protein kinases and phosphoproteins can often be studied biochemically even when they are present at very low abundance because of the high sensitivity of assay techniques based on the use of $^{32}P$. However, obtaining the sequence of the gene encoding a specific protein kinase through techniques that depend on prior protein purification is often extremely difficult. An alternative approach to the study of protein kinases is to begin with gene cloning using one of a number of techniques based on the sequence homology between different protein kinases (see Hanks and Lawton, 1993), and a number of workers have used the polymerase chain reaction (PCR) with primers based on conserved protein kinase catalytic domain sequence motifs in the isolation of protein kinase genes from *T. brucei*. The various investigations into protein phosphorylation in *T. brucei* that have begun with gene cloning are described below.

### 1.3.2.1 Nrk (Nek1 and NimA related kinase)

Gale and Parsons (1993) used PCR with degenerate primers based on protein kinase catalytic subdomains VIII and IX to amplify *T. brucei* gene fragments similar to protein kinase genes. A gene-specific primer, based on the sequence of one of the clones obtained that was similar to protein kinase genes, was then used with a primer based on the *T. brucei* spliced leader to PCR amplify a longer fragment from cDNA. This fragment was cloned and used to screen genomic and cDNA libraries. Sequence analysis of one genomic clone, pNrkA, obtained by this approach showed that it contained an open reading frame encoding a putative protein kinase in which all of the conserved protein kinase catalytic subdomains could be recognised. Sequencing of the single cDNA clone obtained (pNrkB) showed that it contained a region of 97% nucleotide sequence identity to pNrkA across the coding region and into the 3′ untranslated region, and also included a poly(A) tail. However, the open reading frame in this clone was terminated by a stop codon before catalytic subdomain VIII, and any protein product would therefore not be expected to be active. This clone also contained sequence 5′ of the spliced leader acceptor site and was therefore assumed to be a precursor mRNA. The 5′ sequence of this clone was very divergent from that of pNrkA suggesting that there might be more than one genetic locus for the putative protein kinase Nrk. It was shown by Southern blotting that *T. brucei* has two nrk loci (named *nrkA* and *nrkB*) and it was also shown that in two different *T. brucei* stocks transcripts could be detected from both loci, but that stock IsTAR 1 was homozygous for the truncated pseudogene at *nrkB* while stock TREU667 had one pseudogene allele, and one allele encoding a potentially functional protein at this locus. No significant difference was found between the level of expression of each mRNA between slender, stumpy and procyclic forms. Database searches showed that the Nrk catalytic domain sequence was most closely related to that of the murine dual-specific protein kinase Nek1 and of the *Aspergillus* protein serine/threonine kinase NimA. (Letwin *et al.*, 1992; Pu and Osmani, 1995; Ye *et al.*, 1995). However, Nrk is not similar to either of these protein kinases outside the catalytic domain, and the C-terminus of Nrk has similarity to pleckstrin homology domains (Gale and Parsons, 1993; reviewed by Musacchio *et al.*, 1993) which are not found in either Nek1 or NimA.

In a later study (Gale *et al.*, 1994b) the expression of the Nrk protein was analysed using an antiserum raised against a NrkA fusion protein which was shown to

specifically immunoprecipitate Nrk from *T. brucei* extracts. The ability of Nrk immunoprecipitated from stumpy bloodstream cells to phosphorylate various substrates was tested. Nrk was found to be capable of phosphorylating both histone H-1 (basic) and β-casein (acidic) to some extent, but showed higher activity with myelin basic protein. Enolase, poly(glu-tyr), immunoglobulin heavy chain and heat-denatured *T. brucei* extracts were poor substrates and no autophosphorylation of Nrk was observed. Phosphoamino acid analysis of myelin basic protein and β-casein phosphorylated by Nrk showed that labelling was on serine and serine and threonine respectively. Nrk was able to use both $Mg^{2+}$ and $Mn^{2+}$ and its activity against myelin basic protein was significantly enhanced by the addition of 100 mM NaCl to the reaction buffer.

An analysis of the expression of Nrk and its activity during the procyclic cell cycle was performed using stationary phase cultured procyclic cells stimulated to re-enter a dividing cell cycle by dilution into fresh medium. Levels of both Nrk protein and its activity were similar throughout the first cell cycle after stimulation, during which time there was a reasonable degree of synchrony. The level of Nrk protein during the cell cycle was also examined in cells from exponential cultures sorted by flow cytometry and again there was no evidence for cell cycle regulation of Nrk abundance. The levels of Nrk protein were also examined in different *T. brucei* life cycle stages. While in the earlier study it was shown that the expression of Nrk mRNAs was constitutive during the life cycle, it was found by immunoblot analysis that the level of the protein was highly regulated during the life cycle with lowest expression of Nrk in slender bloodstream forms, somewhat higher expression in procyclics, but much higher expression in stumpy forms. It was shown by analysing bloodstream cells from a naturally differentiating stock harvested at different time points during differentiation from slender to stumpy form that Nrk protein expression increased early in the differentiation process. The expression of Nrk was also examined in the monomorphic IsTAR 1 stock which can be induced to differentiate to morphologically stumpy forms by DFMO treatment, and which had been previously shown to have truncated pseudogenes at both *nrkB* alleles. Nrk protein was expressed at a low level in the slender bloodstream forms of this stock, but was strongly induced by DFMO treatment, indicating that the upregulation of the level of Nrk protein is at least partly due to the *nrkA* locus. The level of Nrk protein kinase activity from the different life cycle stages was found to parallel that of the protein's abundance. The

results of biosynthetic labelling experiments indicated that developmental regulation of Nrk abundance is primarily due to changes in the rate of its biosynthesis rather than its degradation. It was also shown that while stumpy forms showed the highest rate of Nrk synthesis, during the labelling period these forms synthesised 11 and 7 times less total protein than the slender and procyclic forms respectively. The Nrk proteins synthesised by slender, stumpy and procyclic cells had identical mobilities in SDS-PAGE when analysed after a 1 hour labelling period. However, when samples were analysed following a 1-4 hour chase period the labelled Nrk protein from procyclic forms had the same mobility in SDS-PAGE as when analysed at the earlier time point, but Nrk from stumpy forms had a slightly lower mobility. Therefore, it seems that Nrk is subject to a post-translational modification (perhaps phosphorylation) in stumpy forms but not in procyclic forms. The lack of developmental regulation of *Nrk* transcript abundance found in the earlier study was confirmed, and it was also determined that in slender, stumpy and procyclic forms all of the *Nrk* transcripts were polyadenylated and in each stage the same spliced leader acceptor site was used. The conclusion reached was that developmental regulation of Nrk expression is at the translational level, a level of control which is unusual in *T. brucei*.

#### 1.3.2.2 cdc2-related protein kinases

Three genes encoding protein kinases with homology to the cdc2 family of protein kinases have been isolated from *T. brucei* (*tbcrk 1-3*) and there is evidence for further cdc2-related protein kinases in this organism (Mottram and Smith, 1995; Mottram, 1994). The genomic organisation of the three genes was analysed by Southern blotting of genomic digests and *T. brucei* chromosomes separated by pulse-field gel electrophoresis. The results of this analysis indicated that each gene is single copy and that the genes are not closely linked although *tbcrk1* and *tbcrk2* could be on the same chromosome. All of the *tbcrk* genes are located on large chromosomes (>3 Mb). A polyclonal antiserum against a peptide corresponding to the conserved PSTAIR box of cdc2 kinases reacted with a major protein band of 32 kDa and minor bands of 43, 65, and 90 kDa in blots of extracts from slender, stumpy and procyclic *T. brucei*, and the recognition of each protein was blocked by a PSTAIR peptide. This result is consistent with there being a family of cdc2-related protein kinases in *T. brucei*. The 32 kDa polypeptide could be

TbCRK1 since the antiserum used in the blotting cross-reacted strongly with a TbCRK1 fusion protein, and TbCRK1 has a calculated molecular weight of 34 kDa.

The proteins have been named cdc2-related protein kinases (CRKs) since they have not yet been shown to associate with cyclin partners. Each of the TbCRKs has the PSTAIR box domain which is implicated in cyclin binding (Ducommun *et al.*, 1991), and sites of regulatory phosphorylation in cdc2 are also conserved (see review by Dunphy, 1994), although in TbCRK1 and TbCRK2 the residue in the equivalent position to T14 in cdc2 is substituted by a serine. The conservation of these features suggests that TbCRKs 1-3 may well be cyclin-dependent kinases also regulated by phosphorylation, and a question of obvious interest is whether any of the TbCRKs might be a functional homologue of cdc2 acting as the key regulator of entry into mitosis. TbCRK1, TbCRK2 and TbCRK3 are respectively 54%, 49% and 53% identical in amino acid sequence to HsCDC2. TbCRK1 is most closely related (88% and 72% identity respectively) to cdc-related kinases from *T. congolense* and *Leishmania mexicana* (Mottram, 1994; Mottram *et al.*, 1993) while TbCRK2 and TbCRK3 are both most closely related to human CDK2 (52% and 59% identity respectively). The three *T. brucei* protein kinases are less closely related to each other than to HsCDC2. A phylogenetic analysis of the TbCRK sequences with those of other cdc2-related protein kinases showed that TbCRK1 and the leishmanial cdc-related protein kinase (LmmCRK1) form an independent outgroup while TbCRK2 and TbCRK3 group with human cyclin-dependent protein kinases CDK4 and CDK5. Thus, by this criterion, none of the TbCRKs seems likely to be a functional cdc2 homologue. LmmCRK1 which groups with TbCRK1 has also been shown to be a poor candidate for a cdc2 functional homologue by various criteria including its failure to complement an *S. pombe* cdc2 temperature sensitive mutant, and its lack of activity towards histone H1 in amastigotes (which are dividing) when this activity could be readily detected in both proliferative and stationary promastigotes (Mottram *et al.*, 1993). It is possible that a primitive eukaryote such as *T. brucei* does not have a functional homologue of cdc2. However, there is evidence for further genes encoding cdc2-related protein kinases in *T. brucei* (see Mottram, 1994; Mottram and Smith, 1995) and a histone H1 protein kinase activity distinct from LmmCRK1 has been detected in *Leishmania mexicana* (Mottram *et al.*, 1993). This kinase binds to p13suc1 beads as well as having greater activity in proliferating *Leishmania* forms than in non-proliferating

forms, and is therefore a candidate leishmanial cdc2 homologue. Affranchino *et al.* (1993) have isolated a *T. brucei* cDNA clone encoding a protein (CYC1) with approximately 25% amino acid identity to human cyclins of A and B classes. Expression of CYC1 rescued an *S. pombe* cyclin mutant and in procyclic *T. brucei* the protein's abundance was found to be cell cycle regulated in a manner consistent with it acting as a mitotic cyclin. An antiserum against CYC1 co-immunoprecipitated a 34 kDa protein that reacted with an anti-PSTAIR serum and a histone H1 kinase activity could also be precipitated by this serum. To summarise, *T. brucei* has a family of cdc2-related protein kinases and although it is not clear whether any of the TbCRKs isolated so far plays a similar role to that of cdc2 in regulating entry into mitosis there is, however, indirect biochemical and immunochemical evidence suggesting that *T. brucei* may have such a protein kinase.

#### 1.3.2.3 A MAP kinase homologue

Hua and Wang (1994) isolated seven amplified cDNA fragments similar to protein kinase coding sequences from *T. brucei* using primers based on conserved catalytic domain peptide sequences. One of these fragments (TbPK-A1) was used to screen a *T. brucei* cDNA library and hybridising clones were isolated. The sequence of the clone with the largest insert contained an open reading frame encoding a putative protein kinase of 350 amino acids, with the catalytic domain extended at the N-terminus by 13 amino acids and at the C-terminus by 51 amino acids. Database searches with the predicted peptide sequence showed that the protein kinase was most closely related to the MAP kinases (MAPKs) KSS1 and FUS3 from *S. cerevisiae* (42% and 40% identity respectively excluding the C-terminal 40 amino acids) and was also similar to the rat MAP kinase, ERK1 (37% identity). No significant similarity was found between the C-terminal 40 amino acids and any sequences in the databases. In addition to the significant overall homology to MAPKs, the *T. brucei* protein kinase has the sequence "TEY" in subdomain VIII. This sequence (sometimes with a different central residue) is characteristic of MAPKs and in other eukaryotes is known to be the site of activating phosphorylations on the threonine and tyrosine catalysed by dual-specificity MAPK kinases (Davis, 1994). Thus on the basis of its sequence the *T. brucei* protein kinase (named KFR1) is a putative MAPK.

*KFR1* was shown to be a single copy gene by Southern blotting while Northern blot analysis showed a single hybridising RNA species from both bloodstream and procyclic *T. brucei* of about 1.4 kb that was 3-4 fold more abundant in the bloodstream stage. The level of expression of the protein in these life cycle stages was also examined by Western blotting with an antiserum raised against the C-terminal sequence and it was found that the protein was about twice as abundant in the bloodstream forms compared to the procyclic forms.

### 1.3.2.4 Partial cDNAs for *T. brucei* protein kinases

The predicted peptide sequences of the remaining six amplified fragments (TbPK-A2-A5; TbPK-B1 and TbPK-B2) isolated by Hua and Wang (1994) were used in database searches for similar sequences. TbPK-A2 was found to be most similar to a protein kinase from *Drosophila* named polo and to a mammalian PKC, which overall are not closely related to each other. A cDNA that includes TbPK-A2 was cloned as part of the present study (Chapter 5) and it is clear from the more complete sequence that the putative protein kinase corresponding to TbPK-A2 is in fact a polo-like kinase.

The assignment of the other sequences to particular protein kinase subgroups is less certain because the cDNA fragments are short, and because the region of the catalytic domain they cover is significantly conserved between even divergent protein kinases. The closest matches found in the original database searches with the predicted peptide sequence for TbPK-A3 were to SNF1 of *S. cerevisiae* (49% identity); a protein kinase of the same family from *Arabidopsis* and to a *Dictyostelium* myosin light chain kinase. TbPK-A4 and TbPK-A5 both showed most similarity to YAK1 of *S. cerevisiae* (53% and 48% identity respectively). TbPK-B1 was most similar to a *Drosophila* member of the protein kinase C family (47% identity) and TbPK-B2 was most similar to COT-1 of *Neurospora* (34% identity). In a Northern blot analysis of *T. brucei* poly(A)$^+$ mRNA probed with the various PCR clones it was shown that transcripts for TbPK-A2, TbPK-A5, TbPK-B2, TbPK-A4 and TbPK-B1 are present at similar levels in bloodstream and procyclic forms (though the signals from the latter two are weak in both life cycle stages), while TbPK-A1 mRNA is 3-4 fold more abundant in the bloodstream forms. In the case of TbPK-A3 a 7.4 kb transcript was common to bloodstream and procyclic forms while two transcripts of 2.2 and 2.5 kb were only found in the

bloodstream forms.

### 1.3.2.5 Protein phosphatase genes from *T. brucei*

*T. brucei* has two almost identical genes encoding the RNA polymerase II large subunit, which are located on different chromosomes (Evers *et al.*, 1989). Evers and Cornelissen (1990) found, in a study of the 5′ flanking regions in the two loci, that in both cases the genes immediately 5′ of the large subunit genes encode putative protein phosphatases. The predicted peptide sequences from the two loci are 96% identical to each other. The protein phosphatase peptide sequence from one locus was found to be 56% identical to rat protein phosphatase 1 (α isoform) and 40% identical to rat protein phosphatase 2A, and removal of 50 N- and C-terminal amino acids from the comparison increased the homology with PP-1α to 64%. It was therefore concluded that the putative *T. brucei* enzymes were of the protein phosphatase 1 family although neither sequence was classified as being of a particular isoform. The sequences immediately 5′ and 3′ of the protein phosphatase coding regions from the two loci were found to be quite different from each other, but it was shown that the overall homology between the two loci extended several kb further 5′ of the protein phosphatase genes. Using a different approach Erondu and Donelson (1991) cloned two different cDNAs from *T. b. rhodesiense* encoding putative protein phosphatase 1 homologues which had 3′ untranslated regions almost identical to the previously described flanking sequences 5′ to the two RNA polymerase II loci, suggesting that these cDNAs are transcribed from the previously described protein phosphatase genes (Evers and Cornelissen, 1990). The results of an S1 nuclease mapping analysis by Evers and Cornelissen (1990) showed that both protein phosphatase 1 genes were co-transcribed with the RNA polymerase II large subunit genes and Erondu and Donelson (1991) were able to detect precursor mRNAs that hybridised to probes from both protein phosphatase and RNA polymerase coding sequences. As described above, the C-terminal domain of the RNA polymerase large subunit is known to be extensively phosphorylated in actively transcribing complexes from *T. brucei* (Chapman and Agabian, 1994) and an endogenous activity dephosphorylating the C-terminal domain was also detected in this study. However, this activity was not characterised, and it is not known whether the *T. brucei* protein phosphatase 1 can dephosphorylate the polymerase large subunit.

Erondu and Donelson (1991) also cloned a further cDNA from *T. b. rhodesiense* with an open reading frame predicting a protein phosphatase 55% identical to mammalian protein phosphatase 2A and with a lower degree of identity (39%) to mammalian protein phosphatase 1. Southern blotting showed that the gene was single copy. One of the protein phosphatase 1 homologues and the protein phosphatase 2 homologue were expressed as fusion proteins. An antiserum against the protein phosphatase 1 fusion protein cross-reacted with mammalian protein phosphatases of the 1, 2A and 2B classes, and with the protein phosphatase 2A fusion protein, while the antiserum against the *T. brucei* protein phosphatase 2A only cross-reacted with the mammalian enzyme of the same class. The native enzyme was shown to have an apparent molecular weight of 35 kDa agreeing with its predicted molecular weight of 34.5 kDa. However, the fusion proteins were not active and the biochemical properties of the enzymes have not been characterised.

### 1.3.3 Conclusions

The biochemical studies of protein kinase activities in *T. brucei* have shown that many protein kinase activities are developmentally regulated and that, as in higher eukaryotes, phosphorylation on serine and threonine residues is more common than phosphorylation on tyrosine. However, in most cases the initial identification of a protein kinase activity or phosphoprotein has not been developed further. The application of homology based approaches (mainly PCR) to the cloning of protein kinase and phosphatase genes from *T. brucei* and other kinetoplastids has recently yielded results (the first *T. brucei* protein kinase sequence was published in 1993), details of which are summarised in Table 1.2. The identification of cdc2-related protein kinases in *T. brucei*, *Leishmania mexicana* and *Crithidia fasciculata*, a putative mitotic cyclin in *T. brucei*, and a *T. brucei* polo-like kinase (see Chapter 5) strongly suggests that kinetoplastids share with other eukaryotes much of the molecular machinery involved in cell cycle regulation. Similarly the isolation of genes for MAPK and MAPKK homologues from *T. brucei* and *Leishmania* respectively suggest that kinetoplastids have MAP kinase cascades. In higher eukaryotes such cascades transduce signals from the cell surface to the nucleus. The *T. brucei* PKAs also show strong homology to PKAs from higher eukaryotes and adenylate cyclase homologues have been identified in both *T. brucei* and *Leishmania* (Alexandre *et al.*,

1990; Sanchez *et al.*, 1995), while a cAMP phosphodiesterase homologue has also been identified from *T. brucei* (M. Boshart, personal communication). It seems reasonable to conclude that *T. brucei* and other kinetoplastids have many of the protein phosphorylation pathways that have been identified in higher eukaryotes and that many more protein kinase genes will be identified in the future, either through further studies using homology based approaches to gene cloning or through large scale genome sequencing. Although the protein kinases of *T. brucei* now appear to be quite conventional considering the early evolutionary divergence of this organism there are likely to be some differences between the roles of protein phosphorylation in *T. brucei* and in higher eukaryotes. For instance, in higher eukaryotes, including unicellular eukaryotes such as yeasts, the downstream targets of many protein phosphorylation cascades are transcriptional regulators. In *T. brucei* most genes are organised in polycistronic transcription units within which differential gene expression (which is common) cannot result from regulation of the rate of transcriptional initiation. It therefore seems likely that if protein phosphorylation cascades are involved in controlling gene expression in *T. brucei* then the downstream targets will be DNA and RNA binding proteins responsible for RNA maturation and stability.

Given the high level of interest in understanding the differentiation events that *T. brucei* undergoes during its complex life cycle (and in the development of ways of interfering with them) research into protein phosphorylation in this organism is likely to expand at an increasing rate. In other organisms the cloning of protein kinase genes has not necessarily rapidly led to an understanding of their cellular role, and the identification of substrates has been a long-standing problem in protein phosphorylation research. However, transfection methods have been developed for *T. brucei* and the application of reverse genetic methods should help greatly in the study of the function of protein kinases and phosphatases in this organism. The large amount of background information available about protein kinases is also helpful in this respect. In summary, the study of protein phosphorylation in *T. brucei* to date has yielded very little information about the parasite's biology, but with the recent isolation of a number of protein kinase and protein phosphatase genes from *T. brucei* and other kinetoplastids, it has become clear that these organisms have many of the protein phosphorylation pathways that operate in higher eukaryotes. The availability of *T. brucei* protein kinase coding sequences allows new

approaches to be taken to the further study of protein phosphorylation in this organism.

### 1.4 Overview of thesis

It is clear that the combination of biochemical, biological, and molecular research is the most powerful combination in terms of understanding the function and properties of specific protein kinases in any organism. At the time that the work reported in this thesis was initiated, very little was known about protein kinases in trypanosomatids but two sets of protein kinase activities from *T. brucei* had been partially characterised at the biochemical level in this laboratory. The first of these protein kinases had been found to be expressed in a stage specific manner with much greater activity in the slender bloodstream forms than in the procyclic forms and furthermore had the interesting property of undergoing apparent autophosphorylation. It was decided to characterise this protein kinase activity in more detail and to develop purification methods so that the purified protein could be microsequenced, or used to raise antibodies which could then be used in screening expression libraries. Chapter 3 describes work performed to extend the earlier characterisation of the activity using the original protein kinase assay system while Chapter 4 describes the partial purification of the protein kinase by liquid chromatography. This work involved both the development of new assays for the protein kinase activity and some further characterisation of the protein kinase.

The second protein kinase activity or set of activities were those with properties suggesting a similarity to PKC (Keith *et al.*, 1990; see Section 1.3.1). These had been more completely characterised and because of their reported instability were not considered suitable for the purification approach taken with the autophosphorylating protein kinases. However, when the sequence of an amplified cDNA fragment predicting peptide sequence similar to PKC catalytic domains became available it was decided to clone the full length gene to complement the earlier biochemical characterisation of *T. brucei* PKC-like activities. This work is described in Chapter 5. Each of the three results chapters has its own discussion which is supplemented by an overall discussion (Chapter 6).

Page 44 became empty following corrections

| | | | |
|---|---|---|---|
| Kingdom: | Protista (Unicellular eukaryotes) | | |
| Phylum: | Kinetoplastida | | |
| Order: | Trypanosomatida | | |
| Genus: | Trypanosoma | | |
| Section: | Salivaria | | |
| Subgenera: | Duttonella | Nannomonas | Trypanozoon |
| Species: | *T. vivax* | *T. congolense* | *T. brucei brucei* |
| | | *T. simiae* | *T. b. rhodesiense* |
| | | | *T. b. gambiense* |
| | | | *T. equiperdum* |
| | | | *T. evansi* |

**Fig. 1.1** Classification of Salivarian trypanosomes

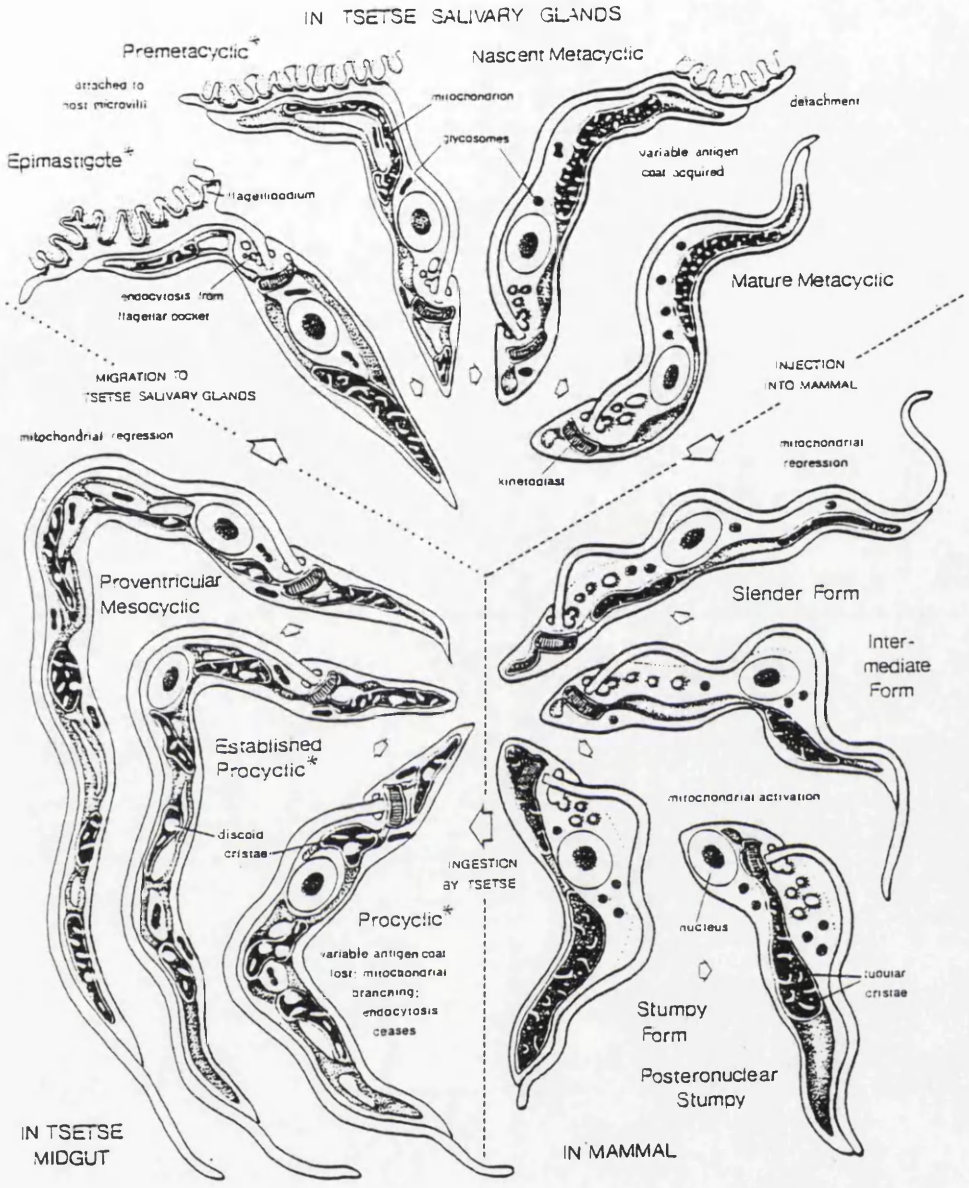



**Fig. 1.2** Developmental cycle of *Trypanosoma brucei* (from Vickerman, 1985)

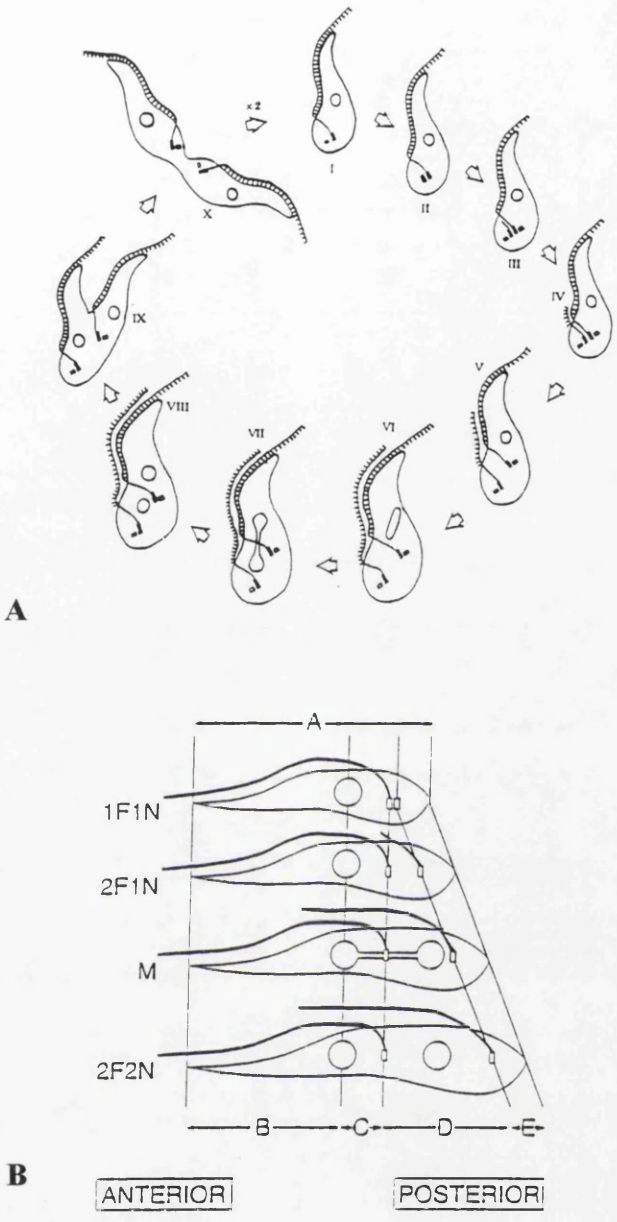



**Fig. 1.3** Morphological changes occurring in the cytoskeleton of the procyclic form of *T. brucei* during the cell cycle. Panel A (from Sherwin and Gull, 1989a) shows a schematic diagram representing the complete cell cycle as deduced from electron micrographs. Panel B (from Robinson *et al.*, 1995) shows procyclic cells at different stages in the cell cycle. Cells at the earliest stage of the cell cycle have 1 flagellum and 1 nucleus (1F1N), the daughter flagellum then begins to lengthen before nuclear division (2F1N). Following mitosis (M) the 2F2N cell continues to elongate before the two cells finally separate. A, B, C, D and E refer to positional descriptors analysed by Robinson *et al.* (1995). Note that the total increase in cell length during cell cycle progression can be accounted for by the increase in inter-basal body distance.

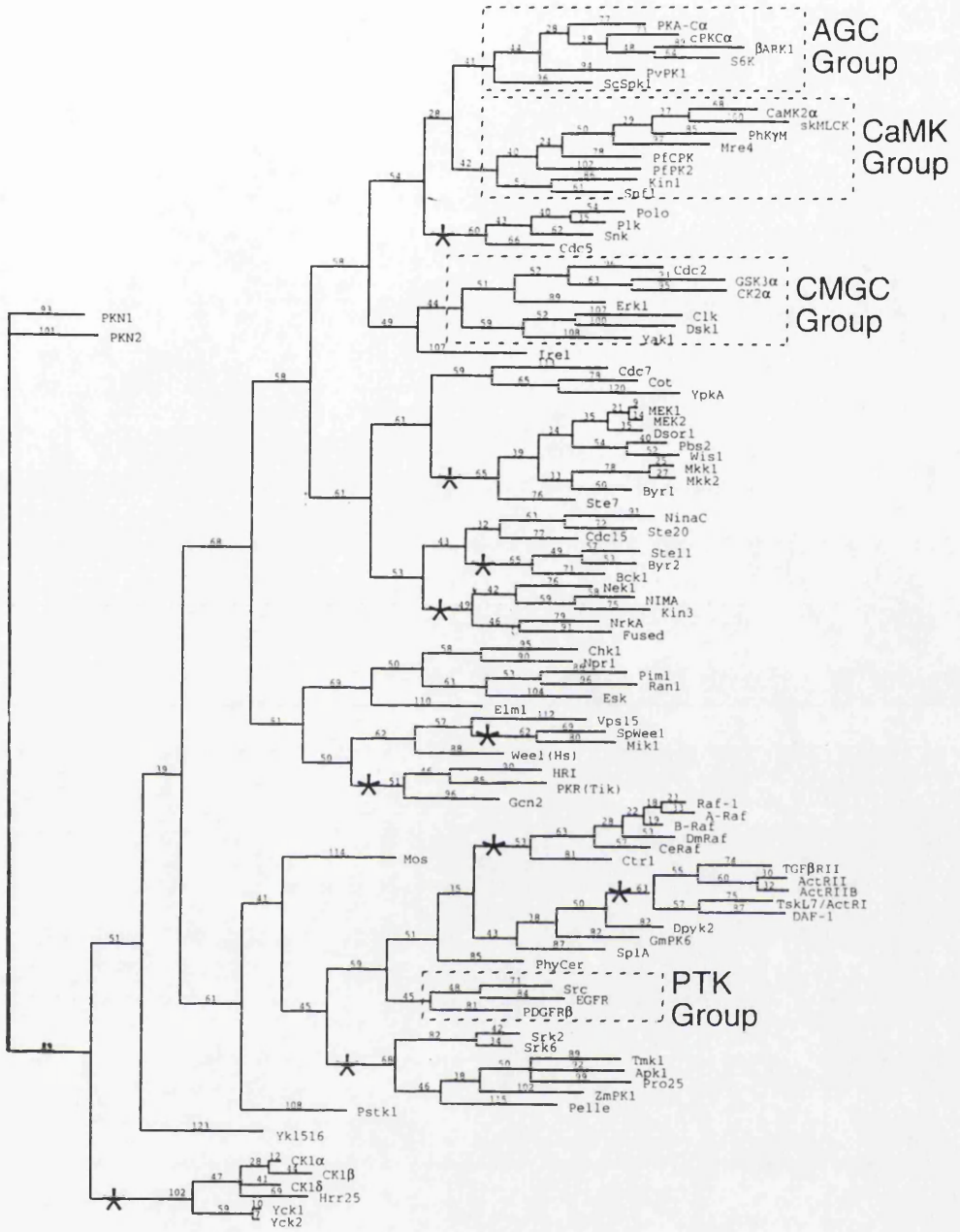



**Fig. 1.4** Phylogenetic tree of the eukaryotic protein kinase superfamily inferred from catalytic domain amino acid sequence alignments (from Hanks and Hunter, 1995a+b). Only representative members have been included for the AGC, CaMK, CMGC and PTK groups of protein kinases (separate trees are given for each of these groups in the references quoted above). Branch lengths indicate the number of amino acid substitutions to reach hypothetical common ancestors at internal nodes. Asterisks indicate branches leading to defined protein kinase families.

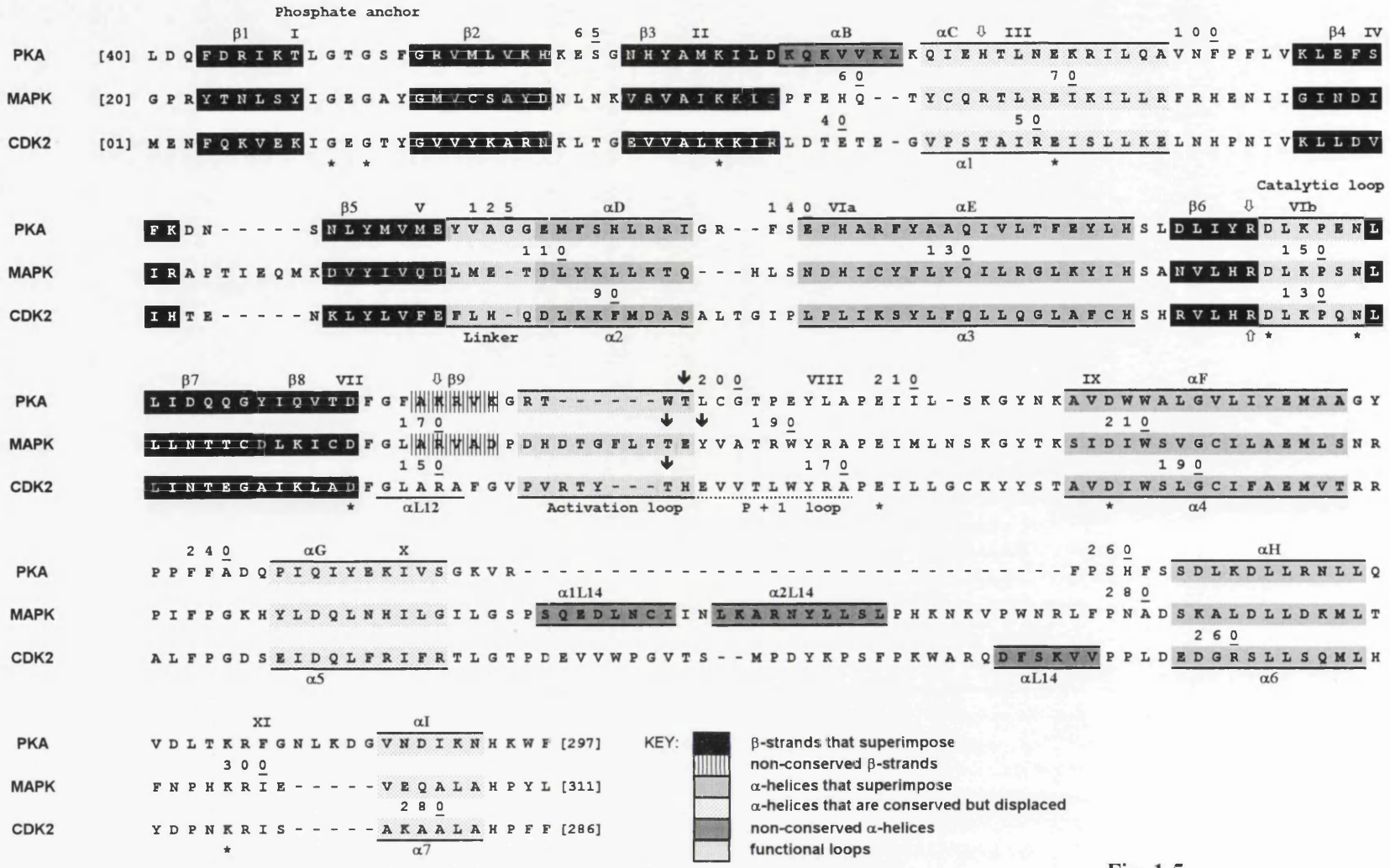

Phosphate anchor
PKA
MAPK
CDK2
Linker
Activation loop
P + 1 loop
Catalytic loop
KEY:
β-strands that superimpose
non-conserved β-strands
α-helices that superimpose
α-helices that are conserved but displaced
non-conserved α-helices
functional loops

Fig. 1.5

**Fig. 1.5** (opposite) Sequence alignment of PKA, MAPK, and CDK2 (slightly adapted from Taylor and Radzio-Andzelm, 1994). Secondary structure motifs are indicated by shading as shown in the key. Highly conserved residues are marked by asterisks, phosphorylation sites in the activation loops are marked by filled arrows, and ligands or potential ligands to the phosphorylation site are marked by open arrows. Roman numerals indicate conserved catalytic subdomains as defined by Hanks and Quinn (1991).

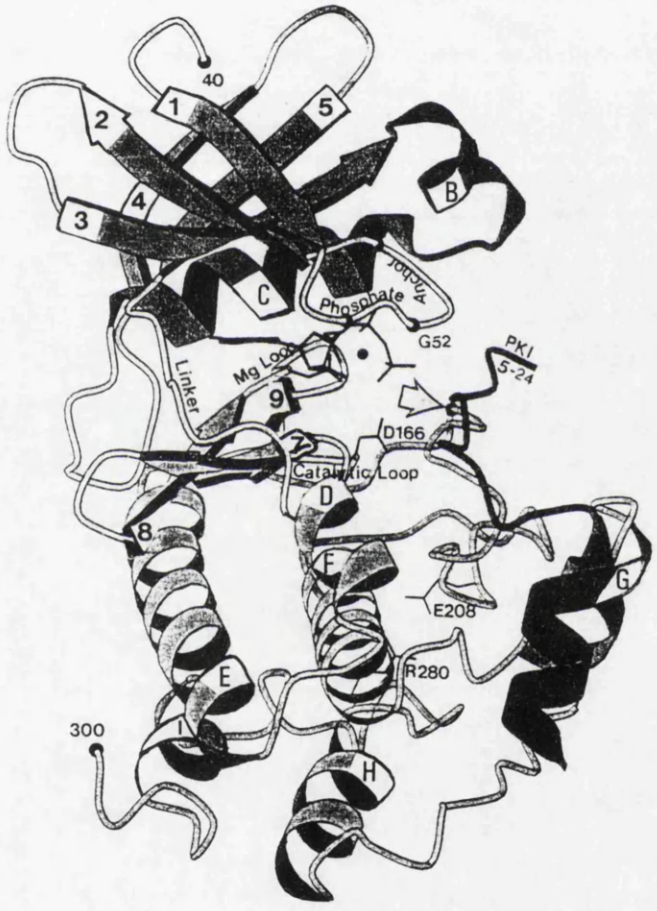



**Fig 1.6** Ribbon diagram of the catalytic subunit of PKA (from Taylor and Radzio-Andzelm, 1994). Only the core of the catalytic subunit (residues 40-300) is shown. ATP and the side chains of Asp166, Asp184, Arg280 and Glu208 are shown in black. The magnesium ion binding the β- and γ-phosphates is shown as a black dot. Phosphorylated Thr197 is indicated by the arrow. PKI (to the right of the figure, labelled) is a small naturally occurring peptide inhibitor of PKA.

| Conserved residue | | | Subdomain | Function |
|---|---|---|---|---|
| PKA | Cdk2 | ERK2 | | |
| $G^{50}$ | $G^{11}$ | $G^{30}$ | I | Loop anchoring the β-phosphate of ATP |
| $G^{52}$ | $G^{13}$ | $G^{32}$ | I | |
| $G^{55}$ | $G^{16}$ | $G^{35}$ | I | |
| $V^{57}$ | $V^{18}$ | $V^{37}$ | I | Lines the adenine binding pocket |
| $K^{72}$ | $K^{33}$ | $K^{52}$ | II | Contacts the α- and β-phosphates of ATP |
| $E^{91}$ | $E^{51}$ | $E^{69}$ | III | Ion pairs with $K^{72}$ |
| $D^{166}$ | $D^{127}$ | $D^{147}$ | VIb | Catalytic base |
| $K^{168}$ | $K^{129}$ | $K^{149}$ | VIb | Interacts with γ-phosphate |
| $N^{171}$ | $N^{132}$ | $N^{147}$ | VIb | Stabilises catalytic loop |
| $D^{184}$ | $D^{145}$ | $D^{165}$ | VII | Chelates $Mg^{2+}$ |
| $E^{208}$ | $E^{172}$ | $E^{195}$ | VIII | Ion pairs with R in subdomain XI |
| $D^{220}$ | $D^{185}$ | $D^{208}$ | IX | Stabilises catalytic loop |
| $R^{280}$ | $R^{274}$ | $R^{229}$ | XI | Forms ion pair with E in subdomain VIII |

**Table 1.1** Highly conserved residues of the protein kinase catalytic domain

| Protein kinases | | | | | |
|---|---|---|---|---|---|
| **Name** | **Organism** | **Class of kinase** | **Size/ kDa** | **Comments** | **Reference** |
| Nrk A | *T. brucei* | NimA-related kinase | 48 | 41% identity with murine Nek1, 30% identity with *Aspergillus* NimA, stage-regulated kinase activity | Gale and Parsons, 1993 Gale *et al.*, 1994b |
| Nrk B | *T. brucei* | NimA-related kinase | 48 | 97% identity with Nrk A | Gale and Parsons, 1993 |
| KFR1 | *T. brucei* | MAP kinase | 40 | 42% identity with KSS1 (*S. cerevisiae*) | Hua and Wang, 1994 |
| | *T. brucei* | MAP kinase | | Different from KFR1 | Accession number Z54341 |
| LdLPK | *L. chagasi* | MAP kinase kinase | 42 | 37% identity with Xenopus MAPKK | Accession number U16029 |
| TbCRK1 | *T. brucei* | cdc2-related kinase | 34 | 54% identity with Hscdc2, homologue of LmmCRK1 | Mottram *et al.*, 1993; Mottram and Smith, 1995 |
| TbCRK2 | *T. brucei* | cdc2-related kinase | 39 | 49% identity with Hscdc2 | Mottram and Smith, 1995 |
| TbCRK3 | *T. brucei* | cdc2-related kinase | 35 | 54% identity with Hscdc2, homologue of LmmCRK3 | Mottram and Smith, 1995 |
| LmmCRK3 | *L. mexicana* | cdc2-related kinase | 35 | 54% identity with Hscdc2 | Mottram and Smith, 1995 |
| CfCRK4 | *C. fasciculata* | cdc2-related kinase | 54 | ~49% identity with Hscdc2, 2 large kinase domain insertions | Brown *et al.*, 1992 |
| TbPK-A3-TbPK-A5 | *T. brucei* | 3 unidentified kinases | ? | PCR products from cDNA, amplified using primers based on kinase subdomains VI and IX | Hua and Wang, 1994 |
| TbPK-B1-TbPK-B2 | *T. brucei* | 2 unidentified kinases | ? | PCR products from cDNA, amplified using primers based on kinase subdomains VI and VIII | Hua and Wang, 1994 |
| TbPLK | *T. brucei* | Polo-like kinase | ? | 53 % identity with CDC5 over kinase domain (TbPK-A2 from Hua and Wang, 1994) | see Chapter 5 |
| TbPKAα | *T. brucei* | Protein kinase A | 38 | 51% - 59% identity with PKAs from yeast to human | M. Boshart, personal communication |
| TbPKAβ | *T. brucei* | Protein kinase A | 38.5 | 52% - 60% identity with PKAs from yeast to human | M. Boshart, personal communication |
| TbPKAγ | *T. brucei* | Protein kinase A | 39 | 51% - 62% identity with PKAs from yeast to human | M. Boshart, personal communication |
| TbZFK | *T. brucei* | Protein kinase C? | 97 | N-terminal $Zn^{2+}$-finger, 24% identity over > 500 amino acids with human PKCι | M. Boshart, personal communication |
| **Protein phosphatases** | | | | | |
| **Name** | **Organism** | **Class of kinase** | **Size/ kDa** | **Comments** | **Reference** |
| TbPP1a | *T. brucei* | Protein phosphatase I | 39 | Type 1 class of catalytic subunit, 66% identity with mammalian PP1 | Evers and Cornelissen, 1990 Erondu and Donelson, 1991 |
| TbPP1b | *T. brucei* | Protein phosphatase I | 39 | Type 1 class of catalytic subunit, > 95% identical with TbPP1a | Evers and Cornelissen, 1990 Erondu and Donelson, 1991 |
| TbPP2 | *T. brucei* | Protein phosphatase 2A | 35 | Type 2A class of catalytic subunit, 55% identity with mammalian PP2A | Erondu and Donelson, 1991 |
| LcPP2C | *L. chagasi* | Protein phosphatase 2C | 42 | Type 2C class of catalytic subunit, 30.8% identity with rat 42 kDa PP2C over 302 amino acids | Burns *et al.*, 1993a |

**Table 1.2** Protein kinase and protein phosphatase genes from kinetoplastids

# Chapter 2

# Materials and Methods

## 2.1 Protein methods

### 2.1.1 *T. brucei* stocks and preparation of cell extracts

Bloodstream form *T. brucei* were prepared from Sprague-Dawley or Wistar rats inoculated with a *T. b. brucei* cloned monomorphic stock TREU 869, originally isolated in Kiboko, Kenya in 1969. Procyclic *T. brucei* were of the cloned stock EATRO 1125 isolated in Uganda (see Hide *et al.*, 1990 for details of these stocks) and were grown in culture in SDM-79 medium (Brun and Shöenenberger, 1979). Bloodstream *T. brucei* were separated from host blood by anion-exchange chromatography (Lanham and Godfrey, 1970), and were then washed twice in phosphate buffered saline pH 8.0 containing 1% glucose (PBSG) before being frozen as a packed cell pellet at -70 °C until use. Cultured procyclic *T. brucei* were washed twice in PBSG and then frozen as a packed cell pellet at -70 °C.

*T. brucei* extracts were made by resuspending the cells (after thawing at room temperature) in ice-cold 20 mM Tris-HCl pH 7.5, 2 mM EDTA, 0.5 mM EGTA with leupeptin and aprotinin added to a final concentration of 25 μg/ml (3 ml buffer/1 ml packed cells). Triton X-100 and NP40 were then added to a final concentration of 0.1%, and a glass homogeniser was used to break up the cells. Crude extracts were stored at -70 °C until use. In order to prepare clarified extracts to load onto chromatography columns, crude extracts were centrifuged at 100 000 g for 45 minutes in a Beckman VTI 65.2 rotor at 4 °C. After centrifugation the supernatant was removed using a syringe and was stored at -70 °C.

### 2.1.2 Isoelectric focusing and *in situ* protein kinase assay

Gels consisted of 0.3 g Agarose IEF (Pharmacia), 3.6 g sorbitol, 27 ml distilled water and 0.38 ml each of ampholines pH 3.5-10 and pH 5-7. The agarose, sorbitol and water were heated in a boiling waterbath until dissolved and the ampholines added immediately before gel pouring. Gels were poured onto the hydrophilic side of Gelbond (FMC) in an 18.5 x 11.3 cm frame and allowed to set for at least 30 minutes before use. For electrophoresis the gels were placed over a thin film of water on the glass cooling plate of a Multiphor II apparatus (Pharmacia). Electrode strips were wetted in 1 M freshly prepared sodium hydroxide (cathode), and 0.05 M sulphuric acid (anode), and laid on the long edges of the gel. Sample papers were soaked in 50 μl or more of sample and

placed on the gel surface 3 cm from the cathode. IEF markers (Pharmacia) were applied to the gel in the same way. Gels were run at a constant 5 W with voltage limited to 1000 V, current to 20 mA, for 1 hour. After electrophoresis, gels were removed from the apparatus and the electrode strips and underlying gel were removed. Gels were then soaked in cold 50 mM Tris-HCl pH 7.5 for 2 x 10 minutes before being placed in 50 ml of 30 mM Tris-acetate pH 7.4, 4 mM magnesium acetate, 10 mM sodium fluoride for 10 minutes at 37 °C. The labelling reaction was then started by the addition of 50 μCi of [γ-$^{32}$P]ATP (6000 Ci/mmol, New England Nuclear) and 2.5 μl of 2 mM ATP and allowed to continue for 15 minutes at 37 °C before the labelling solution was poured off. Gels were fixed in 10% trichloroacetic acid, 1% phosphoric acid and then soaked in this solution for two hours or more before being rinsed again and left overnight in 30% methanol, 10% acetic acid. The gels were then partially dried by blotting with filter paper and tissues, and finally dried completely at 60 °C. The dried gels were exposed to X-Omat-S film (Kodak) at -70 °C

To obtain molecular weights for radiolabelled proteins separated in IEF gels, the relevant gel bands were located using the autoradiograph, and cut out of the dried gel with scissors. Gel bands were cut into smaller pieces and re-hydrated in separate tubes by soaking in 50 mM Tris-HCl pH 7.2, 3% SDS, 10% glycerol for 1-2 hours. Prior to gel loading SDS-PAGE sample buffer was added and samples were boiled for five minutes before removal of Gelbond fragments and loading onto a SDS-PAGE gel while the agarose was still liquid.

Alternatively the separated proteins were analysed by blotting. Following the labelling step of the assay, gels were rinsed in 50 mM Tris-HCl pH 7.5 and then soaked in a fresh change of this solution for 15 minutes. Gels (still attached to Gelbond) were then taped around the circumference onto a plastic tray. A piece of PVDF paper (previously dipped in methanol and soaked in 50 mM Tris-HCl pH 7.5 for 10-15 minutes) was laid on top of the gel and then overlaid with 1 sheet of 3MM (Whatman) paper moistened in 50 mM Tris-HCl pH 7.5, two sheets of dry 3MM paper, a block of tissues, and finally a weight. The stack was left overnight to allow proteins to transfer to the PVDF, the sheet of PVDF was then rinsed 5 x with distilled water and air-dried before autoradiography.

### 2.1.3 Phosphoamino acid analysis

Bands of phosphate labelled protein were located on PVDF blots by overlaying the autoradiograph on the blot and were then cut out with scissors. Each labelled band was cut up into small pieces that were dipped briefly into methanol, rinsed in water, then placed in separate screw-capped microcentrifuge tubes. A minimal volume of 5.7 M HCl sufficient to cover the PVDF strips was added and the tubes were then heated for 1 hour at 110 °C with the caps tightly closed. The PVDF strips were then removed from the tubes which were then centrifuged briefly to collect the liquid. The tubes were then placed in a vacuum oven at 80 °C without lids until the liquid had evaporated.

The samples were then analysed by 2-dimensional electrophoresis on thin-layer chromatography plates according to the method of Cooper *et al.* (1983). Buffer 1.9 was 4.4 ml 100% formic acid, 15.6 ml glacial acetic acid made up to 200 ml with distilled water. Buffer 3.5 was 1 ml pyridine and 10 ml glacial acetic acid made up to 200 ml with water. Each dried sample was re-suspended in 2 µl of Buffer 1.9 containing 10 mg/ml each of phosphoserine, phosphothreonine, and phosphotyrosine as internal standards. The samples were spotted onto cellulose thin-layer chromatography plates (four evenly spaced sample origins/plate) with care being taken to minimise the spreading of each sample. For the 1st dimension electrophoresis each TLC plate was wetted with Buffer 1.9 by blotting with 3MM paper soaked in the buffer. The plate was then placed on the aluminium cooling plate of a Multiphor II apparatus (Pharmacia) and run at 1400 V (limited to 40 W) for 22 minutes with water-cooling. The TLC plate was then dried with a stream of warm air and wetted with Buffer 3.5. It was then placed back in the apparatus at 90° to its previous orientation and electrophoresed for 30 minutes at 1000 V. Following the second dimension run the plate was again dried, and then sprayed with 0.5% ninhydrin in butan-1-ol to develop the internal standards, and then autoradiographed for 1-3 weeks at -70 °C.

### 2.1.4 Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was performed using the discontinuous buffer system of Laemmli (1970). Large format gels were run in Bio-Rad Protean II vertical slab gel apparatus, while minigels were run in Bio-Rad Mini-Protean II apparatus. The gels poured were 1.5 mm thick in both cases. Initially, 12% polyacrylamide gels were prepared using bis-

acrylamide as the crosslinking agent but in some experiments 7.5% gels were prepared using piperazine di-acrylamide as the crosslinker.

Stock solutions made up were 1.5 M Tris-HCl pH 8.8, 10% w/v SDS, and 3% w/v polyacrylamide. A commercial 30% acrylamide/bis-acrylamide (37.5:1) stock (Northumbria Biologicals) was stored in the dark at 4 °C. To prepare 30 ml of a 12% resolving gel, 12 ml of acrylamide stock, 6.9 ml distilled water, 7.5 ml 1.5 M Tris-HCl pH 8.8, 300 µl of 10% SDS, 3 ml of 3% polyacrylamide, and 300 µl of fresh 10% ammonium persulphate were mixed before addition of 12 µl of tetra-methyl-1,2-diaminoethane (TEMED, Sigma) immediately before pouring. Resolving gels were overlaid with water-saturated 2-butanol and allowed to set for about 1 hour. The top of the resolving gel was washed with distilled water before the stacking gel was applied. The mixture used for 10 ml of stacking gel (4%) was 1.7 ml acrylamide stock, 5.8 ml distilled water, 1.25 ml 1 M Tris-HCl pH 6.8, 1 ml 3% polyacrylamide, 100 µl SDS, 100 µl 10% ammonium persulphate, and 10 µl TEMED.

For acrylamide/piperazine di-acrylamide gels a 30% acrylamide/piperazine di-acrylamide stock solution (reagents from Bio-Rad) was made up in distilled water according to the manufacturer's instructions, filtered and stored in the dark at 4 °C. For 30 ml of a 7.5% resolving gel mixture, 7.5 ml of the acrylamide stock, 14.5 ml of distilled water, 7.5 ml of 1.5 M Tris-HCl, 0.3 ml of 10% SDS and 150 µl of freshly prepared 10% ammonium persulphate were mixed before the addition of 15 µl of TEMED to start the polymerisation reaction. For 10 ml of a 4% stacking gel, 1.3 ml of acrylamide solution, 6.1 ml of distilled water, 2.5 ml of 0.5 M Tris-HCl pH 6.8, 100 µl of 10% SDS and 50 µl of 10% ammonium persulphate were mixed before addition of 10 µl of TEMED.

Samples were prepared for SDS-PAGE by addition of SDS-PAGE sample buffer and then boiling for 5 minutes, before gel loading or storage at -20 °C. Radiolabelled protein molecular weight markers (14.3 kDa-200 kDa) were from Amersham, while non-radioactive markers were from Sigma (29 kDa-205 kDa and 14.2 kDa-66 kDa). The 4 x sample buffer stock consisted of 200 mM Tris-HCl pH 6.5, 8% w/v SDS, 8 mM EDTA, 20% v/v 2-mercaptoethanol, and 40% v/v glycerol, with bromophenol blue added to 0.002%. The running buffer was 50 mM Tris, 384 mM glycine, 2 mM EDTA, and 0.1% SDS. Large gels were electrophoresed overnight at 50-60 V and run at 100 V the next

day until the bromophenol blue band had reached the desired position. Minigels were run at 100 V for approximately 2 hours with the position of the bromophenol blue band again being used to judge the progress of electrophoresis. Polyacrylamide gels were silver stained using a commercial silver stain kit (Sigma) following the manufacturer's instructions. Following any post-electrophoresis process, such as silver staining, gels were either dried onto 3MM paper in a Bio-Rad vacuum gel dryer, or between cellophane sheets in a Hoefer Easy-breeze gel dryer. Gels containing radioactive material were exposed to film (Kodak X-Omat S) in a gel cassette at -70 °C.

### 2.1.5 2-dimensional protein electrophoresis

1st dimension gels for 2-dimensional protein electrophoresis were prepared in glass tubes 7.5 cm long with an internal diameter of 1.5 mm using a Hoefer apparatus according to the manufacturer's instructions. The gels contained (per 3.5 ml): 1.95 g urea, 467 µl 30% acrylamide solution (37.5:1 acrylamide:bis-acrylamide, Northumbria Biosciences), 700 µl 10% NP-40, 177 µl 40% ampholines (50:50 mixture of ampholines of pH ranges 3.5-10 and 5-7, Pharmacia), 800 µl distilled water, 10 µl ammonium persulphate, 7 µl TEMED. Samples (10-30 µl) were resuspended in 9.8 M urea, 4% NP-40, 2% ampholines (pH 3.5-10), 100 mM DTT (final concentrations). The electrode buffers were 20 mM NaOH (cathode) and 0.085% phosphoric acid (anode). Gels were prerun at a constant 250 V for 30 minutes before sample loading and were then run at 500 V for 3 hours. Gels were extruded from the tubes using a syringe and adaptor (Hoefer), equilibrated in 0.25 M Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 50 mM DTT and frozen until use. 2nd dimension gels were 7.5% SDS-PAGE minigels prepared using piperazine di-acrylamide as crosslinker as described in Section 2.1.4

### 2.1.6 Liquid chromatography

The HPLC system used was a Waters 625 LC system with a single wavelength (280 nm) UV absorbance spectrophotometer and an automatic fraction collector (Gilson). This system can accurately mix buffers from four components and can be programmed to mix buffer gradients (all gradients used were linear). Following each chromatography experiment, the system was flushed out with water and then filled with 10% isopropanol

to prevent bacterial contamination. Tris and NaCl stock solutions were stored at 4 °C and all solutions were filtered through 0.22 μm filters shortly before use.

Anion-exchange chromatography was performed using a Waters AP-1 (1 cm internal diameter) glass column pre-packed with Protein-Pak DEAE 8HR resin at a flow rate of 0.8 ml/min and fractions were collected every half minute. Buffers used for anion-exchange chromatography were mixed by the HPLC fluid handling system from separate bottles containing 100 mM Tris-HCl, 100 mM Tris-base, 1 M NaCl and distilled water. Size-exclusion chromatography was performed using a Waters Protein Pak 300SW glass column (8 x 300 mm). The flow rate was 0.8 ml/minute and fractions were collected every 30 seconds. In order that HPLC could be performed using H-9 Sepharose (prepared as described in Section 2.1.8) a Waters AP minicolumn (5 cm long, 5 mm internal diameter) was filled using the manufacturer's method for slurry packing. Separations were undertaken using flow rates of up to 0.5 ml/minute.

### 2.1.7 Slot blot kinase assay

Suitably sized pieces of PVDF to fit into the slot blot apparatus (Gibco BRL 24-well filtration manifold) were soaked in methanol for 15 minutes and then left submerged in 30 mM Tris-acetate, 4 mM magnesium acetate, 10 mM sodium fluoride until use. Immediately before sample blotting, a PVDF sheet was removed from the soaking buffer and placed on the lower bed of the apparatus. The upper plate of the blotter was then screwed down and each slot was loaded with an aliquot from a single HPLC fraction (usually 100 μl). A vacuum was then applied briefly to draw the liquid through the PVDF but without causing excessive drying. After blotting the PVDF sheets were removed from the apparatus and were returned to a solution of 30 mM Tris-acetate, 4 mM magnesium acetate, 10 mM sodium fluoride to keep the membrane wet. All procedures between sample blotting and the labelling step were performed as rapidly as possible to minimise protein denaturation or diffusion. The blots were then transferred to plastic bags containing 10 ml of 30 mM Tris-acetate, 4 mM magnesium acetate, 10 mM sodium fluoride and incubated at 37 °C for 10 minutes before a corner was cut off each bag to allow the addition of the label (0.5 μl 2 mM ATP/1 μl [$\gamma$-$^{32}$P]ATP 6000 Ci/mmol in labelling buffer). After re-sealing, the bags were incubated for a further 15 minutes at 37 °C. Following the labelling step, the labelling solution was poured away and the blots

were rinsed 3 times in distilled water and then air dried. The dried membranes were exposed to Kodak X-Omat S film at -70 °C.

### 2.1.8 Liquid kinase assay for column fractions

Column fractions (usually 25 µl aliquots in 20 mM Tris-HCl pH 5.7, with varying NaCl concentrations) were mixed with an equal volume of 60 mM Tris-HCl pH 7.4, 8 mM $MgCl_2$. An ATP labelling mix was prepared by adding 1 µl [$\gamma$-$^{32}$P]ATP (6000 Ci/mmol, NEN) and 0.5 µl 2 mM ATP to 198.5 µl water, and 1 µl of this was added per 50 µl of 1:1 diluted column fractions. Reactions were incubated for 15 minutes at 37 °C, the reaction stopped by addition of SDS-PAGE sample buffer and boiling (5 minutes). Samples were then loaded onto SDS-PAGE gels or stored at -20 °C until use. In some experiments the 2x Tris-HCl/$MgCl_2$ buffer normally used was substituted by 5x or 10x buffer with appropriate adjustment to the volume of column fraction used. The labelling step of the liquid kinase assay was also used to prepare pre-labelled (autophosphorylated) protein kinase for experiments to test the binding of the 60 kDa protein kinase to chromatography matrices (Sections 4.2.2 and 4.2.4). Following the 37 °C incubation, buffer exchange was performed on the samples using Centricon 30 concentrators (see Section 2.1.9) to remove excess unincorporated labelled ATP and, if necessary, to change the buffer conditions. Following the Centricon step the samples were either used immediately or frozen at -20 °C.

### 2.1.9 Preparation of H-9 Sepharose

H-9 Sepharose 4B was prepared essentially according to the method of Inagaki *et al.* (1985). Cyanogen bromide-activated Sepharose 4B (Sigma) was washed with distilled water and re-suspended in 50 mM borate buffer pH 8.0, 0.5 M NaCl. H-9 (Biomol) dissolved in the borate buffer (10 µmol in 250 µl) was added to 0.9 ml of the settled Sepharose which was then made up to a total volume of 2.5 ml with 50 mM borate buffer pH 8.0, 0.5 M NaCl. The reaction was allowed to proceed for 2 hours at 30 °C with regular mixing by inversion of the tube. After incubation, the Sepharose was washed with 30 ml of the borate buffer, re-suspended in excess 1 M ethanolamine pH 8.0 and incubated for 8 hours at room temperature with regular mixing. The Sepharose was then washed extensively in distilled water and stored in 25 mM Tris-HCl pH 7.0 containing 2

mM EGTA and 50 mM 2-mercaptoethanol. A control Sepharose matrix was prepared similarly except that borate buffer without H-9 was used in the initial incubation before blocking with ethanolamine.

### 2.1.10 Protein concentration/buffer exchange

Centricon 30 concentrators (Amicon) were used for concentration and buffer exchange of column fractions and liquid kinase assays according to manufacturer's instructions.

## 2.2 Nucleic acid techniques

### 2.2.1 Growth and transformation of *E. coli*

*E. coli* strains used for plasmid work were either XL-1 Blue or INVαF′ (Invitrogen) and were grown in LB medium or on LB agar plates (Sambrook *et al.*, 1989). Ampicillin was added to a final concentration of 50 μg/ml for cells transfected with ampicillin resistant plasmids. XL-1 Blue cells were made competent by treatment with 50 mM $CaCl_2$ and were transformed according to standard protocols. Manufacturer's instructions were followed for transformation of INVαF′ supercompetent cells. *E. coli* strain Y1090($r^-$) were used for growth of λgt11 and was cultured in LB media containing ampicillin. LB agar plates used for growth of Y1090 cells infected with phage contained 10 mM $MgSO_4$.

### 2.2.2 Bacterial vectors and preparation of vector DNA

Bacterial vector pCRII (Invitrogen) was used for cloning of PCR products according to the manufacturer's protocol. pGEM7Zf(+) (Promega) or Bluescript II SK(+) (Stratagene) were used for all other DNA cloning. Cells containing these plasmids were grown as 10 ml cultures in LB media containing 50 μg/ml ampicillin in 20 ml plastic universal tubes which were incubated overnight at 37 °C with vigorous shaking. Wizard Minipreps (Promega) were used for the extraction of plasmid DNA (3 ml cell culture per miniprep) according to the manufacturer's protocol.

### 2.2.3 Manipulations of plasmid DNA

DNA restriction digests were performed using enzymes and buffers from Gibco BRL or from Boehringer Mannheim according to manufacturer's instructions. Ligations were

performed using T4 ligase and ligase buffer from Gibco BRL and were incubated overnight at 16 °C, except for ligations performed in the construction of deletion sets using the Erase-A-Base kit (Promega), where the manufacturers' reagents (containing polyethylene glycol) and protocols were used to allow shorter incubation times. When required, the ends of linear vector DNA were dephosphorylated to prevent vector self-ligation using calf intestinal phosphatase (Boehringer Mannheim) according to standard protocols (Sambrook *et al.*, 1989).

### 2.2.4 Electrophoresis of DNA in agarose gels

Gels were prepared with electrophoresis grade agarose from Gibco BRL and were run in 40 mM Tris-acetate, 1 mM EDTA (TAE) buffer prepared according to Sambrook *et al.* (1989) containing 0.5 μg/ml ethidium bromide in apparatus from Pharmacia or BRL. Type I sample loading buffer was prepared and used according to Sambrook *et al.* (1989). DNA size markers (1 kb ladder) were from Gibco BRL. DNA bands were visualised by UV transillumination and gels were photographed using a Polaroid camera. The Geneclean II kit (Bio 101) was used for all DNA extraction from agarose gels according to the manufacturer's protocol.

### 2.2.5 Southern blotting of agarose gels

Agarose gels were washed twice (15 minutes per wash) in denaturing solution (0.5 M NaOH, 1.5 M NaCl) and then twice (15 minutes per wash) in neutralising solution (0.5 M Tris-HCl pH 7.5, 3 M NaCl). Gels were then placed, inverted, on 3MM paper (Whatman) on a platform with the ends of the paper immersed in 20x SSC (175.3 g NaCl, 88.2 g sodium citrate made up to 1 litre in distilled water with the pH adjusted to 7.0 with sodium hydroxide). A sheet of Hybond-N membrane of the same size as the gel was then placed on top of the gel and this was then covered by 3 sheets of 3MM paper soaked in 3x SSC, a block of paper towels and finally a weight. Transfer was allowed to proceed overnight. Blots were dried in an oven at 80 °C for 10 minutes and the DNA was cross-linked to the membrane by placing on a UV transilluminator (DNA side down) for 7 minutes. Hybridisations using 22mer oligonucleotide probes prepared by end-labelling or larger random primed DNA fragments (see Section 2.2.6) were performed as follows. Blots were prehybridised in 0.5 M Phosphate buffer pH 7.0, 7% SDS (prepared

by adding equal volumes of 1 M Phosphate buffer and 14% SDS; 1M phosphate buffer was prepared by titrating 1 M $Na_2HPO_4$ with 1 M $NaH_2PO_4$ to pH 7.0) for 30 minutes at 65 °C in a rotating bottle in a hybridisation oven. The buffer was then poured off and replaced by 10 ml of the same buffer containing the labelled probe (denatured by boiling immediately before use). Hybridisation was allowed to proceed overnight at 65 °C. Blots were then rinsed in 2x SSC, 0.1% SDS and then washed 3 times (15 minutes per wash) at 65 °C in the same solution before being either dried or sealed wet in a bag for autoradiography. When required blots (kept wet after hybridisation) were stripped by immersion in 500 ml of 0.1% SDS and heating in a microwave oven at full power for 15 minutes. This step was then repeated with fresh 0.1% SDS, and the blot then transferred to prehybridisation buffer.

### 2.2.6 Labelling of oligonucleotide and DNA probes

Short oligonucleotides (22mers) were prepared as probes by end-labelling with T4 polynucleotide kinase (Gibco BRL) according to the protocol described in Sambrook *et al.* (1989) and were then ethanol precipitated before use. Larger DNA fragments contained in LMP agarose were labelled with [$\alpha$-$^{32}$P]dCTP by random priming using a kit (Boehringer Mannheim) according to manufacturer's instructions.

### 2.2.7 Polymerase chain reaction

Oligonucleotide primers were supplied by the oligonucleotide synthesis service of the University of Glasgow Biochemistry Department in a solution of ammonium hydroxide. Primers were de-salted by ethanol precipitation according to standard protocols. PCR reactions were prepared according to the instructions in the Perkin-Elmer GeneAmp kit, using polymerase, buffer, and dNTPs from Perkin-Elmer or Advanced Biotechnologies. In all reactions a "hot-start" was performed by denaturing at 94 °C for 5 minutes before addition of the polymerase. In all PCR reactions the denaturation step was 94 °C for 1 minute and the annealing step was also for 1 minute. A temperature of 72 °C was used for the extension step. Annealing temperatures, extension times, and cycle numbers specific to particular reactions are described in Section 5.2. A Hybaid thermal cycler was used.

### 2.2.8 DNA sequencing

Deletion sets for DNA sequencing were prepared using the Erase-A-Base kit (Promega) according to the manufacturer's protocol. Sequencing reactions were performed on double-stranded plasmid DNA using the Sequenase kit (USB) with [$\alpha$-$^{35}$S]-dATP as the label according to manufacturer's instructions. Sequencing products were analysed by vertical polyacrylamide gel electrophoresis using 8% gels prepared in the BRL S2 apparatus and were run in 45 mM Tris-borate, 1 mM EDTA (TBE) buffer prepared according to Sambrook *et al.* (1989) at a constant 65 W. The migration of "Sequenase" tracking dyes was used to judge the run length required. Gels were fixed for 20 minutes in 10% methanol/10% acetic acid, were then rinsed briefly with water and blotted to 3MM paper (Whatman) before vacuum drying in a Bio-Rad gel dryer. Gels were then exposed to Kodak X-Omat S film overnight or longer at room temperature. Sequences were read manually and were analysed using the Wisconsin GCG software package.

### 2.2.9 Isolation of *T. brucei* genomic DNA and RNA, purification of poly(A)$^+$ RNA and synthesis of 1st strand cDNA

*T. brucei* genomic DNA was isolated from bloodstream form cells using a standard procedure (Borst *et al.*, 1980). *T. brucei* total RNA was isolated from bloodstream forms using the RNAgents kit (Promega) and the manufacturer's protocol was followed. *T. brucei* poly(A)$^+$ RNA was purified from *T. brucei* total RNA using the Poly(A) Quik kit (Stratagene) according to the manufacturer's instructions. First strand cDNA was prepared using the Riboclone cDNA synthesis kit (Promega) according to the manufacturer's instructions, except that the second strand cDNA synthesis step was omitted.

### 2.2.10 Screening of a $\lambda$ cDNA library

*E. coli* strain Y1090 were prepared as plating cells from a 50 ml culture grown in LB medium containing 50 $\mu$g/ml ampicillin to an optical density (650 nm) of 0.5. The cells were pelleted by centrifugation and re-suspended in 5 ml SM buffer (5.8 g NaCl, 2 g $MgSO_4.7H_2O$, 50 ml 1 M Tris-HCl pH 7.5, 5 ml 2% gelatin made up to 1 litre with distilled water) and stored at 4 °C for up to 1 week. To prepare plates (8.5 cm diameter) of bacteria infected with phage, serial dilutions of the phage stock were prepared in SM

buffer and 100 µl of each phage dilution were mixed with 100 µl of plating cells and incubated at 37 °C for 10-15 minutes. After incubation, 2.5 ml of top agarose (containing 50 µg/ml ampicillin) at 55 °C was added to each tube and the agarose containing the infected bacteria was poured immediately onto pre-warmed LB agar plates. The plates were then incubated overnight at 37 °C and plaques were counted the next day to titre the phage stocks.

Plaques were blotted to nylon filters (GeneScreen Plus, NEN) and DNA was immobilised according to the method described in Sambrook *et al.* (1989). Filters were hybridised to a 171mer oligonucleotide probe prepared from a PCR product similar to TbPK-A2 (Section 5.2.1) by a random priming labelling method (Section 2.2.6). The hybridisation solution used was (per 100 ml): 50 ml Phosphate buffer stock solution (71 g $Na_2HPO_4$, 2 ml 85% orthophosphoric acid made up to 1 litre with distilled water), 35 ml 20% SDS, 0.2 ml 0.5 M EDTA pH 8.0, 14.8 ml distilled water. Filters were prehybridised in the hybridisation solution at 65 °C in a hybridisation oven for 30 minutes before addition of the probe (denatured immediately before use by boiling). The hybridisation reaction was allowed to proceed overnight at 65 °C. Filters were then rinsed in preheated 1 x SSC/5% SDS and were then washed 3 times (15 minutes per wash) in the same solution at 65 °C. Filters were then dried and autoradiographed. Agarose plugs from hybridising plaques were picked using a glass pasteur pipette. The agarose from each plaque was placed in a microcentrifuge tube containing 1 ml SM buffer and a small volume of chloroform. Tubes containing λ stocks were stored at 4 °C.

### 2.2.11 Isolation of λ DNA

DE-52 cellulose (Whatman) was prepared for use by addition of several volumes of 0.05 M HCl until the supernatant was below pH 4.5. Concentrated NaOH was then added with gentle stirring of the suspension until the pH reached 7.0. The supernatant was poured off and several volumes of LB medium were added. This was repeated with fresh LB medium until there was no pH difference between the LB medium and the DE-52 slurry. The slurry was adjusted to approximately 75% resin/25% LB medium and stored at -20 °C until use.

Bacteria were infected with phage from plaques picked as described in Section 2.2.10 at a phage concentration known to give confluent lysis from previous titrations,

and were spread in 0.3% top agarose onto LB agar plates containing 10 mM magnesium sulphate. After overnight incubation, 2 ml of SM buffer was added to each plate and the top agarose was scraped off into a tube and centrifuged to pellet the solid material. The supernatant was then added to the equilibrated DE-52 resin in microcentrifuge tubes (0.8 ml of supernatant to 0.5 ml of slurry) and the tubes were inverted about 50 times. The tubes were then centrifuged for 5 minutes and the supernatant from each was put in a fresh tube. Extraction buffer (0.5 M Tris-HCl pH 8.0, 0.25 M EDTA, 2.5% SDS was added (150 µl to 800 µl of supernatant) and the tubes were heated at 65 °C for 15 minutes. After cooling to room temperature, 200 µl of 8 M ammonium acetate was added to each tube and the tubes were left on ice for 15 minutes to precipitate the phage proteins and RNA. The tubes were then centrifuged for 10 minutes at 10 000 g, the supernatant removed and extracted twice with phenol/chloroform. Isopropanol was added to the aqueous supernatant from the second extraction and the tubes were incubated for 10 minutes at room temperature. The DNA was then pelleted by centrifugation at 10 000 g for 10 minutes and washed with 1 ml 70% ethanol before being resuspended in a small volume of 10 mM Tris-HCl pH 7.4, 1 mM EDTA.

# Chapter 3

## Identification of an autophosphorylating protein kinase from *Trypanosoma brucei*

## 3.1 Introduction

To investigate protein kinase activities in *T. brucei*, Keith *et al.* (1990) adapted an isoelectric focusing gel *in situ* protein kinase assay (Grove and Mastro, 1987). When the procedure was modified for the investigation of autophosphorylating protein kinase activities by the exclusion of exogenous substrates, up to eight gel bands containing phosphate labelled proteins were detected from total extracts of bloodstream *T. brucei*. Furthermore, when these bands were individually cut out of the dried gels, re-hydrated, and run in SDS-PAGE, it was found that the predominant phosphate labelled protein from each IEF gel band migrated with an apparent molecular weight of 60 kDa. Thus it seemed that all of protein kinase activities detected might be isoenzymes of each other. When identical conditions were used to detect protein kinase activities in extracts of procyclic *T. brucei* only two bands containing protein kinase activity were detected which had isoelectric points identical to those of two of the bloodstream activities, although the phosphate labelling was considerably weaker (see below and Hide *et al.*, 1994).

The apparent stage-specificity of these protein kinase activities was very interesting. In higher eukaryotes, cellular growth and differentiation are partly controlled by protein kinases and phosphatases, often acting in cascades. Far less is known about the role of protein phosphorylation in evolutionarily primitive eukaryotes such as *T. brucei*, but a stage-regulated protein kinase activity might play a role in controlling the transition from one life-cycle stage to the next. Based on these initial results, it was decided to perform experiments to confirm the profile of protein kinases obtained for bloodstream and procyclic *T. brucei* and then undertake a more extensive characterisation of these activities.

The procedure for the IEF gel *in situ* assay and subsequent analysis of labelled proteins is summarised in Fig. 3.1 (described in detail in Chapter 2). As depicted in panel A, electrophoresis is performed in a thin horizontal IEF gel with samples and isoelectric point markers loaded onto the gel surface soaked into sample application papers. Following electrophoresis the gel is washed to remove ampholines and other contaminants and is then equilibrated in a protein kinase assay buffer at 37 °C before the kinase reaction is started by addition of [$\gamma$-$^{32}$P]ATP. As shown in panel B, the steps subsequent to the labelling reaction depend on the nature of the analysis required. The

gel is either washed under protein fixing conditions and dried or is washed briefly under non-fixing conditions and capillary blotted to PVDF. The dried gel or blot is then autoradiographed to reveal the positions of proteins that have become phosphate labelled. Molecular weight estimates for phosphate labelled proteins were obtained by cutting out the agarose containing the labelled proteins and loading it into SDS-PAGE gels. Phosphoamino acid analysis was performed on labelled proteins which were capillary blotted onto PVDF membranes from IEF gels.

### 3.2 Analysis of autophosphorylating protein kinase activities in *Trypanosoma brucei* using the IEF gel *in situ* assay

*T. brucei* extracts were prepared and assayed for autophosphorylating protein kinase activity (see Materials and Methods). Fig. 3.2 panel A depicts an autoradiograph of a dried-down IEF gel used in the *in situ* protein kinase assay with [$\gamma$-$^{32}$P]ATP as the label, and in the absence of exogenous substrate. Equal amounts of procyclic (P) and bloodstream (BS) *T. brucei* total extracts were loaded at the origin (O). The positions and isoelectric points of markers used to calibrate the gel are shown to the right of the figure. Two labelled bands were detected in the assay of procyclic *T. brucei* extracts and up to eight in the assay of bloodstream extracts. Six of the bloodstream bands (1-6) form a cluster with isoelectric points between 5.2 and 6.5, of which bands 1 and 2 were the most strongly labelled and reproducibly detected in subsequent assays. The activity at isoelectric point 7.1 was always lower than the activity of bands 1 and 2 but was routinely detected unless the overall activity of the extract was low. The activity at isoelectric point 9.25 was often low but was seen in the majority of the assays. The isoelectric points of the procyclic bands 1 and 2 are 7.1 and 9.25 respectively, and these bands were always weakly labelled. These possibly represent the same activities as those of bloodstream bands 7 and 8 at the same isoelectric points. However, procyclic band 1 and bloodstream band 8 both focus near the cathode and it is possible that they are at a pH discontinuity in the gel.

The molecular weights of proteins in $^{32}$P-labelled IEF gel bands were determined by SDS-PAGE under reducing conditions. Individual bands were cut out of dried IEF gels, re-hydrated in SDS-PAGE sample buffer and fractionated by SDS-PAGE. Fig. 3.2 panel B shows the autoradiograph of an SDS-PAGE gel used to analyse phosphate

labelled proteins from a bloodstream *T. brucei* extract. Lanes 1-6 contain bands 1, 2, 3, 4, 7, and 8 respectively. The positions and sizes (in kDa) of protein molecular weight markers are indicated on the left of the figure. It can be seen that the predominant phosphate labelled protein in each lane has an apparent molecular weight of 60 kDa, although lane 3 clearly has a labelled band below 60 kDa in size, and the main band could be a doublet. It was found consistently in many experiments that a 60 kDa band accounted for almost all of the labelling of any of the IEF bands and was frequently the only phosphate labelled protein present. Similarly, the two IEF bands from procyclic extracts were also found to migrate with an apparent molecular weight of 60 kDa confirming the earlier results.

To investigate whether the autophosphorylating protein kinase activities are cytosolic or membrane associated, the *T. brucei* total extracts were fractionated by ultracentrifugation. Following extraction with TX-100 and NP-40 a total extract was centrifuged at 100 000 g to spin down membrane fractions. Fig. 3.3 shows the autoradiograph of an IEF gel used in an assay to compare a total extract and the 100 000 g supernatant fraction of an extract. Lanes 1 and 2 are loaded with duplicate samples of total extract, lanes 3 and 4 (also duplicates) are loaded with the high speed supernatant fraction. There is high background labelling around the origin (O) but the pattern of labelled bands migrating towards the anode side of the origin is clearly very similar between the two pairs of samples. This result suggests a cytoplasmic rather than a nuclear or membrane location for the protein kinase activity.

### 3.3 Phosphoamino acid analysis

Serine, threonine and tyrosine are the residues phosphorylated by members of the main eukaryotic protein kinase superfamily. In order to determine whether the autophosphorylating activity was a serine/threonine or tyrosine kinase, an analysis was carried out to determine which phosphoesters of these amino acids could be detected in preparations of the 60 kDa protein. The method used was essentially that of Cooper *et al.* (1983) in which protein hydrolysates are analysed by 2-dimensional electrophoresis on thin-layer chromatography plates except that proteins were hydrolysed directly from PVDF blots of IEF gels rather than from solution. A description of the methods used for both blotting of the IEF gels and phosphoamino acid analysis is given in Chapter 2.

Fig. 3.4 panels A and B show areas of a ninhydrin stained thin-layer chromatography plate used in the analysis of the phosphate labelled proteins from bloodstream bands 1 and 2 respectively while panels C and D show the corresponding autoradiographs. The large arrows indicate the direction of electrophoresis in buffers of pH 1.9 and pH 3.5. The identities of the three ninhydrin-stained internal standards are marked on panels A and B while the arrows on panels C and D mark the positions of phosphate labelled spots that match with the ninhydrin stained spots on the plates. The nature of the other radiolabelled sample components detected in the autoradiographs is not known but it is normal to detect phosphopeptides as well as free phosphoamino acids. The marked phosphate labelled spots in both C and D have the same positions as the phosphoserine and phosphothreonine standards of panels A and B but in neither case is there phosphate labelling at a position corresponding to that of the phosphotyrosine standard. As discussed below, the phosphoamino acid analysis performed on the IEF bands is incomplete but phosphoserine and phosphothreonine were detected in each of four separate analyses of bloodstream bands 1 and 2. IEF bands 3 and 6 were not sufficiently well resolved from blots of IEF gels to be separately analysed, while single analyses on bands 4, 5, and 7 showed labelling of serine and threonine. The results of three analyses of band 8 showed labelling only on serine.

### 3.4 Discussion

The profile of labelled bands obtained using the IEF assay to detect autophosphorylating activities and the SDS-PAGE analysis of the labelled bands confirmed the earlier results. The phosphoamino acid analysis confirmed that phosphate had been covalently transferred to protein and, for the bands where thorough analysis was possible, phosphate transfer was either to serine and threonine residues or only to serine. The combined results of the earlier analyses and of the work reported here are published in Hide *et al.* (1994).

The original finding that the predominant phosphate labelled protein in each labelled IEF gel band migrated at 60 kDa in SDS-PAGE was surprising but was confirmed by the work described here. The common size of the labelled proteins suggests that they are related and while some proteins are substrates for many protein kinases the separation of proteins in the *T. brucei* extracts by isoelectric focusing would

be expected to separate most protein kinases from their substrates. Two possible trivial explanations of the results were addressed in controls experiments performed previously and not repeated here. These were: 1) that the 60 kDa band protein was a contaminant of the agarose used to form the gels; 2) that the 60 kDa protein was the VSG or another major protein coincidentally co-migrating with multiple protein kinase activities. The first of these possibilities was ruled out by experiments in which the assay was performed in polyacrylamide gels. The results obtained using these gels were very similar results to those obtained using agarose gels. The 60 kDa protein was shown not to be VSG by its lack of reactivity with anti-VSG antibodies in a Western blot of an SDS-PAGE gel containing the 60 kDa protein. Staining for total protein showed no major protein band at the position of the 60 kDa labelled band indicating that the 60 kDa protein is not highly abundant in *T. brucei*. Thus while the possibility that the 60 kDa labelled proteins co-migrate with unassociated protein kinase activities cannot be absolutely ruled out, by far the most likely explanation is that the labelling of the 60 kDa protein is due to an autophosphorylating protein kinase activity. The simplest case for this type of protein kinase activity in which the substrate and protein kinase are physically associated is that of autophosphorylating protein kinases where the protein kinase catalytic subunit phosphorylates its own amino acids in an intramolecular reaction. This type of activity is very common among protein kinases, but trans-autophosphorylation is also consistent with the results since identical molecules of a protein kinase will have the same isoelectric point and focus to the same region of the gel. It is also possible that the 60 kDa protein is not itself a protein kinase catalytic subunit but is a substrate that co-migrates with one or more types of catalytic subunit due to a physical association with the catalytic subunit. Most protein kinase substrates would only transiently associate with a protein kinase but a protein kinase regulatory subunit, for instance, could be much more strongly associated.

There were obviously differences between the labelled proteins (or between oligomeric complexes that included the labelled protein) detected in the IEF gels since there were multiple bands of different isoelectric point. If the 60 kDa polypeptide is part of an oligomeric complex, the 60 kDa polypeptide itself could be identical in each of the IEF gel bands detected, since the charge differences could be explained by the association of the 60 kDa polypeptide into different homo- or hetero-oligomers. The

denaturation of the proteins before SDS-PAGE (and reduction of any disulphide bonds) would separate the components of the oligomers allowing common subunits to run similarly in the SDS-PAGE gels.

If the differences in isoelectric point between the IEF bands are due to differences between the 60 kDa polypeptides, the differences must affect the isoelectric points much more than the mobilities in SDS-PAGE, unless protein modification during the IEF assay is responsible for the similar mobilities of the labelled proteins in SDS-PAGE (see below). There was no clear evidence for more than one band near 60 kDa when the SDS-PAGE analysis of the IEF bands was undertaken (the lower mobility SDS-PAGE band from seen in gel lane 3 in Fig. 3.2 panel B (from IEF band 3) was not reproducibly detected) but it is possible that small mobility differences could have been missed. The procedure for transferring protein contained in the agarose of the IEF gels to SDS-PAGE can lead to adjacent wells being loaded with slightly different amounts of sample buffer which is not ideal when comparing bands of very similar or identical mobility. Also, long exposures of autoradiographs were often necessary for detection of the weaker bands so that the stronger bands were overexposed, possibly obscuring tightly spaced multimers. The simplest pre-translational mechanism for producing slightly different products is the expression of multiple closely related genes. There are many examples known of this and a small number of closely related gene products, if capable of forming oligomers, could produce a larger number of oligomeric isoenzymes. Alternate splicing is another pre-translational mechanism that could produce slightly different forms of a protein. Phosphorylation is one of the commonest post-translational modifications of proteins in higher eukaryotes. The addition of a phosphate to a protein makes the protein more acidic and shifts its position in isoelectric focusing towards the anode. The charge shift is typically about 0.3 pH units per phosphate in the pH 6-7 region of the gel (Garrison and Wagner, 1982). It is quite possible that some of the different IEF gel bands reflect different levels of phosphorylation of the 60 kDa protein prior to IEF. Phosphorylation differences can lead to detectable differences in protein mobility in SDS-PAGE and the addition of phosphates at particular sites can lead to both increases and decreases in mobility, the latter being more common. However, protein kinase activity during the labelling step in the IEF assay could have equalised the extent

of phosphorylation of the 60 kDa protein. This is one possible explanation of the similar mobilities of the IEF bands when cut out and run in SDS-PAGE.

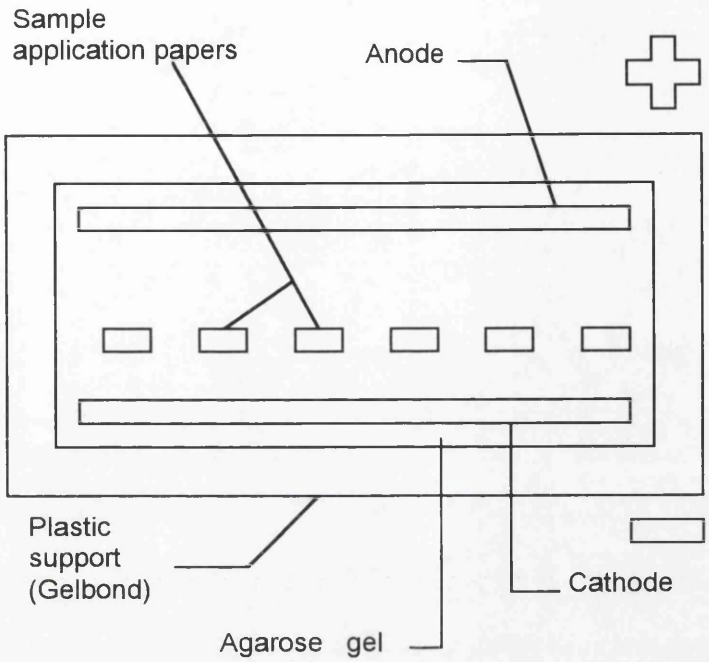
To summarise, based on their common 60 kDa molecular weights in SDS-PAGE the predominant phosphate labelled proteins in each of the labelled IEF gel bands appear to be related, while there is no reason to believe that a 60 kDa protein that might act as substrate for multiple protein kinases is distributed throughout the IEF gels. It therefore appears that the 60 kDa proteins are either protein kinase catalytic subunits or are quite strongly associated with protein kinase catalytic subunits and that they become labelled in an autophosphorylation reaction. The differences in isoelectric point between the IEF bands containing the 60 kDa proteins are unexplained, but a number of different mechanisms that might explain the observed results can be proposed. Because the results are most consistent with the various IEF gel bands representing different isoenzymes of a protein kinase that autophosphorylates on a 60 kDa polypeptide, the activities of different isoelectric point detected in the IEF assay are referred to below collectively as those of the 60 kDa autophosphorylating protein kinase.

The earlier work was also extended by phosphoamino acid analysis of the labelled IEF bands from bloodstream *T. brucei*. Labelling of serine and threonine was detected for IEF bands 1, 2, 4, 5 and 7 while band 8 was apparently labelled only on serine, but no phosphotyrosine was detected. It therefore appears that the protein kinase activity detected is that of a serine/threonine protein kinase. The phosphoamino acid analysis of the radiolabelled IEF bands was however found to be problematic and, despite a number of attempts, the data obtained was limited and no reliable estimate of the relative amounts of phosphoserine and phosphothreonine in each IEF gel band was obtained. A serious problem was the weakness of labelling of some of the IEF bands. The protein hydrolysis conditions are a compromise between incomplete hydrolysis of protein and excessive hydrolysis of phosphoamino acids so it is difficult to obtain results if very little radiolabelled phosphate is incorporated in the sample proteins. A possible solution to this problem would have been to use a much higher ratio of radiolabelled to non-radiolabelled ATP in the IEF assay. A further difficulty was the unreliability of the 2-D electrophoresis procedure. Many experiments were uninterpretable because the resolution of the phosphoamino acid spots was not good enough to allow unambiguous identification. A major problem was in wetting the plates evenly and then preventing uneven drying out of

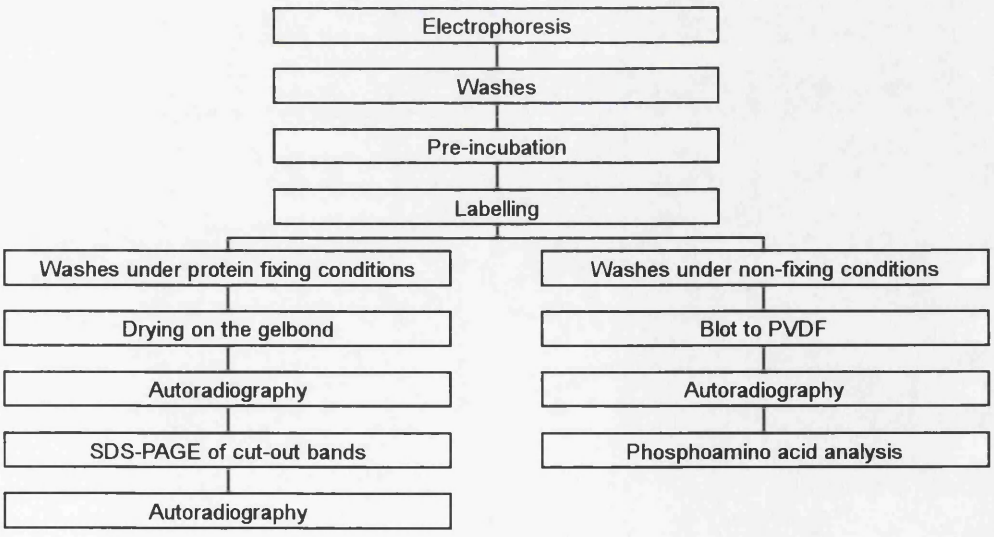
the plates during electrophoresis. Adjustments to the voltage applied and the technique for wetting the plates did not improve the results greatly. Purpose-built equipment is available that addresses the problem of unequal fluid distribution across the electrophoresis plate and use of this equipment would probably have given better results.

The IEF assay could have been used for further characterisation of the protein kinase activity detected but it was decided to purify the protein kinase before further characterisation was undertaken. The protein kinase activity was of interest because of the great differences observed between the activities in the bloodstream and procyclic extracts. Although any information that might have been gained about the properties of the 60 kDa autophosphorylating protein kinase in further studies using the IEF assay would have been potentially useful, there were problems with reproducibility both between batches of extract and between individual gels loaded with the same extract. This meant that not every band of activity could be reliably detected and resolution of the bands was also limited. This may have been more to do with diffusion during the various washes between electrophoresis and gel fixation or blotting than with the electrophoresis itself. For these reasons each experiment had to be repeated a number of times to give a reliable result and it seemed more promising to proceed directly to partial purification of the protein kinase before undertaking further characterisation.

**Fig. 3.1** Isoelectric focusing gel *in situ* protein kinase assay. Panel A shows schematically the arrangement of the gels used in the isoelectric focusing gel *in situ* protein kinase assay. The method is described in detail in Chapter 2. Panel B summarises the post-electrophoresis steps in the protein kinase assay.


Sample
application papers
Anode
Plastic
support
(Gelbond)
Cathode
Agarose gel

A


Electrophoresis
Washes
Pre-incubation
Labelling
Washes under protein fixing conditions
Drying on the gelbond
Autoradiography
SDS-PAGE of cut-out bands
Autoradiography
Washes under non-fixing conditions
Blot to PVDF
Autoradiography
Phosphoamino acid analysis

B

**Fig. 3.1**

**Fig. 3.2** Analysis of autophosphorylating protein kinase activities from bloodstream and procyclic *T. brucei* using the isoelectric focusing gel *in situ* protein kinase assay. Panel A shows the autoradiograph of a typical IEF gel used for *in situ* assay of protein kinase activity in the absence of added substrate. Procyclic (lane P) and bloodstream (lane BS) samples were loaded at the origins (0) and electrophoresed before assay. The positions of the anode (+) and cathode (-) are marked. Isoelectric points (pI) were determined by running pI markers alongside the samples. The phosphate-labelled bands are numbered 1-2 for the procyclic samples and 1-8 for the bloodstream. Panel B shows the autoradiograph of an SDS-PAGE gel loaded with bands cut out from a dried down IEF gel containing autophosphorylating protein kinase activity from bloodstream *T. brucei.* Lanes 1-6 contain bands 1, 2, 3, 4, 7, and 8 respectively. The sizes and positions of protein molecular weight markers are indicated to the left of the figure.

**Fig. 3.3** Autoradiograph of an IEF gel used to compare protein kinase activity in a bloodstream total extract and the 100 000 g supernatant of a total extract. Lanes 1 and 2 are identical lanes loaded with a total extract, lanes 3 and 4 are identical lanes loaded with the supernatant of a total extract.

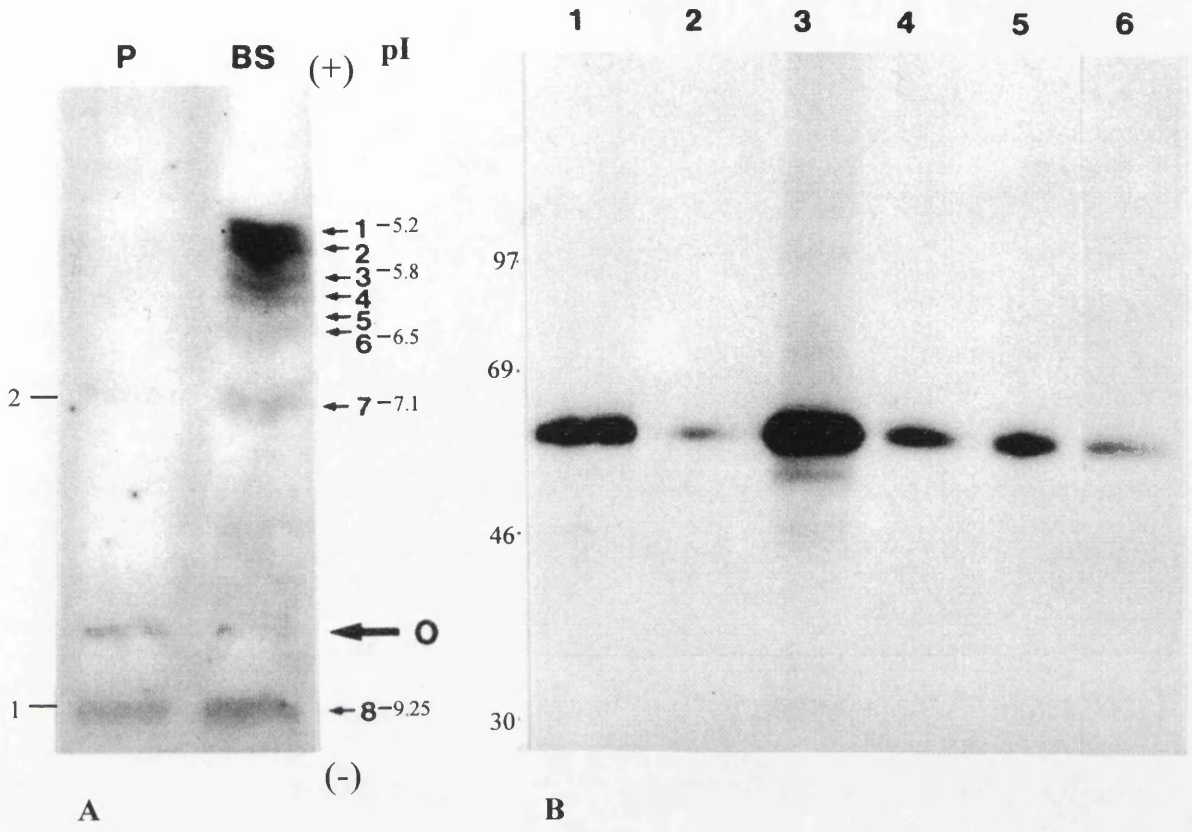



Fig. 3.2

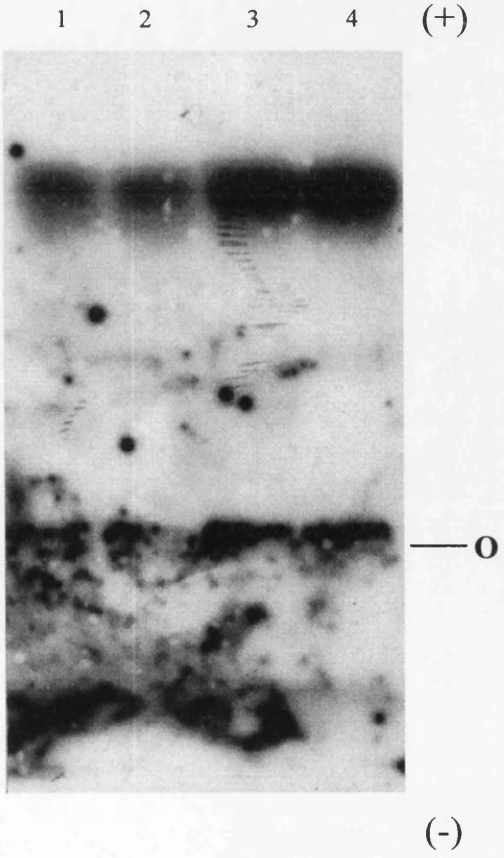



Fig. 3.3

**Fig. 3.4** Phosphoamino acid analysis of the 60 kDa autophosphorylating protein kinase. Panels A and B show sections of a ninhydrin-stained thin-layer chromatography plate loaded with hydrolysates of bloodstream IEF bands 1 and 2 respectively (see Fig. 3.2) mixed with phosphoserine, phosphothreonine, and phosphotyrosine standards. Panels C and D show the autoradiographs corresponding to panels A and B respectively. Labelled spots corresponding to the phosphoserine and phosphothreonine standards are indicated by arrows. The directions of electrophoresis in buffers 1.9 and 3.5 are also indicated.

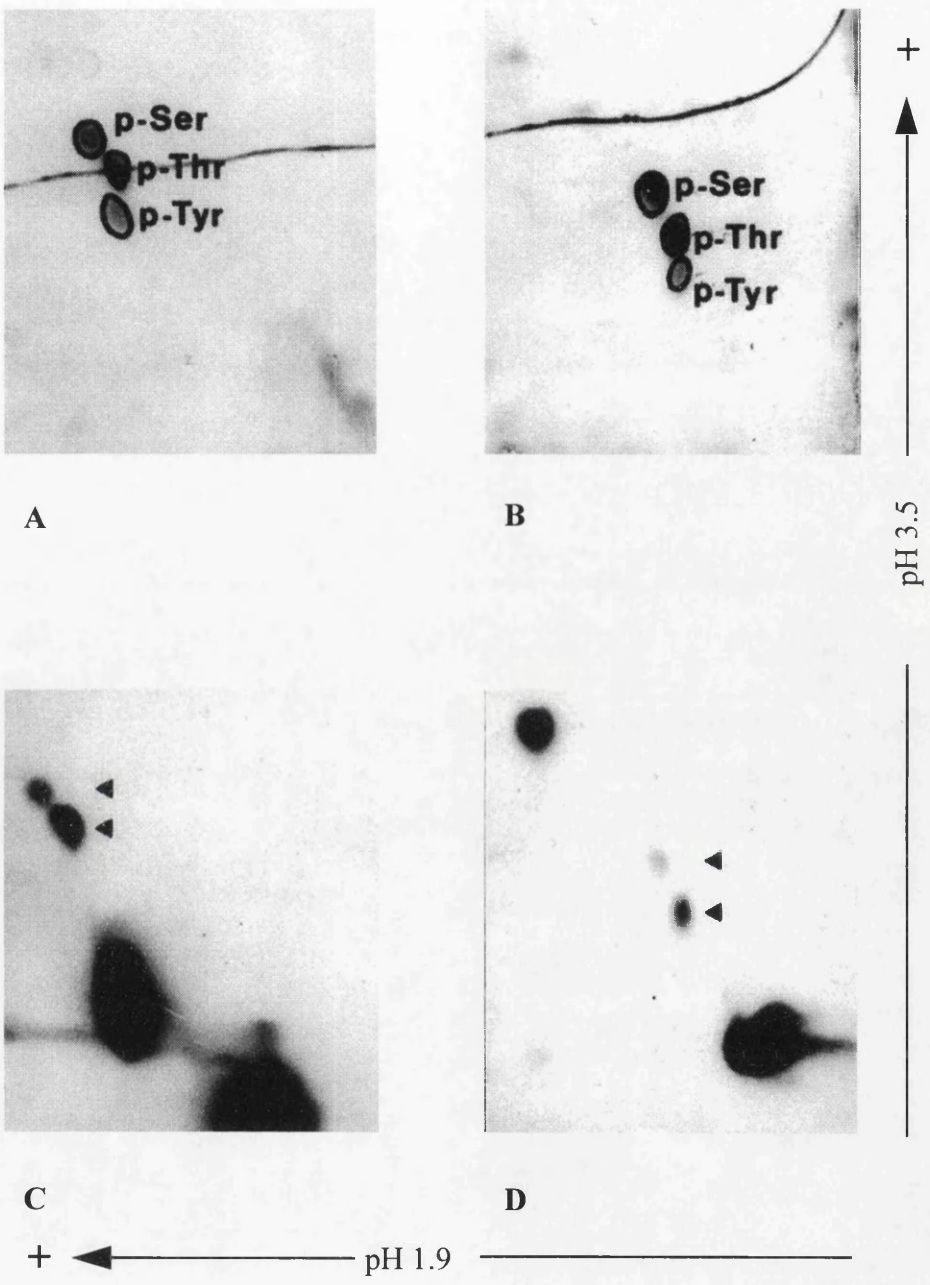



**Fig. 3.4**

# Chapter 4

# Purification of the 60 kDa protein kinase

## 4.1 Introduction

### 4.1.1 Aims

As described in the previous chapter the 60 kDa autophosphorylating protein kinase identified using the IEF assay showed significant stage-specific differences in its expression. Up to eight bands of activity were detected from bloodstream forms while only two bands were detected from procyclic forms. Therefore, the study of the protein kinase was interesting from the point of view of defining biochemical differences between slender bloodstream and procyclic forms (no experiments had been performed to detect the protein kinase in stumpy bloodstream forms). Protein phosphorylation in *T. brucei* has only been studied to a limited extent (see Chapter 1) while in other eukaryotes major developmental changes such as those that *T. brucei* undergoes in its complex life cycle are co-ordinated by regulatory networks that include protein kinases. Therefore, in order to better understand differentiation in *T. brucei* there is a need for more research into protein kinases that are expressed at different levels between life cycle stages.

One major goal in characterising the 60 kDa protein kinase was to clone and sequence the gene encoding it; the approach chosen to achieve this was to purify the protein kinase in small quantities in order to determine the amino acid sequence by microsequencing and to use as an antigen for raising specific antibodies. With a small amount of known sequence, specific (although degenerate) oligonucleotides can be designed to allow gene cloning through PCR or library screening. If antisera specific to the protein can be raised, expression library screening can be used to isolate clones from which the gene sequence can be determined, in addition, a specific antiserum could be used in the purification of greater quantities of pure protein for microsequencing. The development of methods for at least partial purification of the 60 kDa protein kinase would also simplify further biochemical characterisation. General considerations in protein purification are discussed in Section 4.1.2.1 while various chromatography techniques for protein purification are described in Sections 4.1.2.2-4.1.2.5. Affinity purification techniques specific to protein kinases are dealt with in Section 4.1.2.6.

### 4.1.2 Approaches to protein purification

#### 4.1.2.1 General considerations

Proteins are large molecules with complex three-dimensional structures, and a mixture of charged, polar and hydrophobic groups exposed at the surface. A wide range of techniques are therefore applicable to their purification, but protein purification also presents particular difficulties. For many purposes it is necessary that the protein being purified retains its folded structure throughout the purification and so gentle separation conditions are often necessary. Protein purification also requires specific means of detecting the protein of interest from a complex mixture of proteins. Because of the large number of proteins in a cell it is usually necessary to use several sequential purification steps in purifying an individual protein to homogeneity. In many cases the overall procedure is a combination of "classical" purification methods such as ion-exchange and size-exclusion chromatography and one or more affinity steps based on the reversible binding of the target protein to an immobilised ligand, such as a substrate, for which it has a specific binding site.

Liquid chromatography techniques are a part of most protein purification procedures and a range of chromatographic media designed for protein separations is available. In most cases the separation property is determined by a specific chemical group (e.g. a charged group for an ion-exchange medium) which is attached to an insoluble matrix. The support matrix for protein chromatography will generally be hydrophilic so that it can be wetted by an aqueous mobile phase but should not be charged or strongly polar as this will result in non-specific protein interactions with the matrix. It must also have mechanical strength and flow properties suitable for the conditions under which it will be used. The best resolution is obtained when the surface area is maximal and since proteins are large molecules the pore size of the support must be large enough that proteins can diffuse freely through the matrix (so that it does not become blocked or show undesirable size-exclusion properties). The best results will be obtained with column chromatography using a pump capable of operating at high pressure and a gradient forming system able to mix a mobile phase accurately from several solvents (HPLC). UV spectrophotometry at 280 nm is commonly used to give an estimate of overall protein concentration in the eluate.

#### 4.1.2.2 Ion-exchange chromatography

Ion-exchange chromatography depends on ionic interactions between charged groups on the chromatographic matrix and charged sample components. In a typical purification an ion-exchange matrix is equilibrated in a buffer that includes counter-ions with respect to the immobilised charged groups of the matrix. The sample is then applied. Molecules that are uncharged or that have the same charge as the ion-exchange matrix fail to bind while those of opposite charge may displace buffer counter-ions and bind reversibly to the chromatography matrix. After washing to remove unbound material, the conditions are changed to elute bound sample molecules. Two types of changes in the buffer are commonly used for elution in the case of protein mixtures: an increase in ionic strength leads to increased competition for binding sites by buffer ions, while a shift in buffer pH towards the isoelectric point of a protein weakens its interaction with an ion-exchange matrix. Once the target protein has been eluted under conditions that maximise separation, more stringent elution conditions may be employed to remove any strongly-bound material, before the matrix is re-equilibrated.

The strength of binding of a protein to an ion-exchanger depends on the number of charged groups interacting with the ion-exchanger and the charge density at the points of interaction. Very high resolution of proteins may be obtained because of their great diversity in charge properties. Proteins are amphoteric and both positively charged (anion-exchange) and negatively charged (cation-exchange) matrices may be used provided that the target protein is stable across the required pH range. Ion-exchangers may be described as strong or weak depending on whether they remain completely ionised over a very wide or a somewhat narrower pH range, respectively.

In developing an ion-exchange step for protein separation the pH chosen for protein adsorption will normally be about 1 unit above (anion-exchange) or below (cation-exchange) the isoelectric point of the target protein. Most proteins begin to elute from an ion-exchanger at 0.5 pH units from their isoelectric point at an ionic strength of 0.1 M NaCl (Lampson and Tytell, 1965). However, it is usually necessary to refine binding and elution conditions empirically. A protein's isoelectric point is a measurement of net charge while protein binding to an ion-exchange material is also dependent on the distribution of charge, so that some proteins will remain bound at their isoelectric points. Salt gradients are commonly used for protein elution and, while buffer pH changes are

also quite commonly used, these are normally step changes since pH gradients are less reproducible.

The advantages of ion-exchange chromatography in protein separations are its wide applicability, the high resolution achievable, and the generally high capacity of ion-exchange media. Ion-exchange separations can usually be performed using "gentle" buffer conditions and recovery of native proteins retaining activity is generally good. It is often possible to elute the desired protein in a relatively small volume. For these reasons it is one of the most commonly used protein separation techniques. Ion-exchange chromatography may be used at any stage of a protein purification and it is not unusual for a protein purification procedure to include more than one ion-exchange step. However, because of the high capacity of ion-exchange matrices it is very frequently the first chromatography step used.

#### 4.1.2.3 Size-exclusion chromatography

Size-exclusion chromatography (also known as gel filtration) separates molecules according to differences in their molecular size. The media used for size-exclusion chromatography contain pores filled by the liquid phase into which sample components can enter. A sample is applied to the top of a column in a small volume and is eluted by a flow of eluant of constant composition. As the sample zone moves down the column, the smallest sample molecules move freely into the pores of the matrix and are retarded in their movement down the column, larger molecules may be partially or completely excluded and travel faster down the column mainly or entirely in the spaces between the matrix particles. Sample molecules are therefore separated on the basis of size with the largest molecules eluting first although molecules of unusual shape may migrate anomalously.

Unlike ion-exchange chromatography which has wide applicability to protein separations, size-exclusion chromatography is limited to separating proteins that differ substantially in size. In a size-exclusion separation important considerations are the pore size of the chromatographic matrix and the column length as well as the sample volume and viscosity. The matrix used should be one that gives best resolution over the size range important for the separation being performed; if this is not known, a matrix with a wide fractionation range may be used initially. Resolution is proportional to the square

root of the column length and so columns for size-exclusion chromatography tend to be relatively long. Sample volume is important since sample resolution depends on the ratio of sample volume to column volume. A sample volume 1-5% of the column volume is normally used for protein separations. Sample concentration does not affect size-exclusion chromatography provided that the viscosity of the sample does not differ from that of the eluant by more than a factor of two but this is a relatively stringent requirement given that sample volume must also be low. The separation achieved in size-exclusion chromatography does not fundamentally depend on eluant composition, although the buffer should have a high enough ionic strength to prevent ionic interactions between sample components and the chromatographic matrix. Because an eluant can be chosen to maximise protein stability, recovery of activity is usually good. It is also often possible to load samples from a previous step without buffer exchange. If the target protein interacts with other proteins in the sample, a buffer may be chosen to minimise or maximise the extent of interaction.

Size-exclusion chromatography is an attractive technique for protein separation provided that the target protein differs sufficiently in size from contaminants to make the technique applicable. The equipment needed is relatively simple, and the development of size-exclusion separations is usually uncomplicated. Recovery of native protein is usually good. The main disadvantages of size-exclusion chromatography, apart from its limited applicability, are the relatively low protein capacity of the technique (for this reason it is usually used in the later stages of protein purification) and its inevitable tendency to dilute the sample. Size-exclusion chromatography is often used for purposes other than protein separation. Particularly common applications are in determining protein molecular weights under non-denaturing conditions and for the de-salting of protein solutions, for which it is much quicker than dialysis.

#### 4.1.2.4 Reversed-phase and Hydrophobic interaction chromatography

Reversed-phase chromatography (RPC) and hydrophobic interaction chromatography (HIC) are related techniques based on the partitioning of sample components between a polar mobile phase and a hydrophobic stationary phase. While the majority of hydrophobic groups will be buried in the core of a folded protein, the side chains of alanine, methionine, tryptophan and phenylalanine do occur at the surface (Hofstee,

1975). The "hydrophobic interaction" between a protein leaving solution to adsorb to a hydrophobic matrix is due to the entropy increase that occurs as water molecules that were previously ordered through contact with the protein become less ordered, although van der Waal's forces can also play a part.

RPC describes chromatography using strongly hydrophobic groups, usually aliphatic chains, attached to the support matrix at high density. In RPC there would normally be a significant organic component to the mobile phase and, with protein samples, denaturation would not be unusual. HIC refers to chromatography with a less hydrophobic stationary phase substituted with shorter chain alkyl groups or with phenyl groups, at a lower density. The mobile phase might include an organic component but could be entirely aqueous. Protein denaturation is less likely to be a problem than for RPC. A common way of performing HIC is to apply the protein sample in the presence of "salting out" salts which promote protein adsorption by trapping water molecules released on protein adsorption and then to elute with a decreasing salt gradient. If a drop in ionic strength is insufficient to elute the target protein an organic component might be added to the mobile phase. These promote elution by reducing the polarity of the solution or by competing more strongly than the protein for sites on the matrix. A reduction in temperature will also decrease the strength of interaction of a protein with a hydrophobic matrix.

Proteins vary widely in the hydrophobic groups exposed at their surfaces and HIC is therefore applicable to many protein separations. The protein capacity of HIC matrices is generally good and HIC, like ion-exchange, can be used in the early stages of a protein purification. An HIC step using salt-promoted adsorption can follow a salting-out step with little adjustment to the sample buffer. The main disadvantages of HIC are that the conditions used are more likely to denature proteins than those of most other chromatography techniques and the development of HIC separations often requires much experimentation.

#### 4.1.2.5 Affinity chromatography

Many proteins interact very specifically with other biomolecules due to ligand binding sites on the protein which provide a surface complementary to the ligand. For instance, enzymes have specific sites which bind substrates. In affinity chromatography a ligand is

immobilised on a support to make the affinity matrix. A protein mixture containing the target protein is passed over the matrix to allow adsorption and the matrix is washed. Elution is achieved either with a selective eluant (such as the free ligand) that competes for the binding site on the protein or by a change in conditions to weaken the binding between the target protein and ligand in a less specific way. For instance, if the binding between a protein and a ligand is mainly ionic then increasing the ionic strength to a high enough level will cause desorption.

Ligands that may be suitable for affinity chromatography include the cellular substrates and regulators of the target protein, natural and synthetic analogues of these and antibodies. A number of factors affect the choice of ligand for a particular application. Availability of the ligand may be a significant factor as a molecule that would otherwise be ideal for affinity chromatography might be very difficult or expensive to obtain in sufficient quantities. The affinity of the interaction between target protein and ligand must be high enough that the protein is adsorbed to a high degree but not so high that harsh conditions are needed for elution. In general, the affinity in solution of the ligand for the target protein should be in the range $10^{-4}$-$10^{-8}$ M. A high degree of specificity for the target protein is desirable but ligands such as nucleotides have often been used for affinity chromatography despite the fact that they bind many proteins. In such cases affinity chromatography is best used after previous purification steps have removed most competing proteins.

A very important consideration in affinity chromatography is the coupling of the ligand to its support to make the affinity matrix. A range of support materials is available for immobilising ligands with different reactive groups but it is also essential that the linkage used is compatible with protein binding. If the ligand is immobilised through a group at, or close to, a part of the ligand that interacts with the target protein then protein binding may be sterically hindered. Therefore, it may be necessary to test matrices with the ligand bound through different groups. Spacer molecules between the support and the ligand may also be needed where the ligand is small and protein binding would otherwise be sterically inhibited by the support matrix. A possible problem associated with the use of spacer molecules is increased non-specific protein binding.

Affinity chromatography potentially gives the highest levels of protein separation of any of the chromatographic techniques. Provided that the affinity is not too strong,

retention of biological activity is usually good and ligand binding stabilises most proteins, although selective eluants can be difficult to remove. The major limitation is that it is difficult to develop affinity chromatography for target proteins that have not been well characterised. Even if a potentially suitable ligand is available, it may not be easy to find a way of immobilising it to give a suitable matrix for purifying the target protein.

#### 4.1.2.6 Purification approaches specific for protein kinases

General considerations in affinity chromatography have been discussed above. In this section affinity techniques applicable to protein kinases are considered in more detail. Protein kinase purification was reviewed in Methods in Enzymology **200**: 159-178 and by Wang and Roach (1993). A number of features are common to many protein kinase purifications. The abundance of protein kinases varies widely (see Ferrari and Thomas, 1991) but they are frequently low-abundance proteins and, if this is the case, it is important to maximise the efficiency of the purification techniques used and one or more affinity steps may be virtually essential. Protein kinases are often regulated by phosphorylation and dephosphorylation which can affect both their activity in assays and their properties in purification procedures. For this reason it may be necessary to homogenise the starting material rapidly in a buffer containing inhibitors of protein kinase and protein phosphatase activity and to keep the protein kinase in a buffer containing phosphatase inhibitors at all times. A lysis buffer would normally contain EDTA which will prevent protein kinase activity by chelating $Mg^{2+}$. Sodium fluoride is commonly used as a protein phosphatase inhibitor and is cheap but, if it is necessary to avoid raising the ionic strength or if the protein kinase is targeted by a protein phosphatase not inhibited by fluoride, it may be necessary to include other inhibitors such as okadaic acid. Vanadate can be used to inhibit protein tyrosine phosphatases.

Possible affinity ligands for the purification of protein kinases are the nucleotide and peptide substrates and effector molecules such as activators and inhibitors. Immobilised adenosine nucleotides have been used in the purification of protein kinases as all use ATP as a substrate (some will also use GTP). ATP affinity chromatography matrices have been described as being of four types depending on whether the coupling is through the N-6 of adenine, the C-8 of adenine, the terminal phosphate or a ribose hydroxyl group (Jenö and Thomas, 1991). A ribose-coupled ATP agarose was used to

give a 20-fold purification of a mitogen-activated S6 kinase (Jenö *et al.*, 1989) although the same enzyme failed to bind ATP agaroses coupled through the N-6 and C-8 positions of adenine under the same conditions. Ribose-coupled ATP agarose was also used to purify CDK2 from insect cells over-expressing the protein (Rosenblatt *et al.*, 1993) while Seger *et al.* (1992) used a C-8 coupled agarose in the purification of a MAP kinase kinase. However, despite the fact that all protein kinases bind ATP, the number of examples of the successful use of ATP affinity matrices for protein kinase purification is limited. The use of ATP affinity matrices for this purpose was recently re-examined by Haystead *et al.* (1993) who pointed out that the structure published for PKA co-crystallised with $Mg^{2+}$ and ATP shows the adenine ring deeply buried in a hydrophobic pocket with the phosphates facing outwards. The ribose hydroxyls at the 2- and 3-positions hydrogen bond with the enzyme and all three phosphates are also involved in interactions with the enzyme. The ATP-binding sites in the other protein kinases analysed by X-ray crystallography are similar, although there are some differences (see review by Taylor and Radzio-Andzelm, 1994). On the basis of these considerations steric blocking of the interaction between protein kinase ATP-binding sites and immobilised ATP would be expected for N-6, C-8 and ribose-coupled ATP affinity matrices (purifications achieved in the past using these matrices might have been due to cation-exchange) but immobilisation of ATP through the terminal (γ) phosphate would be expected to be the most suitable linkage. However, Haystead *et al.* (1993) were unable to find any published work describing the use of such affinity matrices for protein kinase purification (or any source for a γ-phosphate-linked ATP matrix) and decided to develop a γ-phosphate-linked ATP Sepharose. The first matrix synthesised was a Sepharose with a non-hydrolysable ATP analogue (adenosine-5′-(γ-4-aminophenyl) triphosphate) bound directly to reactive groups on the support through the phenylamino group without the use of a spacer. Although the ATP analogue had been previously shown to inhibit four different protein kinases in solution and several ATP-binding proteins were shown to bind to the matrix, none of the protein kinases tested were found to bind. However, when the same ATP analogue was linked to Sepharose in the same way but with an additional hexyl spacer, a matrix was obtained that was used to give a 31-fold purification of a mammalian MAPKK. ATP affinity matrices of this type have since been

used in the purification of a rat liver AMP-activated protein kinase (Davies *et al.*, 1994) and a human IL-1 activated protein kinase (Kracht *et al.*, 1994).

A number of triazinyl dyes bind enzymes at sites naturally occupied by purine nucleotides. Cibacron blue F3GA has been shown by crystallography to bind alcohol dehydrogenase at the sites normally occupied by the adenine and ribose rings despite the dye's general lack of structural similarity to NAD (Biellmann *et al.*, 1979). Cibacron blue F3GA and other dyes of this group that act as ADP-ribose analogues have been used in the purification of protein kinases, although the structure of the nucleotide binding site in protein kinases is different from that in NAD binding proteins (see Bossemeyer, 1994), and it is not known whether the interaction between these dyes and protein kinases is at the protein kinase ATP-binding site. Protein binding to triazinyl dyes can involve both ionic and hydrophobic interactions, so unless a bound protein kinase elutes with a low concentration of a specific eluant such as ATP a range of elution conditions may have to be tested.

Protein and peptide substrates have been used as affinity ligands for the purification of a number of protein kinases. Specific protein substrates are likely to give highest binding affinity but are often difficult to obtain at high purity in sufficient quantities. The large size of proteins, and the variety of ionic, polar, and hydrophobic regions present at their surfaces can cause non-specific interactions with contaminating proteins as well as there being a higher probability of recognition sites for multiple protein kinases on a whole protein than on a peptide. An example of the use of a protein substrate affinity matrix was the 10-fold purification of glycogen synthase kinase-3 on a glycogen synthase-agarose (Hemmings *et al.*, 1981). Matrices with immobilised model substrates, such as protamine, are commercially available and so are easy to test as affinity matrices. Protamine-agarose, for example, has been used in the purification of protein kinase C (Wooten *et al.*, 1987). The use of peptides based on phosphorylated sites in protein kinase substrates for protein kinase affinity purification was reviewed by Woodgett (1991). Where the sequences at sites phosphorylated by the target protein kinase are known, similar peptides can be tested as substrates for the protein kinase in solution before being tested as affinity ligands. Peptides corresponding to protein kinase recognition motifs but with the phosphorylated residues replaced by residues that cannot accept phosphate groups have been used as protein kinase inhibitors and such motifs

(pseudosubstrate domains) occur naturally in a number of protein kinases. A peptide corresponding to the pseudosubstrate domain of the naturally occurring peptide inhibitor of PKA (PKI) was effective as an affinity ligand for purification of this protein kinase (Olsen and Uhler, 1989). Compared to whole proteins, peptides are simple to obtain in large amounts and at high purity; they are small and are therefore less likely to bind non-specifically to proteins other than the target protein and can also be coupled to supports at higher density. For both proteins and peptides the coupling group used is important, and a spacer may also be needed although useful affinity matrices can be obtained when only a proportion of the ligands are in the correct orientation.

As well as matrices with immobilised proteins and peptides, Sepharoses with immobilised threonine have been used in the purification of protein kinase C and ribosomal S6 kinase (Lane and Thomas, 1991; Koide *et al.*, 1992). While threonine is one of the substrate amino acids of these protein kinases, it is not clear that threonine Sepharose acts as a true affinity matrix. Another type of matrix that may interact with sites on some protein kinases that are involved in peptide substrate recognition is phosphocellulose. This is an ion-exchange material that has been used for purifying a number of different types of protein but it may act as an affinity matrix for protein kinases such as casein kinase I (Xu *et al.*, 1995) and glycogen synthase kinase. These protein kinases phosphorylate residues in peptide motifs that include phosphoserine, and might have sites that can specifically bind to the phosphorylated cellulose.

Various molecules that bind to protein kinases as regulators of enzyme activity can be used as ligands in affinity chromatography. Protein kinase regulatory subunits have been used in a number of cases. For example, calmodulin-Sepharose has been used for the purification of several calcium/calmodulin-dependent protein kinases (e.g. Nairn *et al.*, 1985) and a group of proteins that interact with cyclin-dependent protein kinases have proved extremely useful as affinity ligands for their purification (Hindley *et al.*, 1987; Azzi *et al.*, 1994). Immobilised EGF has also been used in purifying its receptor (Cohen *et al.*, 1982) but the usefulness of peptide growth factors as affinity ligands for the purification of receptor protein kinases is limited by both cost and the fact that they may bind so strongly to their receptors that elution without loss of activity is impossible. Non-peptide effectors of protein kinase activity have also been used as affinity ligands. An example is the use of immobilised forms of cAMP and cGMP for the purification of

the cyclic nucleotide dependent protein kinases. cAMP and cGMP bind to these protein kinases with high affinity and a very substantial level of purification can be achieved (e.g. Lincoln, 1983). Phosphatidylserine is a natural effector of protein kinase C that has been used as an affinity ligand for its purification (Wise *et al.*, 1982). Heparin is an inhibitor of casein kinase II *in vitro* and immobilised heparin has been used for its affinity purification (Takio *et al.*, 1987). Non-physiological protein kinase inhibitors can also be useful affinity ligands for protein kinase purification. The isoquinolinesulfonamide, H-9, inhibits PKA, PKG, and PKC. Inagaki *et al.* (1985) used an H-9 Sepharose in purifying PKC from rabbit brain to near homogeneity and were also able to elute PKC, PKA and PKG separately from the same column. Chijiwa *et al.* (1989) demonstrated binding of casein kinase I to Sepharoses to which two further isoquinolinesulfonamide inhibitors (CKI-7 and CKI-8) had been attached and used the latter matrix in the purification of this protein kinase.

A significant proportion of protein kinases are themselves phosphoproteins and this property may be exploited in their purification. Phosphotyrosine antisera have been used in the purification of protein tyrosine kinases (e.g. Wedegaertner and Gill, 1989) although, the best known use of such antibodies, is probably in screening expression libraries to isolate genes for protein tyrosine kinases. Antibodies to phosphoserine and phosphothreonine are also available but are less likely to be useful for protein kinase purification due to the presence of these phosphoamino acids on many proteins. Phosphoproteins can also show selective adsorption to iron (III) metal chelate chromatography matrices (Muszynska *et al.*, 1986). This technique appears to have been little used for protein kinase purification although it was used by Songyang *et al.* (1994) to select phosphopeptides from a peptide library.

#### 4.1.2.7 Summary

The classical chromatography techniques described above are among the most commonly used techniques for protein purification that have general applicability. However, there are many other protein purification methods that have not been mentioned. Among these, an important technique is 2-D gel electrophoresis in which the first dimension is an IEF gel and the second dimension is SDS-PAGE. This technique can give extremely high resolution of complex protein mixtures so that even if the starting material has not been

previously fractionated at least a proportion of proteins are likely to be resolved as individual spots free of contaminating protein. It is therefore potentially possible to isolate pure protein from the SDS-PAGE gel or from a blot of the gel. However, unless the resolution of the electrophoresis is very high, previous purification steps are likely to be required before a large proportion of proteins will be resolved as individual spots and considerable development may be needed if the technique is to be used preparatively rather than analytically.

A large number of affinity techniques have been used for the purification of protein kinases (Section 4.1.2.6), the main difficulty with the development of such methods being the requirement for some prior information about the properties of the protein kinase being purified. All protein kinases share the property of using ATP as substrate and $Mg^{2+}$ as a co-factor but no affinity matrix with general applicability for protein kinase purifications has been available in the past. The recently developed γ-phosphate-linked ATP Sepharose may prove to have such general applicability. Most protein kinase purifications are likely to involve both classical purification techniques and affinity techniques. It may sometimes be necessary to develop a partial purification procedure using classical techniques before performing further characterisation to assist in the development of affinity purification steps.

## 4.2 Results

### 4.2.1 Development of anion-exchange chromatography

As described in Chapter 3 a number of properties of the 60 kDa autophosphorylating protein kinase from bloodstream *T. brucei* had been determined. 1) The protein partitioned in the 100 000 g supernatant of *T. brucei* extracts. 2) The protein kinase activity from this supernatant was stable for at least 1 month at -20 °C or -70 °C. 3) The isoelectric points of the various isoenzymes were known. 4) The molecular size of the protein kinase had been determined by SDS-PAGE. 5) The protein kinase had been found to be able to utilise both ATP and GTP (data not shown) as the nucleotide substrate and to autophosphorylate on serine and threonine residues. Protein kinase activities that were apparently similar to two of the isoenzymes of the 60 kDa protein kinase from bloodstream forms were also detected from extracts of procyclic *T. brucei* but the activity was much lower. Based on these known parameters, the high speed

supernatant of bloodstream *T. brucei* extracts was chosen as a suitable starting material for protein purification that was free of particulate matter and could be applied to chromatography columns. Ion-exchange chromatography was chosen for the first purification step because of its high protein capacity (the starting material had only been enriched for the 60 kDa protein kinase by removal of insoluble material) and because the estimation of the isoelectric points of the 60 kDa protein kinase isoenzymes from the IEF protein kinase assay could be used to judge the likely behaviour of the 60 kDa protein kinase in ion-exchange. A DEAE anion-exchange column (DEAE is a weak anion-exchanger) was available and appeared suitable.

#### 4.2.1.1 Initial anion-exchange separation

In the initial anion-exchange separation using the DEAE column, a priority was to elute the protein kinase in a small volume at a predictable point in the chromatography run because only a small number of fractions (10) could be assayed immediately for the 60 kDa protein kinase activity using the IEF gel *in situ* protein kinase assay. The protein kinase was known to be stable to freezing in extracts of high protein concentration but it could not be assumed that the same would be true in column fractions. The chromatography conditions used are shown in Fig. 4.1 panel B. The sample was loaded in 20 mM Tris-HCl pH 7.65 as the isoelectric points of the protein kinase isoenzymes that showed highest activity in the IEF assay are between 5.2 and 5.8. As described in Section 4.1.2.2 a starting pH of about 1 pH unit above the isoelectric point of the target protein is usually chosen in anion-exchange. Thus the lowest isoelectric point isoenzymes would be likely to bind quite strongly to the column under the starting conditions, and some of the other isoenzymes (isoelectric points between 5.8 and 7.1) should also bind. Co-elution of multiple isoenzymes in column fractions to give a pattern of labelled bands in the IEF assay similar to that seen when total extracts were assayed would aid identification of the 60 kDa protein kinase (although SDS-PAGE size data could also be used). It seemed unlikely that the isoenzyme of isoelectric point 9.25 would bind to the column under the buffer conditions chosen. Two buffer changes were introduced to elute the protein kinase. The first of these was a pH step to 20 mM Tris-HCl pH 5.7, and the second a salt step to 500 mM NaCl.

The absorbance trace from the anion-exchange separation is shown in Fig. 4.1 panel A with absorbance (at 280 nm) on the vertical axis and time on the horizontal axis. Absorbance units full scale (AUFS) on the spectrophotometer was set initially to 0.2 but the spectrophotometer was rescaled for a short time following the NaCl step (indicated in the figure). The sample was loaded in multiple injections and the run was started at the time of the last injection. The times at which buffer changes were introduced are indicated on the horizontal axis. The trace shows a large absorbance peak following sample injection representing unbound protein and by 20 minutes into the run the absorbance had returned to near baseline level. The pH step introduced at 25 minutes did not lead to a significant increase in absorbance but the 500 mM NaCl step introduced at 33 minutes was followed by a large absorbance peak. Because the pH step had failed to lead to an absorbance peak, it was decided to only analyse fractions following the salt step. Fig. 4.2 panel A shows the autoradiograph of the IEF gel used in the assay of fractions 40-44, which were the first few fractions to show elevated absorbance following the 500 mM NaCl step. The gel lanes are labelled with the fraction number loaded. The autoradiograph shows that there was no phosphate labelled protein in the lanes loaded with fractions 40-41.5 and 43.5-44 but there is labelling in the lanes loaded with fractions 42, 42.5, and 43. Because it was necessary to analyse as many fractions as possible on a single gel it was not possible to include lanes loaded with the 100 000 g supernatant used as the starting material for chromatography or with isoelectric point markers, so the identification of any of the phosphate labelled bands as being due to isoenzymes of the 60 kDa protein kinase from the autoradiograph of the IEF gel is tentative. However, comparison of the phosphate labelled bands from fraction 42 with the autoradiographs of IEF gels loaded with total bloodstream extracts (e.g. Fig. 3.2 panel A) shows some similarity. In both cases the strongest labelling is of the bands focusing nearest the anode and the broad strongly labelled band from fraction 42 could represent bands 1 and 2 from assay of the total bloodstream extract. The very weak band of slightly lower isoelectric point could represent IEF band 6 while the slightly stronger band focusing nearest to the cathode could represent IEF band 7. Following autoradiography of the IEF gel, the gel bands containing the phosphate labelled protein were cut out from the dried gel, rehydrated, and loaded onto an SDS-PAGE gel. The autoradiograph of the SDS-PAGE gel lane loaded with the phosphate labelled protein

from fraction 42.5 is shown in Fig. 4.2 panel B (an attempt was made to analyse the labelled protein from fraction 42, but the transfer of the agarose from the IEF gel to the SDS-PAGE gel was poor). The 60 kDa size observed for the phosphate labelled bands strongly suggests that the activity detected was indeed that of the 60 kDa protein kinase.

#### 4.2.1.2 100 mM NaCl step elution

Having established that the 60 kDa protein kinase would bind to the DEAE column when loaded in a starting buffer of pH 7.65 and that it could be eluted by a 500 mM NaCl step (at a buffer pH of 5.7), the next aim was to refine the conditions to obtain a higher level of purification. It was decided to retain the initial binding conditions and the pH step change but then to elute the protein kinase with successive 100 mM NaCl steps. Fig. 4.3 panel A shows the absorbance trace for the first anion-exchange separation performed using 100 mM NaCl steps for elution while panel B shows the buffer conditions used. As previously, the sample was loaded in multiple injections and sample loading was followed a large absorbance peak. However, in this separation, the pH step (at 28 minutes) was followed by a small absorbance peak. At 40 minutes the first NaCl step (to 100 mM NaCl) was introduced, and this was followed by further 100 mM increases in the buffer NaCl concentration at 50, 60, and 70 minutes. Each NaCl step was followed by a significant absorbance peak.

In this experiment it was not possible to predict at what point in the run the 60 kDa protein kinase was likely to elute. Therefore, the number of fractions that required assay for the 60 kDa protein kinase was larger than could be easily assayed using the IEF protein kinase assay and, to deal with this problem, a new protein kinase assay (referred to as the slot blot assay) was developed. In this assay system (similar to that of Glover and Allis, 1991), a slot blot apparatus is used to blot aliquots from individual column fractions to a PVDF membrane. The liquid is drawn through the membrane under vacuum leaving proteins adsorbed to its surface. The blot is then incubated in a protein kinase assay buffer containing [$\gamma$-$^{32}$P]ATP and is then washed and dried. Phosphate labelling is detected by autoradiography of the blot but, as non-denaturing conditions are used in washing the blot, labelling of protein could be due to ATP-binding activity as well as protein kinase activity. The autoradiographs of the blots used to assay the column fractions are shown in Fig. 4.4. The first few fractions to show elevated absorbance after

sample application contained significant activity but elution of this activity was complete by 6 minutes. No further activity was detected in fractions collected while the column was being washed with the starting buffer or in the fractions showing elevated absorbance following the pH step (fractions 31-40). A set of fractions (fractions 50-53.5) collected following the 100 mM NaCl step introduced at 40 minutes contained activity and a further set of fractions containing activity (fractions 58.5-65) was collected following the 200 mM NaCl step. Small numbers of fractions containing weak activity were also collected following the 300 mM (fractions 69-70.5) and 400 mM salt steps (fractions 78.5-79.5). The slot blot assay could not be used to positively identify the 60 kDa protein kinase and so the IEF assay was used to test for the presence of the 60 kDa protein kinase in some of the fractions identified as containing activity in the slot blot assay. The fractions chosen for assay using the IEF system were fractions 1-2.5 and 60-61.5. Fractions 1-2.5 (collected shortly after sample loading) would obviously not contain the 60 kDa protein kinase isoenzymes that eluted after the 500 mM NaCl step in the first anion-exchange separation but were analysed because they showed a high level of activity in the slot blot assay. Fractions 60-61.5 were the only other fractions that showed a high level of activity in the slot blot assay and might contain the 60 kDa protein kinase isoenzymes that were eluted by the 500 mM NaCl step in the first anion-exchange separation. Fig. 4.5 panel A shows the autoradiograph of the IEF gel used in the assay. The autoradiograph shows no phosphate labelling in the lanes loaded with fractions 1-2.5 and so the activity detected in the slot blot assay of these fractions is entirely distinct from that of the 60 kDa protein kinase. However, labelled proteins were detected in the gel lanes loaded with fractions 60-61.5 which are at a similar position in the gel to the labelled bands detected in IEF assay of fractions from the first anion-exchange separation (Fig. 4.2) and, as previously, a higher isoelectric point activity was only detected in the earliest-eluting fraction. The labelled bands from fractions 60 and 60.5 were analysed by SDS-PAGE and autoradiography (Fig. 4.5 panel B). The 60 kDa size of the single labelled protein detected from each fraction confirms that the activity identified in the IEF assay of the column fractions was that of the 60 kDa autophosphorylating protein kinase.

The result of the first anion-exchange separation using 100 mM NaCl steps was confirmed in a further chromatography run in which with longer periods were allowed

between successive 100 mM NaCl steps. Fig. 4.6 panel A shows part of the absorbance trace with the fractions that were analysed by the IEF assay indicated, panel B shows the buffer conditions while panel C shows the autoradiograph of the IEF gel used in the assay of fractions 67.5-71. It can be seen from panel C that the 60 kDa protein kinase was detected in fractions 69.5-71 but not in fractions 67.5-69, showing that the 60 kDa protein kinase elutes slightly after the main absorbance peak following the 200 mM NaCl step.

#### 4.2.1.3 150 mM and 200 mM NaCl steps

The results of the two chromatography runs using 100 mM NaCl steps for protein kinase elution at a buffer pH of 5.7 showed that isoenzymes of the 60 kDa protein kinase with significant activity in the IEF protein kinase assay could be eluted by a 200 mM NaCl step and that these isoenzymes were not eluted at 100 mM NaCl. In order to refine the elution conditions further, an anion-exchange run was performed in which 100 mM, 150 mM, and 200 mM NaCl steps were used for elution. A new protein kinase assay was developed to assay the column fractions from this run. In this assay column fractions are mixed with a protein kinase assay buffer and the protein kinase labelling reaction is started by addition of [$\gamma$-$^{32}$P]ATP. Following the labelling incubation, the reaction is stopped by addition of SDS-PAGE sample buffer and boiling. The samples are then loaded onto SDS-PAGE gels and phosphate labelled proteins are detected by autoradiography of the dried gels. This assay is referred to as the liquid kinase assay.

Fig. 4.7 panel A shows the absorbance trace for the latter part of the run while panel B shows the buffer conditions used. The sample was loaded in 20 mM Tris-HCl pH 7.65 as in the earlier separations but the pH step to pH 5.7 was introduced earlier. The 100 mM, 150 mM, and 200 mM NaCl steps each produced absorbance peaks. Fig. 4.8 shows the autoradiographs of SDS-PAGE gels used to analyse the liquid kinase assays of fractions 54.5-76.5. The autoradiograph of a gel loaded with assays of fractions 50-54 is not included in the figure but showed no phosphate labelling at any position. The autoradiographs of two sets of fractions showed a labelled band of 60 kDa (and no labelled bands of any other size). The first set (fractions 57-60.5) correspond to the latter part of the absorbance peak following the 150 mM NaCl step while the second set (fractions 73.5-76.5) correspond to the main peak following the 200 mM NaCl step.

The 60 kDa size for the phosphate labelled bands indicated that the liquid kinase assay had identified the 60 kDa protein kinase and the lack of labelling of other proteins showed that the assay was specific in spite of the availability to the protein kinase of all the proteins in the fractions assayed. This specificity is probably due to the activity of the 60 kDa protein kinase being specifically autophosphorylating. The detection of two peaks of activity eluting after 150 mM and 200 mM NaCl steps was not predicted from the previous results but was not particularly surprising considering that multiple isoenzymes were detected in the IEF assay. To simplify discussion of the result from this run the two protein kinase activities have been named PKI (eluted by 150 mM NaCl) and PKII (eluted by 200 mM NaCl).

#### 4.2.1.4 100 mM-200 mM NaCl gradient elution

In order to confirm the result of the 150 mM/200 mM NaCl step run that showed that the 60 kDa protein kinase eluted as two peaks (PKI and PKII), and to determine more precisely the NaCl concentrations required for elution, a further anion-exchange separation was performed in which a linear NaCl gradient (100 mM-200 mM NaCl) was used to elute the protein kinase. Fig. 4.9 panel A shows the absorbance trace and panel B shows the buffer conditions. In this run the starting buffer was 20 mM Tris-HCl pH 5.7 where 20 mM Tris-HCl pH 7.65 had been used previously. This change was introduced because the results of the previous separations suggested that the use of a higher pH starting buffer was unnecessary. The buffer changes during the run and the fractions that were assayed using the liquid kinase assay are indicated in Fig. 4.9. The autoradiographs of the SDS-PAGE gels used to analyse the liquid kinase assays showed a phosphate labelled band at 60 kDa and there was very little labelling at other molecular weights. Following autoradiography, gel slices containing the labelled 60 kDa band were cut from the SDS-PAGE gels and counted by scintillation counting. The results of this analysis are shown overlaid on the absorbance trace (see Fig. 4.9 legend). The profile of radioactive counts clearly confirms the result of the 150 mM/200 mM NaCl step run, although unfortunately a few fractions eluting just after the end of the gradient that clearly would have contained some PKII activity were not assayed. PKI elutes at 148 mM NaCl while PKII elutes at 175 mM NaCl.

To briefly summarise the results of the DEAE anion-exchange chromatography, conditions were identified under which the 60 kDa protein kinase would bind the anion-exchange column and, using increases in buffer NaCl concentration to elute the protein kinase in a buffer that was otherwise 20 mM Tris-HCl pH 5.7, it was shown that the 60 kDa protein kinase could be eluted as two peaks (PKI and PKII) at 148 mM and 175 mM NaCl. During the development of the anion-exchange chromatography new assays were also developed for analysis of the 60 kDa protein kinase activity from column fractions. The liquid assay was found to be robust and informative. The anion-exchange chromatography could be developed further to give optimal purification (see Discussion) but the fractions containing the protein kinase eluted by the finer NaCl steps or by a 100-200 mM NaCl gradient were suitable as the starting material to use in the development of other chromatography methods for purification of the 60 kDa protein kinase.

### 4.2.2 Hydrophobic interaction chromatography (HIC)

The initial aim in the investigation of hydrophobic interaction chromatography for the purification of the 60 kDa protein kinase was to identify a suitable chromatography matrix. In order that this could be achieved rapidly a test kit containing disposable gravity fed columns of different agaroses substituted with straight chain alkyl groups was used. Protein kinase activity under high salt conditions was also investigated since high salt conditions can be used to promote protein adsorption to hydrophobic matrices. The starting material used was the concentrated anion-exchange chromatography fractions containing the PKII activity eluted by a 100-200 mM NaCl gradient (Section 4.2.1.4); this material was pre-labelled (see Section 2.1.8) and adjusted to the starting buffer conditions. The starting buffers were 10 mM Tris-HCl pH 7.2 with 0.5 M or 2 M NaCl.

In order to test whether the protein kinase activity was stable under high salt conditions, the liquid protein kinase assay for the 60 kDa protein kinase was performed with different concentrations of NaCl added to the assay buffer. Fig. 4.10 shows the autoradiograph of the SDS-PAGE gel used to analyse the assays with the final NaCl concentration in the assay buffer indicated above each lane. The positions and sizes of molecular weight markers are indicated to the left of the figure. The labelling of the 60 kDa band in each lane shows that the protein kinase was active at all of the NaCl concentrations tested. The labelling intensity in each lane appears to be similar. Thus the

60 kDa protein kinase is active at high NaCl concentrations and NaCl might be used to promote its interaction with hydrophobic matrices.

The most hydrophobic HIC matrix that was tested was the agarose substituted with dodecyl carbon chains. The sample containing pre-labelled protein kinase was loaded and then the column was washed with 10 column volumes of the starting buffer (10 mM Tris-HCl pH 7.2, 0.5 M NaCl). Fig. 4.11 panel A shows a silver-stained SDS-PAGE gel loaded with aliquots from consecutive fractions eluting during application of the starting buffer. There is very little eluted protein apparent and the labelled kinase was not detected by autoradiography (not shown). Radioactive counts were detected at the top of the column (using a Geiger counter) after washing with starting buffer and in order that this radioactive material could be analysed the column was dismantled and the agarose containing the counts was boiled in SDS-PAGE sample buffer. The sample buffer and suspended agarose was then loaded onto an SDS-PAGE gel. Fig. 4.11 (panel B) shows two silver-stained SDS-PAGE gel lanes (left) loaded with different portions of boiled agarose and autoradiographs of these gel lanes (right). The silver stains show a dark smear through which protein bands can just be seen, while the autoradiographs show the phosphate labelling in a low molecular weight band and in a smear near the top of the gel. It was noticed in dismantling the gel assembly after electrophoresis that the agarose left in the wells was still significantly radioactive. Therefore, it seems likely that most of the 60 kDa protein kinase was bound irreversibly to the agarose.

The binding of the pre-labelled kinase to an agarose substituted with hexyl carbon chains was examined next. Fig. 4.12 panel A shows silver-stained gels containing aliquots from consecutive fractions eluting in the starting buffer (which in this experiment contained 2 M NaCl) with the fraction numbers labelled above the gel lanes. The leftmost lane contains an aliquot of the sample material applied to the column. The autoradiographs of these gels are shown in panel B and show a prominent phosphate labelled band at 60 kDa. After washing of the column with 10 column volumes under the starting conditions the buffer was changed to 10 mM Tris-HCl pH 7.2. Autoradiography of SDS-PAGE gels loaded with these fractions showed no phosphate labelling. These results show that the 60 kDa protein kinase failed to bind to the hexyl agarose under the conditions used.

The binding of the pre-labelled 60 kDa protein kinase to octyl agarose was also tested. Fig. 4.13 panel A shows a silver-stained SDS-PAGE gel containing consecutive fractions eluting in the starting buffer (10 mM Tris-HCl pH 7.2, 0.5 M NaCl). No phosphate labelled protein was detected by autoradiography of this gel. After washing of the column with several column volumes of the starting buffer, the buffer was changed to 10 mM Tris-HCl pH 7.2. Lanes 1-8 in the gel shown in panel B were loaded with aliquots from consecutive fractions eluting after the buffer change. Autoradiography of this gel again detected no radioactive protein. In order to elute the 60 kDa protein kinase, the agarose was removed from the top of the column and boiled in sample buffer before being loaded onto a gel. Fig. 4.13 panel C shows the silver stained SDS-PAGE gel separation of the boiled agarose (left) and the autoradiograph of the gel (right). The autoradiograph shows a prominent phosphate labelled band of 60 kDa although, due to over-concentration of the anion-exchange fraction before labelling, other proteins have also become significantly labelled.

The results from these experiments and possible future experiments to develop HIC for the purification of the 60 kDa protein kinase are discussed in section 4.3. It was found that the 60 kDa protein kinase bound irreversibly to the dodecyl agarose but failed to bind the hexyl agarose when applied in a 2 M NaCl buffer. However, the protein kinase did bind to the octyl agarose but was only eluted under denaturing conditions; elution under non-denaturing conditions was not tested.

### 4.2.3 ATP affinity chromatography

The use of ATP affinity chromatography for the purification of protein kinases is discussed in Section 4.1.2.6. The ATP affinity matrix that was tested was of the type with the ATP linked to the support through the N-6 amino group of the adenine ring. The sample material used to evaluate whether the ATP agarose could be used in the purification of the 60 kDa protein kinase was the anion-exchange fractions containing PKII that had been pooled, concentrated and shown to contain the 60 kDa protein kinase activity. Three separate binding reactions were performed in 20 mM Tris-HCl pH 7.0 with either no NaCl, 10 mM NaCl or 50 mM NaCl. A salt such as NaCl is often included in chromatography buffers to reduce non-specific protein-matrix interactions. However, low salt buffers have been found to be necessary in some cases for protein kinase

purifications using ATP matrices (see Haystead *et al.*, 1993). The same buffer containing 800 mM NaCl was used to elute the protein kinase (the protein kinase activity was known to be stable to high NaCl concentration). The binding incubations were carried out for 20 minutes at 4 °C after which time the agarose was spun down and the supernatant removed before addition of the elution buffer. After incubation in the elution buffer the suspension was again centrifuged and the supernatant removed. Liquid kinase assays were then performed on the supernatants from the binding and elution steps and then analysed by SDS-PAGE.

Fig. 4.14 panel A shows the silver-stained SDS-PAGE gel loaded with the products of the kinase assays while panel B shows the autoradiograph of the gel. Lanes 1 and 2 contain molecular weight markers, lanes 3, 4, 5 were loaded with kinase reactions performed on supernatants from the binding reactions performed in 0, 10, and 50 mM NaCl buffers respectively, while lanes 8, 9, and 10 were loaded with the corresponding reactions from the elution step. There is a phosphate labelled protein of 60 kDa in all three lanes loaded with protein kinase reactions from the supernatants of the binding reactions (panel B) but no labelling of any protein in those containing supernatants from the elution step. Therefore, under the conditions used, the protein kinase did not bind the ATP matrix but was readily detectable in the supernatants of the binding reactions.

### 4.2.4 H-9 affinity chromatography

The protein kinase inhibitor N-(2-Aminoethyl)-5-isoquinolinesulphonamide (H-9) was found to be inhibitory towards the autophosphorylating activity of the 60 kDa protein kinase in preliminary tests of the effect of various protein kinase inhibitors on this activity. In order to test H-9 as an affinity ligand for the purification of the 60 kDa protein kinase, H-9 was coupled to cyanogen-bromide activated Sepharose following the method of Inagaki *et al.* (1985) as described in Chapter 2. A batch of control Sepharose was prepared identically except that it was reacted only with the blocking agent (ethanolamine) and not with H-9.

#### 4.2.4.1 Initial tests of protein kinase binding to H-9 Sepharose

A preliminary test of the binding of the 60 kDa protein kinase to H-9 Sepharose was performed with material eluted from the DEAE anion-exchange column by a 500 mM

NaCl step that had been pre-labelled and then concentrated. The sample was incubated with a small volume of H-9 Sepharose in 25 mM Tris-HCl pH 7.0 and the Sepharose was then washed 6 times with more buffer. Significant radioactive counts could be detected associated with the H-9 Sepharose (using a Geiger counter) following the washes. The Sepharose was then boiled in SDS-PAGE sample buffer and loaded onto an SDS-PAGE gel. Fig. 4.15 panel A shows the silver-stained gel lane containing the agarose and an autoradiograph of the gel. It can be seen that the predominant phosphate labelled protein (which is a doublet) is 60 kDa in size. Therefore, it appeared from this experiment that at least some of the 60 kDa protein kinase had bound to the H-9 Sepharose, but the results did not prove that this binding was specific and the silver-stained gel shows many protein bands of other sizes.

In order to test whether the binding of the 60 kDa protein kinase to the H-9 Sepharose was specific, the binding of pre-labelled protein kinase to H-9 Sepharose was compared with its binding to control Sepharose. In this experiment the sample used was pooled anion-exchange fractions (eluted by a 100 mM-200 mM NaCl gradient) containing both PKI and PKII activities that were concentrated and then pre-labelled. Fig. 4.15 panel B shows a silver stained SDS-PAGE gel lane loaded with an aliquot of the pre-labelled sample and the corresponding autoradiograph. The predominant labelled band at 60 kDa is very broad but appears to be a doublet. Aliquots of this material were incubated with equal portions of H-9 Sepharose and control Sepharose. The final buffer conditions in the binding reactions were 30 mM Tris-HCl pH 7.0, 25 mM NaCl. Following the binding incubations, the two Sepharoses were washed four times with 25 mM Tris-HCl pH 7.0. Fig. 4.15 panel C shows silver stained SDS-PAGE gel lanes and autoradiographs used in the analysis of the supernatants from the binding incubations and first washes. It can be seen from the autoradiographs that unbound 60 kDa protein kinase was readily detectable in both supernatants from the binding incubations. However, comparison of the autoradiographs of the supernatants from the first washes suggested that more of the 60 kDa protein kinase was eluted from the control Sepharose than from the H-9 Sepharose. In addition, after four washes, radioactive counts could be detected bound to the H-9 Sepharose (using a Geiger counter) while almost no counts were detected bound to the control Sepharose. Thus it appeared that there was specific binding of the 60 kDa protein kinase to the H-9 Sepharose. To investigate different

elution conditions, the H-9 Sepharose with the bound 60 kDa protein kinase was divided into two portions. One portion was incubated in buffer containing 1% Triton-X100 while the other was incubated in buffer containing 100 mM NaCl. The 100 mM NaCl removed most of the counts (as judged with a Geiger counter) while the detergent removed very few. The supernatants from the two elution conditions were analysed by SDS-PAGE and autoradiography. The silver stained gel lanes and autoradiographs used in the analysis of the Triton X-100 and 100 mM NaCl eluates are shown to the left of panels D and E respectively (Fig. 4.15). The labelled 60 kDa protein kinase is barely detectable in the Triton X-100 eluate (the silver stained gel also shows that very little other protein was eluted) but was the only phosphate labelled protein detected, while the autoradiograph of the gel lane loaded with the 100 mM NaCl eluate shows a prominent doublet at 60 kDa and one much weaker band at <45 kDa. Since there were still detectable counts associated with the Sepharose that had been incubated in 1% Triton X-100, this Sepharose was then incubated in buffer containing 2 mM ATP as a further test of elution conditions. The silver stained SDS-PAGE gel lane and autoradiograph used to analyse the supernatant from this incubation are shown in panel D alongside those used to analyse the supernatant from the incubation in 1% Triton X-100. It can be seen that 2 mM ATP was a much more effective eluant of the 60 kDa protein kinase than the detergent. The silver-stained gel lanes containing the 100 mM NaCl and 2 mM ATP supernatants showed a relatively small number of protein bands and so 5x the volume of these supernatants were used on a larger format SDS-PAGE gel to determine whether a polypeptide corresponding to the 60 kDa protein kinase could be detected. The silver stained gel lanes and autoradiographs are shown to the right of panels D and E. There is a faint silver-stained band at the position of the upper band of the 60 kDa protein kinase doublet detected by autoradiography from both NaCl and ATP supernatants (which could be due to a single protein species) while the lower band of the doublet corresponds in position to a broader silver-stained band which presumably consists of more than one species. This analysis was extended by running further portions of the two supernatants in 2-dimensional gel electrophoresis. The volume of supernatant loaded onto the 1st-dimension gels was the same as was loaded onto the smaller format SDS-PAGE gels shown in Fig. 4.15 panels D and E. Fig. 4.16 panel A shows the silver-stained 2nd dimension (SDS-PAGE) gel used in the analysis of the 100 mM NaCl supernatant while

panel B shows the corresponding autoradiograph. The 60 kDa protein kinase doublet was the only phosphate labelled protein detected by autoradiography of the gel. The two spots in the autoradiograph correspond almost exactly in position with two spots in the silver-stained gel (arrowed) in the vertical (SDS-PAGE) dimension, and are close to, but do not correspond exactly in position, to these spots in the horizontal (IEF) dimension. The upper of the spots in the autoradiograph is further away from the arrowed silver-stained spot at the same vertical position than is the lower (which suggests that the discrepancy in position cannot be due to shrinkage of the gel). The autoradiograph of the 2nd dimension gel from analysis of the ATP eluate showed the same two spots as seen in the Fig. 4.16 panel B but no protein spots were visible at any position in the silver stained gel. Thus the conclusion from the 2-dimensional gel electrophoresis is that an amount of the 60 kDa protein kinase that was detectable by autoradiography was not detectable by silver-staining although it is conceivable that the two silver-stained spots are forms of the 60 kDa protein kinase that do not autophosphorylate in the liquid kinase assay.

#### 4.2.4.2 ATP gradient elution of the 60 kDa protein kinase from an H-9 Sepharose column

An HPLC minicolumn packed with H-9 Sepharose was prepared so that ATP gradient elution could be tested for the purification of the 60 kDa protein kinase. The starting material used in this experiment was a portion of the pre-labelled protein kinase analysed by SDS-PAGE (shown in Fig. 4.15 panel B) except that an additional buffer exchange step was used to reduce the NaCl concentration to approximately 2 mM. Fig. 4.17 panel B shows the run conditions used. Following sample application the column was washed with approximately 12.5 column volumes of the starting buffer at a flow rate of 0.5 ml/minute. The flow rate was then reduced to 0.2 ml/minute and a 0-10 mM ATP linear gradient was applied between 25 and 45 minutes. Washing of the column with buffer containing 10 mM ATP was continued between 45 and 55 minutes and this was followed by a change to buffer containing 1 M NaCl at 55 minutes. Fig. 4.17 panel A shows the profile of radioactive counts (taken on whole fractions) eluted from the column between 25 and 65 minutes with counts per minute on the vertical axis and time on the horizontal. No absorbance measurements were taken due to an equipment failure. Although the

radioactive counts eluted are only shown for the latter part of the run all fractions were counted. The first few fractions collected after sample application contained a high level of radioactivity but subsequent analysis by SDS-PAGE and autoradiography showed that this was not due to elution of the 60 kDa protein kinase (see below). It can be seen from panel A that the peak of radioactivity detected in the latter part of the run began to elute towards the end of the ATP gradient and was mostly eluted when the ATP concentration in the eluate had already reached 10 mM. The 1 M NaCl step failed to produce a second peak, suggesting that the ATP had eluted all of the bound 60 kDa protein kinase. The fractions that contained >1000 cpm were analysed by SDS-PAGE and autoradiography. Fig. 4.18 shows silver-stained SDS-PAGE gels (panels A and C) and autoradiographs (panels B and D) used in this analysis with the fraction number loaded indicated above each lane and the positions and sizes of protein molecular weight markers shown to the left. Fractions 3, 4, and 5 contained significant counts but the autoradiograph shows no phosphate labelling. It is likely that the counts detected in these fractions were due to radioactive ATP or phosphate. The elution profile of the doublet at 60 kDa in the autoradiographs of the gel lanes loaded with fractions 45-57 corresponds well with the profile of radioactive counts shown in Fig. 4.17 panel A. Although no absorbance trace was obtained for this experiment, it can be seen by comparing the silver-stained gel lanes loaded with fractions 3, 4 and 5 with those loaded with fractions that were found to contain the 60 kDa protein kinase, that a major contaminating protein at 66 kDa (and some proteins of other sizes) were excluded from the protein kinase containing fractions. Therefore, the H-9 affinity chromatography gave a degree of purification of the 60 kDa protein kinase, although this has still to be quantified. The result of this experiment also confirmed the earlier result showing that ATP could be used to elute the 60 kDa protein kinase from H-9 Sepharose (H-9 is known to inhibit a number of protein kinases competitively with ATP) but it has not been shown that elution of the 60 kDa protein kinase is due to ATP binding at the protein kinase ATP-binding site.

In summary, these experiments demonstrate that the pre-labelled 60 kDa protein kinase binds specifically to H-9 Sepharose. The protein kinase was partially eluted from H-9 Sepharose by incubation in 2 mM ATP but using column chromatography it was shown that the bulk of the protein kinase eluted as a single peak at a higher ATP concentration. The protein kinase was also effectively eluted from H-9 Sepharose by 100

mM NaCl. An experiment in which an NaCl gradient was used to elute pre-labelled 60 kDa protein kinase from the H-9 Sepharose column was unsatisfactory due to unexplained weak labelling of the 60 kDa protein kinase (data not shown) but the results appeared to suggest that the protein kinase begins to elute at approximately 25 mM NaCl. While this experiment would have to be repeated in future work, the preliminary result is consistent with the finding that 100 mM NaCl is an effective eluant. It would also provide a possible explanation of why a large amount of protein kinase failed to bind to the H-9 Sepharose in the experiment in which protein kinase binding to H-9 and control Sepharoses was compared (Fig. 4.15, panel C), since the NaCl concentration in the binding reactions was approximately 25 mM (in later experiments NaCl remaining after elution of the protein kinase from the anion-exchange column was more thoroughly removed).

### 4.2.5 Size-exclusion chromatography

Size-exclusion chromatography is generally most useful in the later stages of a protein purification since sample capacity is relatively low, but since a size-exclusion column suitable for protein work (separation range 10 kDa-300 kDa, 30 cm long, 8 mm internal diameter) was available it was decided to perform a test run at an early stage on the high speed supernatant of the total *T. brucei* bloodstream extract. The run was performed under isocratic conditions using a 20 mM Tris-HCl pH 5.7, 200 mM NaCl buffer. The flow rate was 0.8 ml/minute and fractions were collected every 30 seconds.

Fig. 4.19 panel A shows the absorbance trace (280 nm) for the size-exclusion run. A single sample injection of the high speed supernatant of the crude extract (~3% of column volume) was made at time 0. Note that there were two changes to the scaling of the spectrophotometer output at about 6.5 minutes (AUFS from 1.0 to 2.0) and at about 14.5 minutes (AUFS from 2.0 to 1.0). The liquid kinase assay was used to analyse fractions from 1-30 minutes. Fig. 4.19 panel B shows the autoradiograph of the SDS-PAGE gel loaded with the liquid kinase assays of the fractions that were found to contain activity. The fraction number analysed is indicated above each gel lane while the positions and sizes of molecular weight markers are indicated to the left of the figure. The fractions containing protein kinase activity eluted at a position corresponding to the latter part of the first absorbance peak following sample loading, suggesting that the

elution profile may be consistent with an apparent molecular weight for the protein kinase in size-exclusion chromatography similar to its apparent molecular weight in SDS-PAGE, rather than with a higher molecular weight. However, the size-exclusion chromatography would have to be repeated after column calibration to give a real molecular weight estimate for the 60 kDa protein kinase. The autoradiograph also shows labelling of proteins at molecular weights other than 60 kDa although the 60 kDa protein kinase (which is clearly a doublet in the original autoradiograph) is still by far the most prominent.

### 4.3 Discussion

The aim of the work described in this chapter was to purify the 60 kDa autophosphorylating protein kinase first identified using the IEF gel *in situ* protein kinase assay (Chapter 3) so that the purified protein could be microsequenced or injected into animals to provide an antiserum. Anion-exchange chromatography was used as the first purification step and an affinity chromatography step was also developed using a Sepharose to which a protein kinase inhibitor (H-9) had been immobilised. Preliminary investigations were also made into the suitability of a number of other chromatography matrices (discussed below). Sequential anion-exchange chromatography, H-9 affinity chromatography and SDS-PAGE were not sufficient to purify the 60 kDa protein kinase so that it could be resolved as a single band of protein that could be separated from other proteins for microsequencing and neither the amount of protein kinase per *T. brucei* cell nor the purification factor achieved were quantified. The results obtained and the future work that would be required to complete the purification of the 60 kDa protein kinase are discussed below.

The reasons for the decision to use the 100 000 g supernatant of the extract of bloodstream form *T. brucei* as the starting material for the purification of the 60 kDa protein kinase and for the choice of anion-exchange chromatography as the first purification step are discussed in Section 4.2.1. The early development of the DEAE anion-exchange chromatography was complicated by the fact that at the start only the IEF assay (Chapter 3) was available to use to identify the 60 kDa protein kinase in the column fractions. This assay is unsuitable for analysis of large numbers of samples, and until more appropriate assays were developed, step changes in buffer conditions rather

than gradients were used so that only column fractions eluted after each step change had to be assayed to give a reasonable chance that the fractions containing the 60 kDa protein kinase would be identified. In the initial run (Section 4.2.1.1) it was found that the 60 kDa protein kinase would bind to the DEAE column in a pH 7.65 buffer, was not eluted by a pH shift to pH 5.7, but could be detected using the IEF assay in fractions eluted after a 500 mM NaCl step. Elution conditions were then refined using 100 mM NaCl steps (Section 4.2.1.2) and at the same time a new assay (the slot blot assay) was developed. This assay was rapid and was successfully used to eliminate large numbers of fractions that showed no activity. However, in addition to those fractions which showed activity in the slot blot assay and which were later shown to contain the 60 kDa protein kinase using the IEF assay, other fractions showed activity in the slot blot assay but did not contain the 60 kDa protein kinase activity. Using the combination of the slot blot and IEF assays it was found that the 60 kDa protein kinase activity eluted after a 200 mM NaCl step (Section 4.2.1.2). Therefore, in spite of the limitations imposed by the lack of more appropriate assays, buffer conditions for protein kinase binding and elution were rapidly identified and partly refined.

At the next stage in the development of the anion-exchange chromatography 150 mM and 200 mM NaCl steps were used to elute the protein kinase (Section 4.2.1.3), and a further assay (the liquid kinase assay) was tested. It was found that the 60 kDa protein kinase eluted as two peaks (PKI and PKII) after the 150 mM and 200 mM NaCl steps respectively and that the 60 kDa protein kinase was the only phosphate labelled protein detected. Finally, in a chromatography run using a 100 mM-200 mM NaCl linear gradient, it was shown that the two forms of the 60 kDa protein kinase elute at 148 mM (PKI) and 175 mM NaCl (PKII). The fractions containing the 60 kDa protein kinase eluted from the anion-exchange column were suitable for use as the starting material for the development of new chromatography methods and, since a complete purification of the 60 kDa protein kinase would inevitably involve a number of purification steps, it was decided to focus on the identification of new chromatography matrices for purification. However, in retrospect, at this stage it would have been better to have performed at least some further development of the anion-exchange chromatography. Three aspects in particular were unsatisfactory: 1) It was not decided whether to purify both PKI and PKII activities (Sections 4.2.1.3 and 4.2.1.4) or just one and some simple experiments to

determine whether it might be possible to distinguish between the autophosphorylated forms of PKI and PKII by SDS-PAGE and autoradiography were not carried out. 2) The buffer conditions used in the chromatography could probably have been developed to give a higher level of protein kinase purification from runs of shorter duration. 3) The chromatography conditions could have been standardised and the liquid kinase assay developed further to allow an estimate of the purification factor obtained in the anion-exchange chromatography to be determined and to provide a standardised partially purified preparation of the 60 kDa protein kinase to use in the development of further chromatography steps. These aspects are discussed in turn below, although in practice they are interrelated and some flexibility would be required in the approach taken to further development of the anion-exchange chromatography.

To behave differently in anion-exchange the 60 kDa polypeptides PKI and PKII (Section 4.1.2.3 and 4.1.2.4) must differ in some way from each other (unless one or both are non-covalently associated with other molecules) and it might be expected that differences in mobility between PKI and PKII labelled by autophosphorylation would be observed in autoradiographs of SDS-PAGE gels (unless the differences in behaviour of PKI and PKII were due to phosphorylation state and the labelling reaction of the liquid kinase assay lead to autophosphorylated PKI and PKII becoming identical). Unfortunately, liquid kinase assays on fractions containing PKI and PKII activities were never analysed in adjacent SDS-PAGE gel lanes and so small mobility differences could have been missed. An obvious experiment would be to label PKI and PKII separately and run three adjacent SDS-PAGE gel lanes containing PKI only, PKII only, and a mixture of the two (mixed after denaturation in sample buffer) and then expose the gel to film. The autoradiograph of such a gel would show whether autophosphorylated PKI and PKII have different mobilities in SDS-PAGE and if so which form had higher mobility and which form had lower. In most of the work to identify new chromatography matrices for protein kinase purification, using protein kinase partially purified by anion-exchange, only fractions containing PKII were used and the 60 kDa phosphate labelled band appeared as a singlet in these experiments. However, in an experiment in which the 100 000 g supernatant of bloodstream *T. brucei* extract was used as the sample in size-exclusion chromatography (Section 4.2.5), the 60 kDa protein kinase appeared as a doublet in the autoradiographs of SDS-PAGE gels used to analyse liquid kinase assays

performed on the column fractions (Fig. 4.19). In addition, in the experiments to test binding of the 60 kDa protein kinase to H-9 Sepharose the samples used were pooled anion-exchange fractions containing both PKI and PKII activities that were pre-labelled before binding to H-9 Sepharose (see Section 4.2.4 for more details) and autoradiography of aliquots of the samples (Fig. 4.15 panels A and B) and of the eluates from the H-9 Sepharose (Figs. 4.15, 4.16, and 4.18) showed the labelled 60 kDa protein kinase as a doublet. Therefore, although it has not been shown whether the doublets observed in the autoradiographs from these experiments do represent the autophosphorylated forms of PKI and PKII, at least two forms of the 60 kDa protein kinase with different mobility in SDS-PAGE have been detected.

It is possible that PKI and PKII only differ as a result of secondary modifications, and if, for instance, their different behaviour in anion-exchange is due to one form being more heavily phosphorylated than the other, then it might be possible to convert the more phosphorylated form to the less phosphorylated form by phosphatase treatment. It would also be interesting to examine the behaviour of the labelled (i.e. autophosphorylated) forms of PKI and PKII in anion-exchange chromatography (and also to perform phosphoamino acid analysis). If the autophosphorylated forms behaved significantly differently from PKI and PKII it might be possible to achieve a useful level of purification by a second anion-exchange step. In summary, a small number of simple experiments would show whether autophosphorylated PKI and PKII have the same mobility or different mobility in SDS-PAGE and it would also be relatively simple to test the behaviour of PKI and PKII in size-exclusion and H-9 affinity chromatography. This would aid interpretation of earlier results and would allow a more informed decision to be made about whether to purify both PKI and PKII or just one activity.

As stated above, the buffer conditions used in the anion-exchange chromatography might be refined further. In particular, the starting buffer conditions might be changed. Throughout most of the development of the anion-exchange chromatography the starting buffer was 20 mM Tris-HCl pH 7.5 and a pH shift to pH 5.7 used in the first separation was retained. In the gradient separation (Section 4.2.1.4) this was altered so that the binding buffer had a pH of 5.7. However, even with this alteration the starting buffer conditions are probably still rather conservative. PKI and PKII remained bound to the DEAE column at 100 mM NaCl after loading in a buffer

lacking NaCl, and 20 mM Tris-HCl pH 5.7, 100 mM NaCl might be tested as a starting buffer. If liquid kinase assays of fractions eluted shortly after sample loading showed significant amounts of unbound protein kinase then starting buffers containing lower NaCl concentrations could be tested. The use of more stringent starting buffer conditions would leave more contaminating protein unbound which could lead to improvements in the level of purification attained, would reduce the duration of the chromatography runs required and would also allow more sample to be applied to the column without overloading.

The IEF and slot blot assays were not suitable for providing a quantitative estimate of the activity of the 60 kDa protein kinase in column fractions. However, the liquid kinase assay with scintillation counting of the SDS-PAGE gel bands containing the 60 kDa protein provides a more suitable assay for a quantitative measurement of protein kinase activity. Using the liquid kinase assay in this manner with some standardisation of the chromatography procedure it would have been possible to perform a more detailed characterisation of the anion-exchange chromatography. With a standard amount of the 100 000 g supernatant loaded in each chromatography run and protein concentration measurements (e.g. by Bradford assays) taken on the fractions found to contain PKI and PKII and the starting material, a purification factor (i.e. amount of protein in the fractions containing the protein kinase activity divided by the total amount of protein loaded) could be calculated for a particular set of run conditions. The availability of this information would make it possible to compare the suitability of anion-exchange runs performed using different sets of run conditions. Once the conditions for a standardised anion-exchange run had been decided on it would be possible to use the liquid kinase assay to assess quantitatively factors such as the effect of freezing on protein kinase activity (which so far has only been assessed qualitatively). It would also be interesting to concentrate samples of the standardised protein kinase preparation and to then assay serial dilutions to determine at what protein concentration proteins other than the 60 kDa protein become labelled. As discussed below the liquid kinase assay might also be used to provide an estimate of the amount of the 60 kDa protein.

The second chromatography step developed for the purification of the 60 kDa protein kinase that would be likely to form part of a complete purification procedure was H-9 affinity chromatography. H-9 Sepharose was tested as an affinity matrix for

purification of the 60 kDa protein kinase because H-9 was found to be inhibitory to the kinase and because H-9 Sepharose is well characterised as an affinity matrix for the purification of PKC, PKA, and PKG (Inagaki *et al.*, 1985). The 60 kDa protein kinase was found to bind to H-9 Sepharose and could be effectively eluted by 100 mM NaCl or by 10 mM ATP (2 mM ATP caused some elution). The results of another experiment (data not shown) using a NaCl gradient suggested that the 60 kDa protein kinase elutes from H-9 Sepharose at about 25 mM NaCl. This result would provide a possible explanation of why much of the kinase failed to bind H-9 Sepharose in an earlier experiment (see Section 4.2.4.2) in which the NaCl concentration was above 25 mM.

The elution of the 60 kDa protein kinase from H-9 Sepharose by a low concentration of NaCl limits the stringency of the conditions that can be used for binding of the protein kinase to H-9 Sepharose and subsequent washes. Therefore it is necessary to try to maximise the efficiency of the elution step by using a specific eluant. ATP was used as a specific eluant and in future work GTP might also be tested as well as the effect of the inclusion of $Mg^{2+}$ ions in the elution buffer. While it was clearly demonstrated that the binding of the 60 kDa protein kinase to H-9 Sepharose was dependent on H-9 in a comparison of the binding of the protein kinase to H-9 Sepharose and control Sepharose, it would be necessary in future work to estimate a purification factor for the H-9 chromatography step, and this would be best performed individually for standardised preparations of PKI and PKII.

A preliminary investigation was carried out to evaluate the feasibility of using HIC for purification of the 60 kDa protein kinase. The autophosphorylating activity of the 60 kDa protein kinase was shown to be stable to 2 M NaCl suggesting that it might be possible to use salt-promoted adsorption in HIC of the 60 kDa protein kinase. The pre-labelled protein kinase failed to bind to the hexyl agarose tested while its binding to the dodecyl agarose was apparently irreversible, even after boiling of the agarose in SDS-PAGE sample buffer. The protein kinase bound to the octyl agarose in a 0.5 M NaCl starting buffer but failed to elute when a similar buffer lacking NaCl was applied to the column. In contrast to the situation with the dodecyl agarose, the protein kinase was completely removed from the agarose by boiling in SDS-PAGE sample buffer. While the stability of the protein kinase activity to high salt conditions suggests that HIC might be a suitable purification method for the 60 kDa protein kinase the development of HIC for

the purification of the 60 kDa protein kinase would require substantial further experimentation. As well as further work with matrices substituted with 6-8 carbon straight chain alkyl groups, matrices substituted with phenyl groups and branched-chain alkyl groups might be tested (no source was identified for a heptyl agarose). In performing further screens for suitable HIC matrices the buffer conditions used should probably also be changed. NaCl was included in the starting buffers to promote protein adsorption but NaCl is a relatively weak salting-out agent and might only be effective at higher concentrations. Ammonium sulphate at 1 M concentration would probably have been a better choice. While salt-promoted adsorption of the protein kinase to an HIC column followed by elution in a decreasing salt gradient represents the ideal, it is quite common in HIC purifications to have to use additives such as ethylene glycol to promote protein desorption and no such agents were tested. In summary, under the conditions tested the 60 kDa protein kinase bound an agarose substituted with octyl carbon groups while failing to bind a hexyl agarose. It is not known whether the protein kinase can be eluted from the octyl agarose without denaturation. Although much more work would be required to develop HIC as a purification step, the protein kinase activity was stable in a buffer containing 2 M NaCl suggesting that the protein is sufficiently robust for HIC to be used for its purification.

As described in Section 4.2.3 the 60 kDa protein kinase failed to bind to an ATP affinity matrix tested. This could have been due to the use of a buffer that was unsuitable for binding but the conditions used were mild and the 60 kDa protein kinase activity was readily detected in assay of the supernatants from the binding incubations (Fig. 4.14). The ATP agarose tested was of the type with the ATP immobilised through the adenine N-6 group. This may not be a suitable linkage to produce an affinity matrix for protein kinase purification as has also been suggested for the linkages involving the adenine C-8 group and the ribose hydroxyls (see discussion in Section 4.1.2.6). However, a γ-phosphate-linked ATP-Sepharose of the type developed by Haystead *et al.* (1993) has become commercially available since the work directed towards the purification of the 60 kDa protein kinase was discontinued. This matrix has been used in the purification of a MAP kinase kinase (Haystead *et al.*, 1993), an AMP-activated protein kinase (Davies *et al.*, 1994) and an IL-1 activated protein kinase (Kracht *et al.*, 1994). It thus appears that the γ-phosphate-linked ATP-Sepharose has promise as a general affinity matrix for

protein kinase purifications, and should certainly be tested for its suitability for the purification of the 60 kDa protein kinase in any future work to complete its purification. Although binding of the 60 kDa protein kinase to H-9 Sepharose is likely to be through the ATP-binding site, the affinity of the protein kinase for the γ-phosphate-linked ATP-Sepharose would probably be higher than for H-9 Sepharose allowing more stringent washing conditions, however, fewer contaminating proteins are likely to bind specifically to the H-9 Sepharose. Therefore, it seems likely that chromatography with both types of matrix would be useful in a full purification procedure.

A single size-exclusion chromatography separation was performed as a preliminary investigation to find out whether protein kinase activity could be recovered after chromatography of the high speed supernatant of the extract of bloodstream forms. Using the liquid kinase assay the 60 kDa protein kinase activity was detected in fractions eluted towards the latter part of the first absorbance peak following sample application (Section 4.2.5, Fig. 4.19). Size-exclusion chromatography represents the mildest purification step used prior to assay of any that were investigated and, in this respect, it is interesting that the protein kinase was detected as a doublet. As discussed above, the two bands of the doublet could represent PKI and PKII (Section 4.1.2.4) but this has not been established. It would be useful to calibrate the size-exclusion column and then to determine molecular weights under non-denaturing conditions for the 60 kDa protein kinase purified by size-exclusion chromatography alone and for PKI and PKII from anion-exchange fractions. It might also be possible to perform experiments to determine whether PKI and PKII interact. Size-exclusion chromatography might be used in the future for purification of the 60 kDa protein kinase but would be best used at a late stage in the purification procedure since it is a low capacity technique. The applicability of size-exclusion chromatography as a purification method for the 60 kDa protein kinase would also depend on its apparent molecular size in size-exclusion chromatography after earlier purification steps. If the 60 kDa polypeptide(s) form multimers or are associated with other polypeptides, then size-exclusion chromatography might be a good purification step, however, if the apparent molecular size of the protein kinase in size-exclusion chromatography is close to 60 kDa this method would probably not be useful for purification since SDS-PAGE would be likely to be the last step of a full purification procedure.

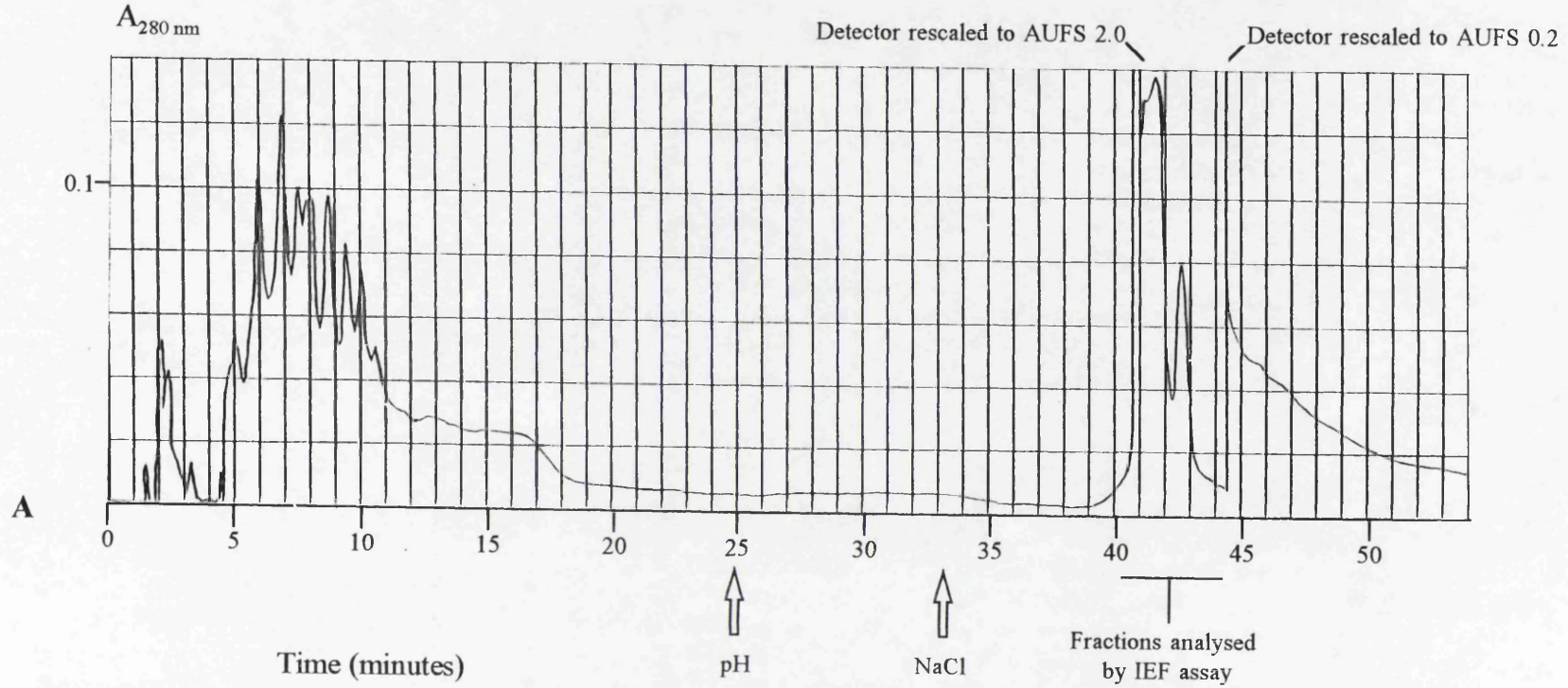
It was decided to discontinue work on the purification of the 60 kDa protein kinase and to investigate the PCR amplification of *T. brucei* protein kinase cDNAs as an alternative strategy for the study of protein kinases in *T. brucei* (Chapter 5). This decision was taken because of time constraints, particularly as only anion-exchange chromatography and H-9 Sepharose affinity chromatography had been developed. It also remains undetermined whether or not the amount of the 60 kDa protein kinase per bloodstream *T. brucei* cell might be so small as to make protein purification in quantities sufficient for microsequencing unrealistic. An estimate of the amount of the 60 kDa protein kinase in an SDS-PAGE gel band could be calculated from a measurement of radioactivity incorporated in the liquid kinase assay (by scintillation counting). However, such a calculation would require assumptions to be made, one of which would be that a certain number of radiolabelled phosphate groups are added to each 60 kDa polypeptide in the liquid kinase assay. Unfortunately, this important factor cannot be determined with any certainty. Another way in which it might be possible to radiolabel the 60 kDa protein kinase for the purpose of estimating the amount of protein would be to use an affinity label (e.g. [$^{14}$C]$FSO_2PhCOAdo$, see Zoller *et al.*, 1981) that incorporated a radioactive group and that would irreversibly inactivate the protein kinase. This would potentially be quite an accurate way in which to quantify the protein since only one label should be incorporated per available active site but this would require more development.

An alternative to estimating the amount of protein kinase purified from given mass of cells following anion-exchange and H-9 affinity chromatography would be to first develop at least one further chromatography step using SDS-PAGE and 2-dimensional gel electrophoresis to analyse fractions containing the 60 kDa protein kinase activity in the hope that it would become possible to visualise the 60 kDa protein kinase as a distinct band or spot by silver staining. Once this became possible, the purification procedure could be scaled up using protein stains with progressively higher detection limits to estimate the amount of protein kinase recovered. If this latter approach were undertaken, the γ-phosphate-linked ATP-Sepharose would be the obvious chromatography matrix to test first since the affinity of the 60 kDa protein kinase for this matrix is likely to be high and the development of the chromatography would probably be uncomplicated (an ATP gradient could be used to elute the protein kinase and if necessary the protein kinase could be pre-labelled). If the protein kinase proved not to

bind to this matrix, HIC or cation-exchange chromatography (which was not tested) would probably be the best choices.

In summary, the priorities in future work would be to fully develop and standardise the procedures for the anion-exchange and H-9 affinity chromatography and to obtain estimates of the purification achieved in each step (with PKI and PKII probably being purified separately) and then to perform experiments to estimate the amount of protein recovered to assess the feasibility of purifying sufficient quantities of the 60 kDa protein kinase for microsequencing or the raising of antibodies. Depending on the approach chosen the latter aim might be simplified by the development of at least one further purification step.

**Fig. 4.1** Absorbance trace and run conditions for the initial anion-exchange separation. Panel A shows the absorbance trace with absorbance (280 nm) shown on the vertical axis and time on the horizontal axis. Buffer changes are marked on the horizontal axis at the time they were introduced at the pump and the fractions that were analysed using the IEF gel *in situ* protein kinase assay are also indicated. Panel B shows the buffer conditions used. The flow rate was 0.8 ml/minute throughout the run and fractions were collected every half minute.

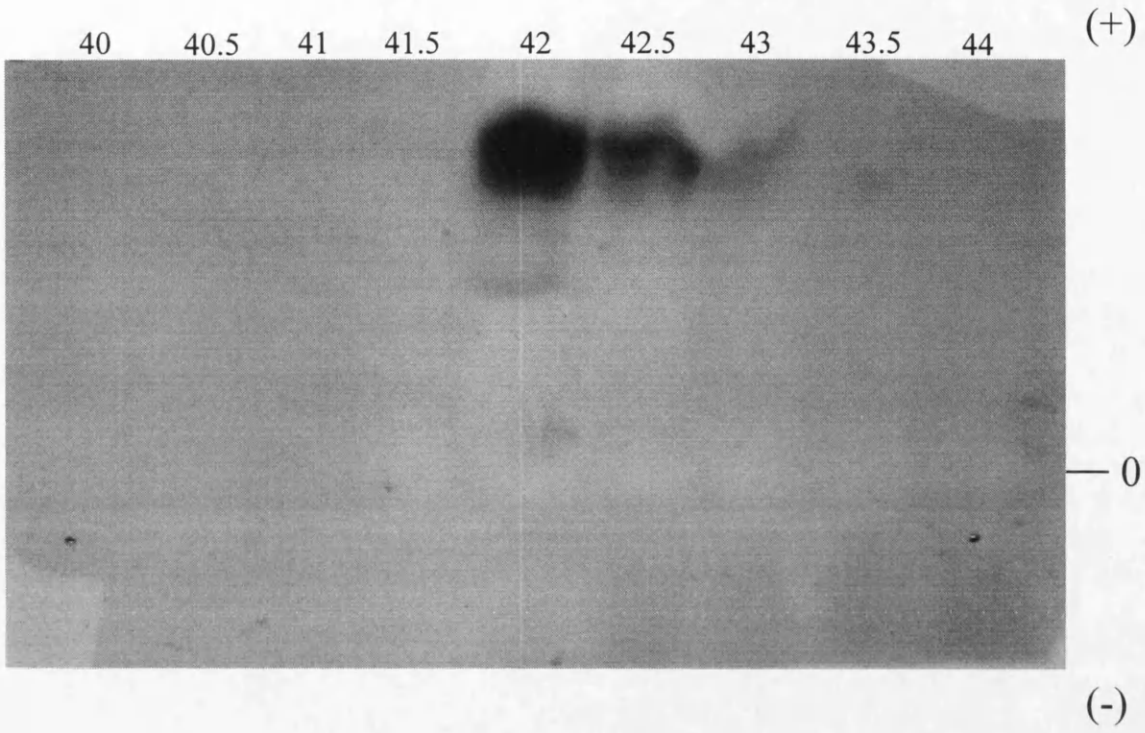



**B**

| Time/minutes | % 100 mM Tris-HCl | % 100 mM Tris-base | % 1 M NaCl | % $H_2O$ |
|---|---|---|---|---|
| 0 | 16 | 4 | 0 | 80 |
| 25 | 20 | 0 | 0 | 80 |
| 33 | 20 | 0 | 50 | 30 |

**Fig. 4.1**

**Fig. 4.2** Assay of fractions from the initial anion-exchange separation using the IEF gel *in situ* protein kinase assay. Panel A shows the autoradiograph of the IEF gel used in the assay of column fractions from the initial anion-exchange separation (Fig. 4.1). The fractions loaded are indicated above the gel lanes and the origin of electrophoresis (O) and the position of the anode (+) and cathode (-) are also indicated. In order to obtain molecular weight estimates for the phosphate labelled proteins detected in the IEF protein kinase assay, gel bands were cut out from the dried IEF gel and loaded onto an SDS-PAGE gel which was then dried and autoradiographed. Panel B (left panel) shows the autoradiographs of a gel lane containing radiolabelled protein molecular weight markers with the molecular weights (in kDa) indicated to the left. The right panel shows the autoradiograph of a gel lane loaded with the phosphate labelled protein from fraction 42.5. The labelled protein from fraction 42.5 (indicated by the arrow) is 60 kDa in size.




A

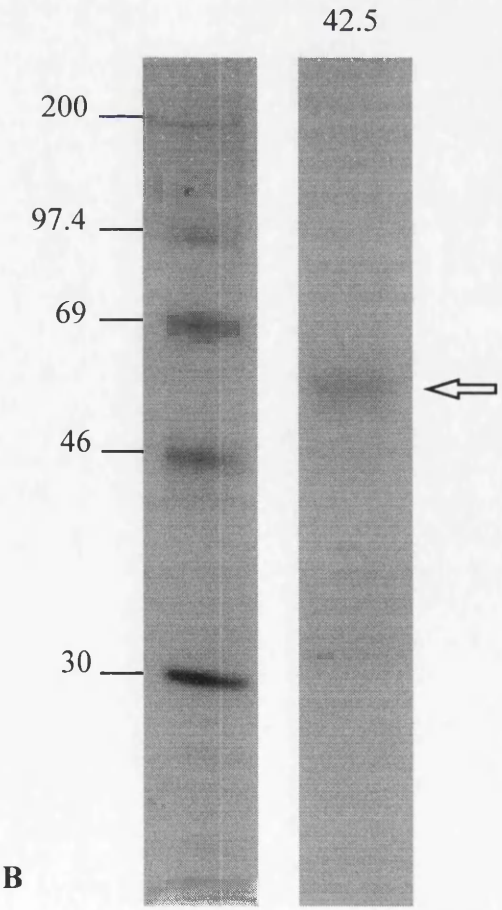
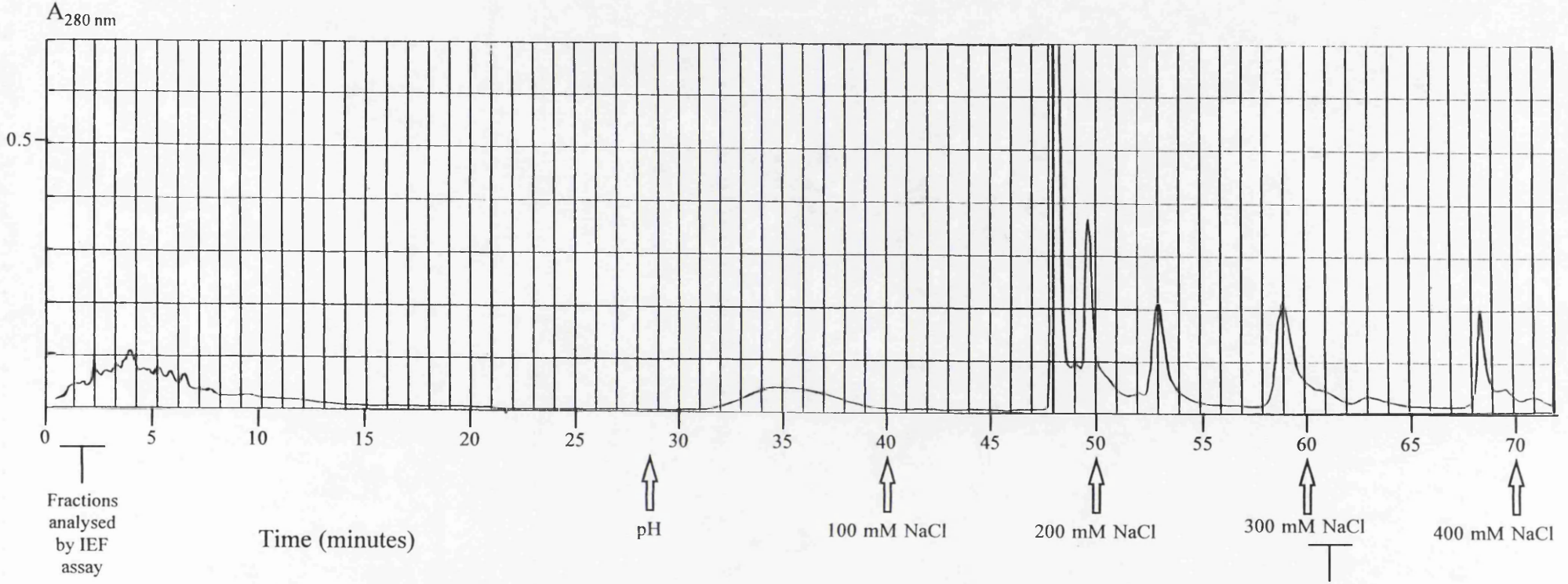



**Fig. 4.2** **B**

**Fig. 4.3** Elution of the protein kinase from the anion-exchange column by 100 mM NaCl steps. Panel A shows the absorbance trace from the second anion-exchange chromatography separation (100 mM NaCl steps). Absorbance (280 nm) is shown on the vertical axis and time on the horizontal axis. Buffer changes are marked on the horizontal axis at the time they were introduced at the pump, and the fractions that were analysed using the IEF gel *in situ* protein kinase assay are also indicated. Panel B shows the buffer conditions used. The flow rate was 0.8 ml/minute throughout the run and fractions were collected every half minute.

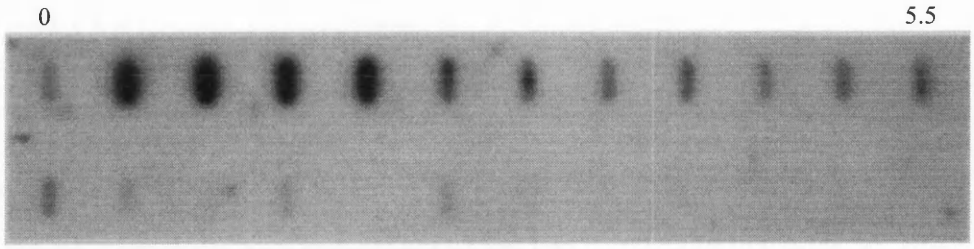


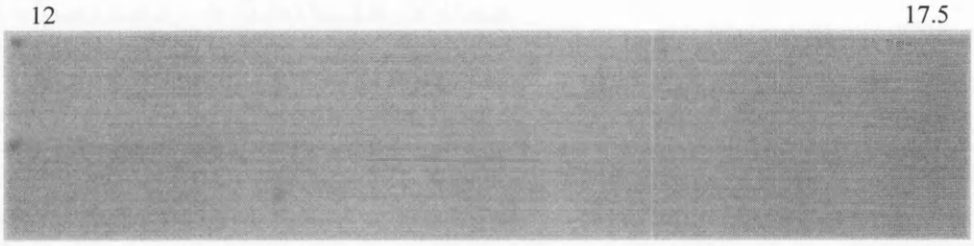

**A**

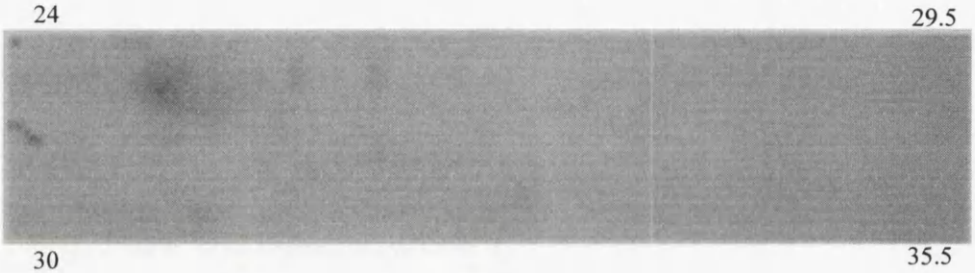
**B**

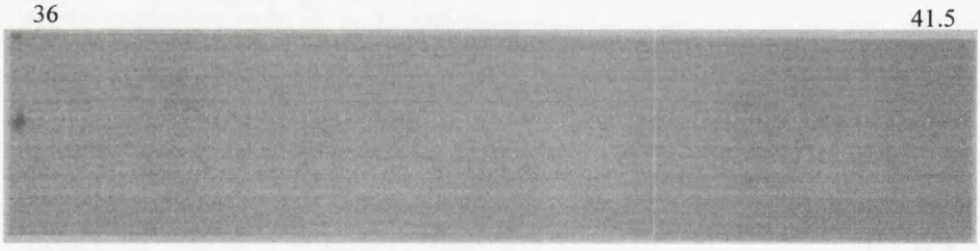
| Time/minutes | % 100 mM Tris-HCl | % 100 mM Tris-base | % 1 M NaCl | % $H_2O$ |
|---|---|---|---|---|
| 0 | 16 | 4 | 0 | 80 |
| 28 | 20 | 0 | 0 | 80 |
| 40 | 20 | 0 | 10 | 70 |
| 50 | 20 | 0 | 20 | 60 |
| 60 | 20 | 0 | 30 | 50 |
| 70 | 20 | 0 | 40 | 40 |

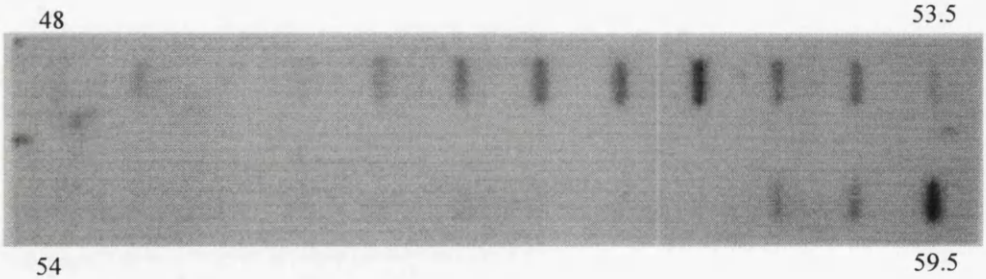
**Fig. 4.3**

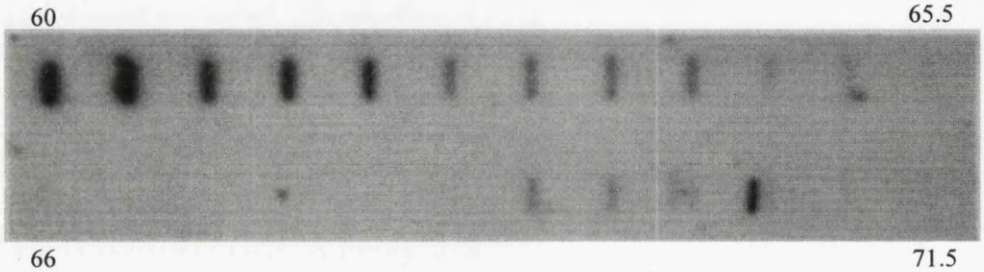

0
5.5
6
11.5


12
17.5
18
23.5


24
29.5
30
35.5


36
41.5
42
47.5


48
53.5
54
59.5


60
65.5
66
71.5

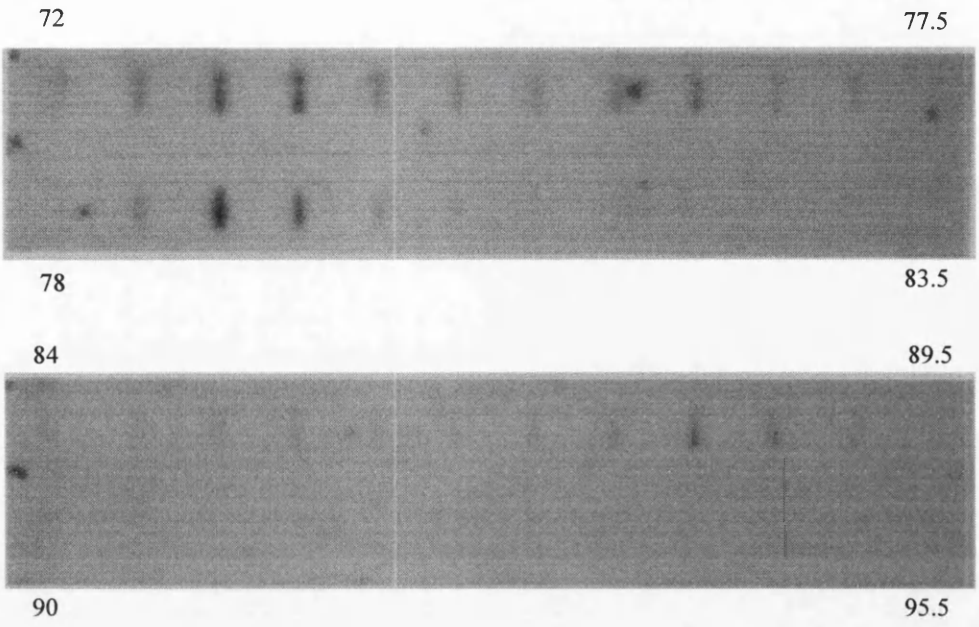
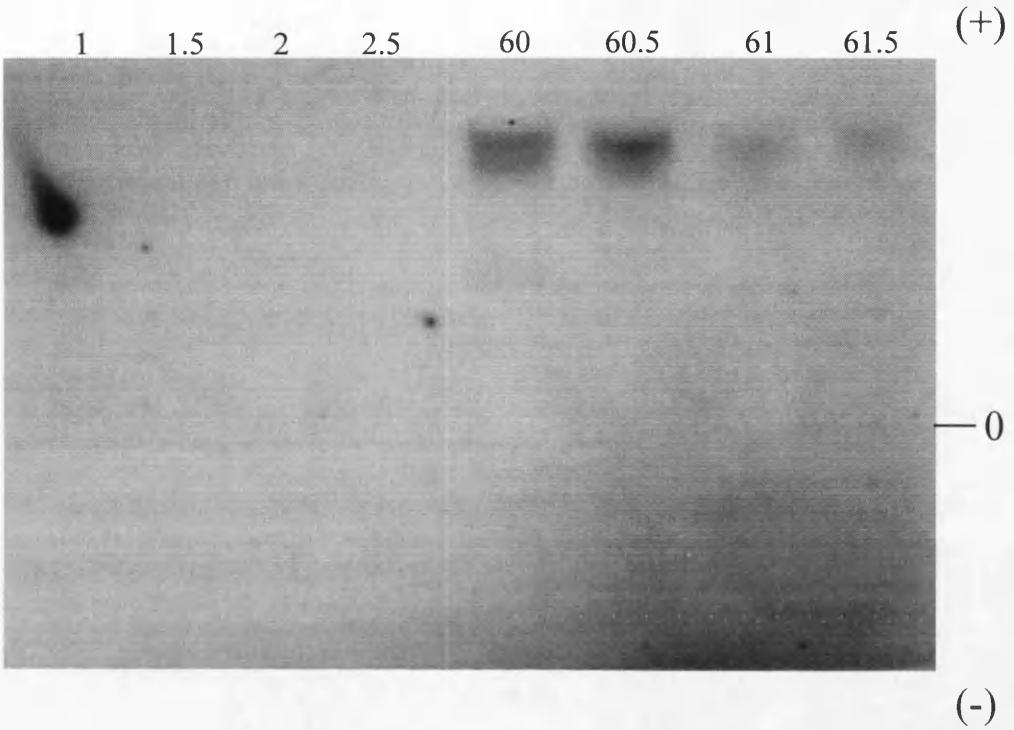



**Fig. 4.4** Assay of fractions from the 100 mM NaCl step separation (Fig. 4.3) using the slot blot assay. The figure shows the autoradiographs of the blots used in the assay. The fractions loaded in each slot are indicated alongside the autoradiographs. Fractions 1-2.5 and 60-61.5 were analysed using the IEF gel protein kinase assay (Fig. 4.5).

**Fig. 4.5** Assay of fractions from the 100 mM NaCl step elution (Fig. 4.3) using the IEF gel *in situ* protein kinase assay. Panel A shows the autoradiograph of an IEF gel used in the assay of column fractions from the second anion-exchange run (100 mM NaCl step elution). Two sets of fractions that showed high activity in the slot blot system were chosen for assay. The gel lanes are labelled above with the numbers of the fraction loaded and the origin of electrophoresis (O) and the position of the anode (+) and cathode (-) are also indicated. To obtain molecular weight estimates for the labelled proteins detected in the IEF assay, the gel bands containing the phosphate labelled protein were cut from the dried IEF gel and loaded onto an SDS-PAGE gel. Panel B shows the autoradiograph of part of the SDS-PAGE gel. The leftmost gel lane was loaded with radiolabelled protein molecular weight markers, the sizes of which (in kDa) are indicated to the left of the figure. The other two lanes shown were loaded with the phosphate labelled bands from fractions 60 and 60.5 (Fig. 4.3). The labelled protein in each lane is 60 kDa in size.




A

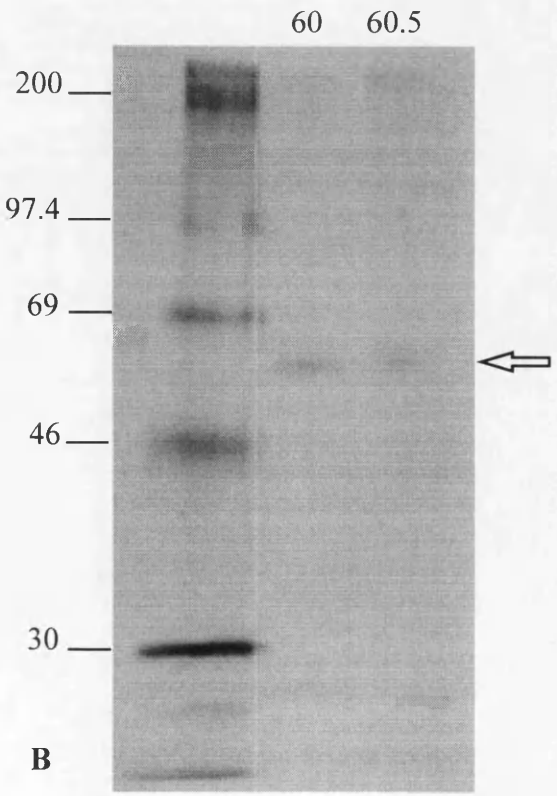
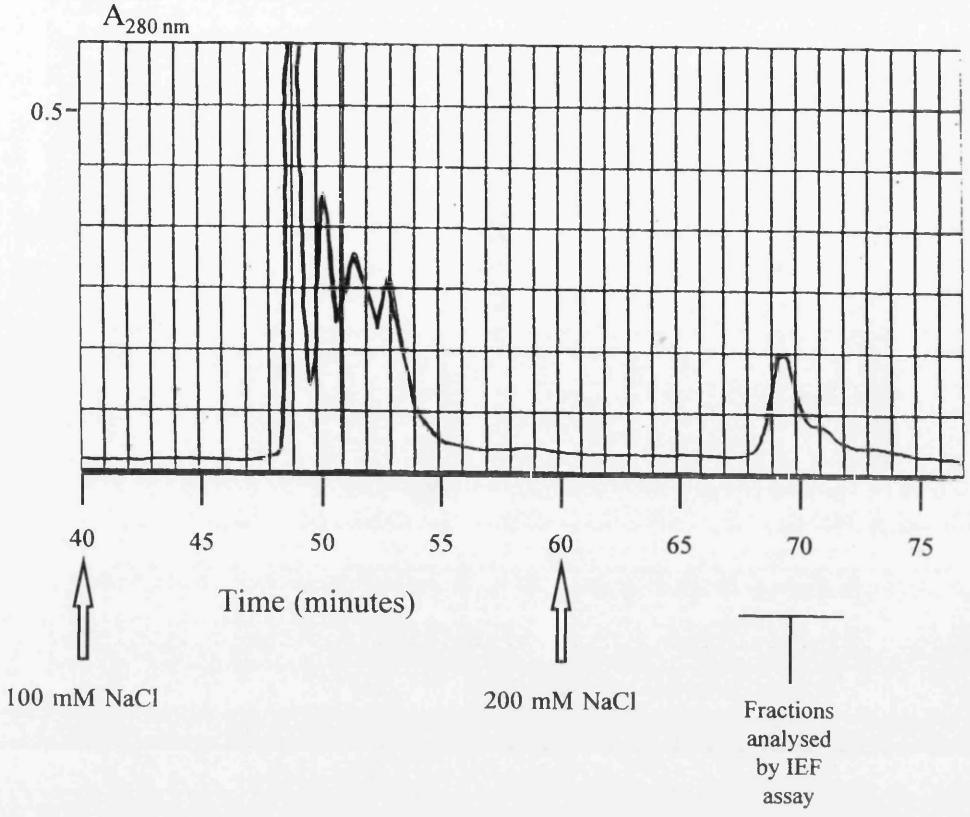



**Fig. 4.5** **B**

**Fig. 4.6** Results of a confirmatory anion-exchange separation using 100 mM NaCl steps for protein kinase elution. Panel A shows the absorbance trace from part of the run that includes the fractions eluting after the 200 mM NaCl step that were assayed for protein kinase activity. The buffer conditions used are shown in panel B. Panel C shows the autoradiograph of the blot of the IEF gel used to assay fractions 67.5-71 for the 60 kDa protein kinase. The leftmost lane was loaded with a total extract as a positive control.

A




B

| Time/minutes | % 100 mM Tris-HCl | % 100 mM Tris-base | % 1 M NaCl | % $H_2O$ |
|---|---|---|---|---|
| 0 | 16 | 4 | 0 | 80 |
| 28 | 20 | 0 | 0 | 80 |
| 40 | 20 | 0 | 10 | 70 |
| 60 | 20 | 0 | 20 | 60 |

C

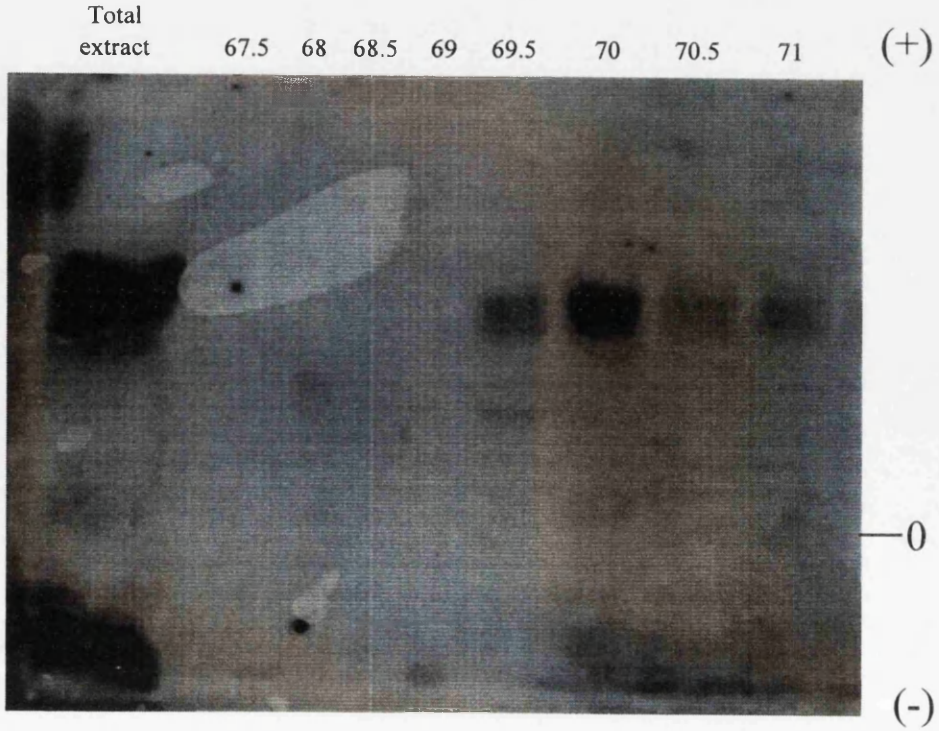
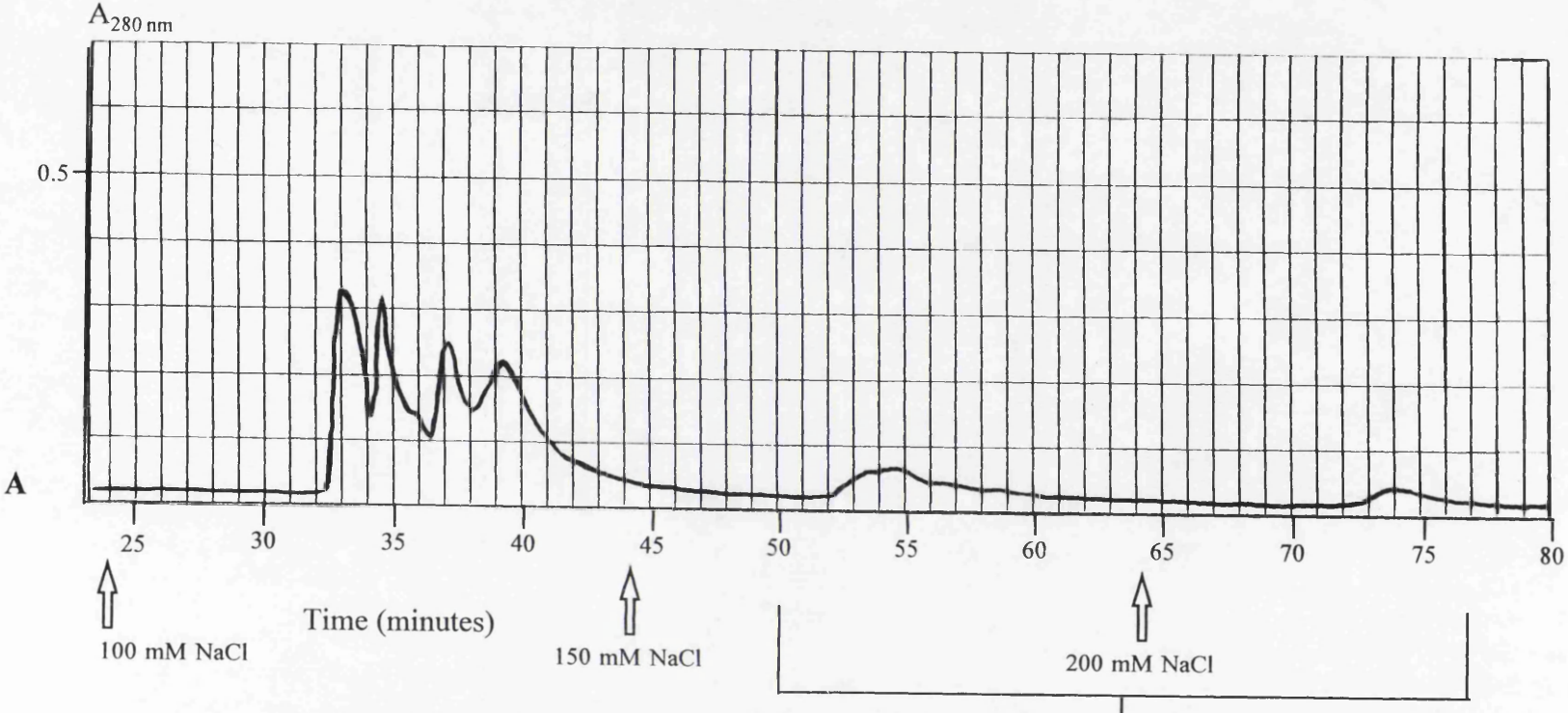



**Fig. 4.6**

**Fig. 4.7** Elution of the protein kinase from the anion-exchange column by 150 mM and 200 mM NaCl steps. Panel A shows the absorbance trace for an anion-exchange chromatography separation in which 150 mM and 200 mM NaCl steps were used to elute the 60 kDa protein kinase. Absorbance (280 nm) is shown on the vertical axis and time on the horizontal axis. Buffer changes are marked on the horizontal axis at the time they were introduced at the pump, and the fractions that were analysed using the liquid kinase assay (see text) are also indicated. Panel B shows the buffer conditions used. The flow rate was 0.8 ml/minute throughout the run and fractions were collected every half minute.




**B**

| Time/minutes | % 100 mM Tris-HCl | % 100 mM Tris-base | % 1 M NaCl | % $H_2O$ |
|---|---|---|---|---|
| 0 | 16 | 4 | 0 | 80 |
| 12 | 20 | 0 | 0 | 80 |
| 24 | 20 | 0 | 10 | 70 |
| 44 | 20 | 0 | 15 | 65 |
| 64 | 20 | 0 | 20 | 60 |
| 84 | 20 | 0 | 30 | 50 |

**Fig. 4.7**

**Fig. 4.8** Assay of column fractions from the 150 mM/200 mM NaCl step separation (Fig. 4.7) using the liquid kinase assay. The figure shows autoradiographs of SDS-PAGE gels used to analyse the liquid kinase assays of fractions 54.5-76.5. Fractions 50-54 were also assayed but the autoradiograph of the SDS-PAGE gel (which showed no phosphate labelling at any position in the gel) is not included in the figure. Radioactive protein molecular weight markers were run alongside the assay samples and the sizes (in kDa) are shown to the left of each row.

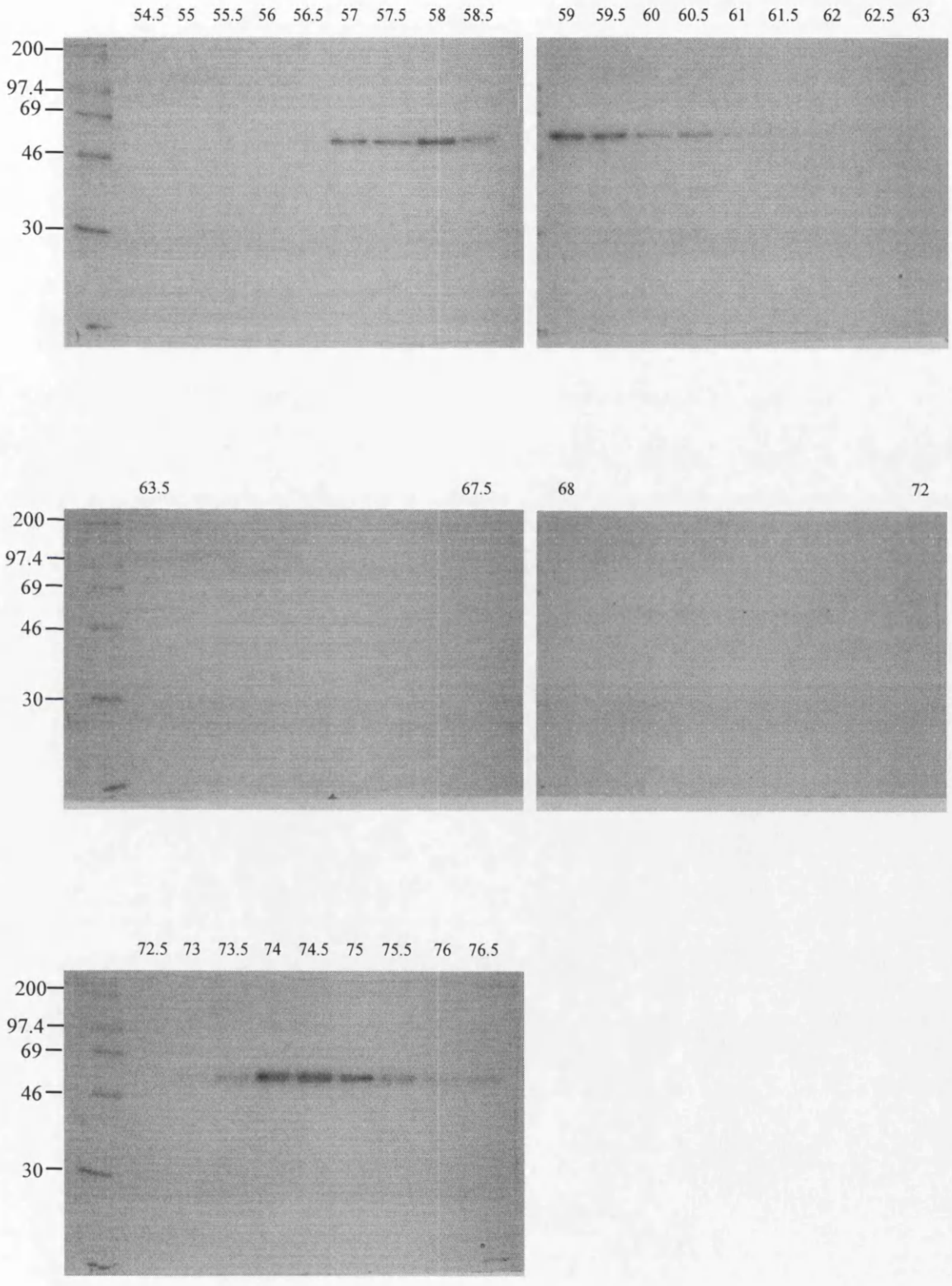
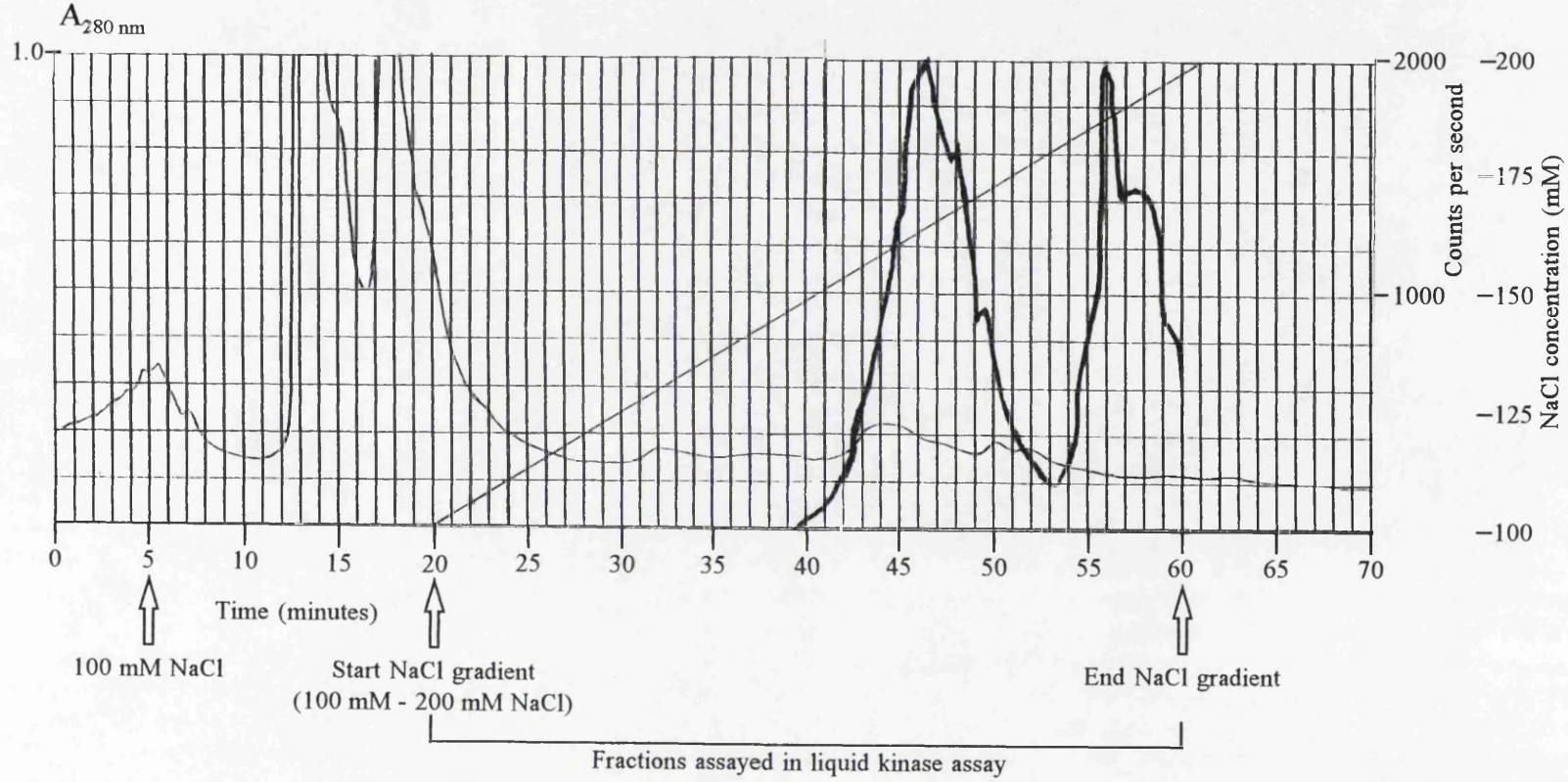

54.5 55 55.5 56 56.5 57 57.5 58 58.5
59 59.5 60 60.5 61 61.5 62 62.5 63
63.5 67.5
68 72
72.5 73 73.5 74 74.5 75 75.5 76 76.5
200
97.4
69
46
30

**Fig. 4.8**

**Fig. 4.9** Elution of the protein kinase from the anion-exchange column by a 100 mM-200 mM NaCl linear gradient. Panel A shows the absorbance trace for the run with absorbance (280 nm) shown on the vertical axis and time on the horizontal axis. Buffer changes are marked on the horizontal axis at the time they were introduced at the pump with the gradient indicated by a line superimposed on the trace. The bolder line shows the radioactive counts associated with the 60 kDa protein kinase determined by scintillation counting of SDS-PAGE gel slices from gels used in the analysis of the liquid kinase assays of the column fractions. The results are expressed as counts per second per fraction and the scale is indicated to the right of the trace. Panel B shows the buffer conditions used. The flow rate was 0.8 ml/minute throughout the run and fractions were collected every half minute.

A




B

| Time/minutes | % 100 mM Tris-HCl | % 100 mM Tris-base | % 1 M NaCl | % $H_2O$ |
|---|---|---|---|---|
| 0 | 20 | 0 | 0 | 80 |
| 5 | 20 | 0 | 10 | 70 |
| 20 | 20 | 0 | 10 | 70 |
| Linear gradient 100 mM NaCl - 200 mM NaCl between 20 and 60 minutes | | | | |
| 60 | 20 | 0 | 20 | 60 |

**Fig. 4.9**

**Fig. 4.10** Activity of the 60 kDa protein kinase in buffers of different NaCl concentration. The figure shows the autoradiograph of an SDS-PAGE gel used in the analysis of liquid kinase assays for the 60 kDa protein kinase in which NaCl was added to the labelling reactions at different concentrations. The sample material was concentrated anion-exchange chromatography fractions containing the PKII activity eluted by a 100-200 mM NaCl gradient (Section 4.2.1.4). The final NaCl concentrations in the different labelling reactions are indicated above the SDS-PAGE gel lanes and the positions and sizes (in kDa) of protein molecular weight markers run alongside the samples are indicated to the left of the figure.

**Fig. 4.11** Binding of pre-labelled 60 kDa protein kinase to the dodecyl agarose column. Panel A shows a silver-stained SDS-PAGE gel loaded with aliquots of consecutive fractions eluting in the starting buffer. The autoradiograph of the gel showed no phosphate labelling of any protein. Panel B shows two gel lanes from a silver-stained SDS-PAGE gel loaded with portions of agarose from the top of the dodecyl agarose column (left), and the corresponding autoradiographs (right). The positions and sizes (in kDa) of protein molecular weight markers run are indicated to the left of panels A and B.

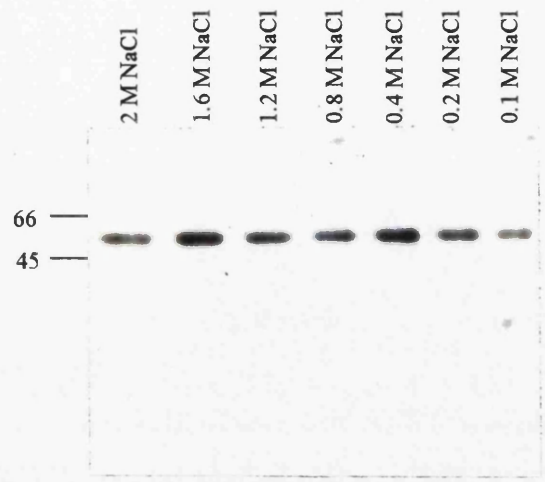



**Fig. 4.10**

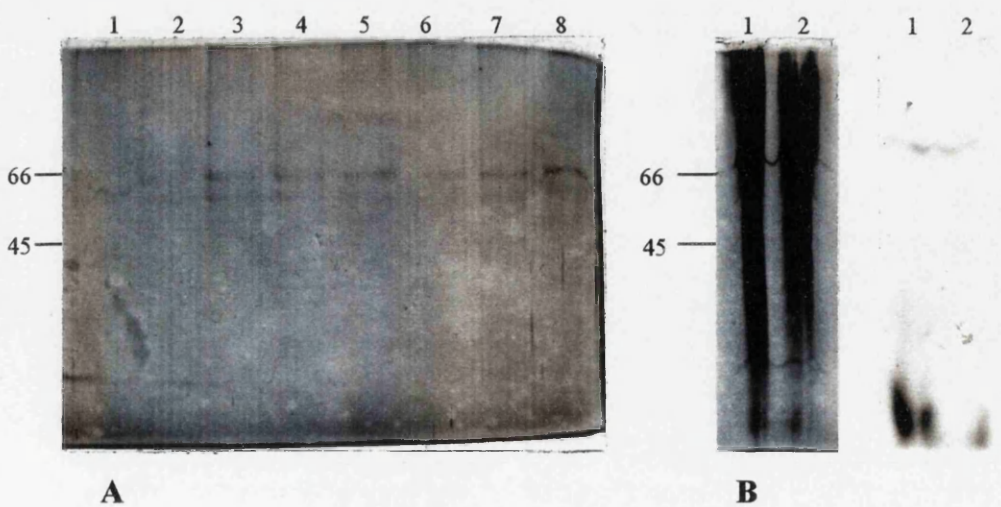



**Fig. 4.11**

**Fig. 4.12** Binding of pre-labelled 60 kDa protein kinase to the hexyl agarose column. Panel A shows silver-stained SDS-PAGE gels loaded with aliquots from consecutive fractions eluting from the hexyl HIC column in starting buffer. The leftmost lane contains an aliquot of the sample that was applied to the column. Panel B shows the autoradiographs of the gels. The positions and sizes (in kDa) of protein molecular weight markers run alongside the samples are indicated to the left of the figure.

**Fig. 4.13** Binding of pre-labelled 60 kDa protein kinase to the octyl agarose column. Panel A shows a silver-stained SDS-PAGE gel loaded with consecutive fractions eluting from the octyl HIC column in starting buffer (10 mM Tris-HCl pH 7.2, 0.5 M NaCl). No labelled protein bands were detected by autoradiography of this gel. Panel B shows an SDS-PAGE loaded with consecutive fractions eluting after the buffer was changed to 10 mM Tris-HCl (pH 7.2). Autoradiography of the gel again showed no labelled bands. The left block of Panel C shows a silver stain of an SDS-PAGE gel lane loaded with agarose removed from the top of the octyl HIC column. The autoradiograph of the gel lane is shown to the right. The positions and sizes (in kDa) of protein molecular weight markers are indicated on the figure.

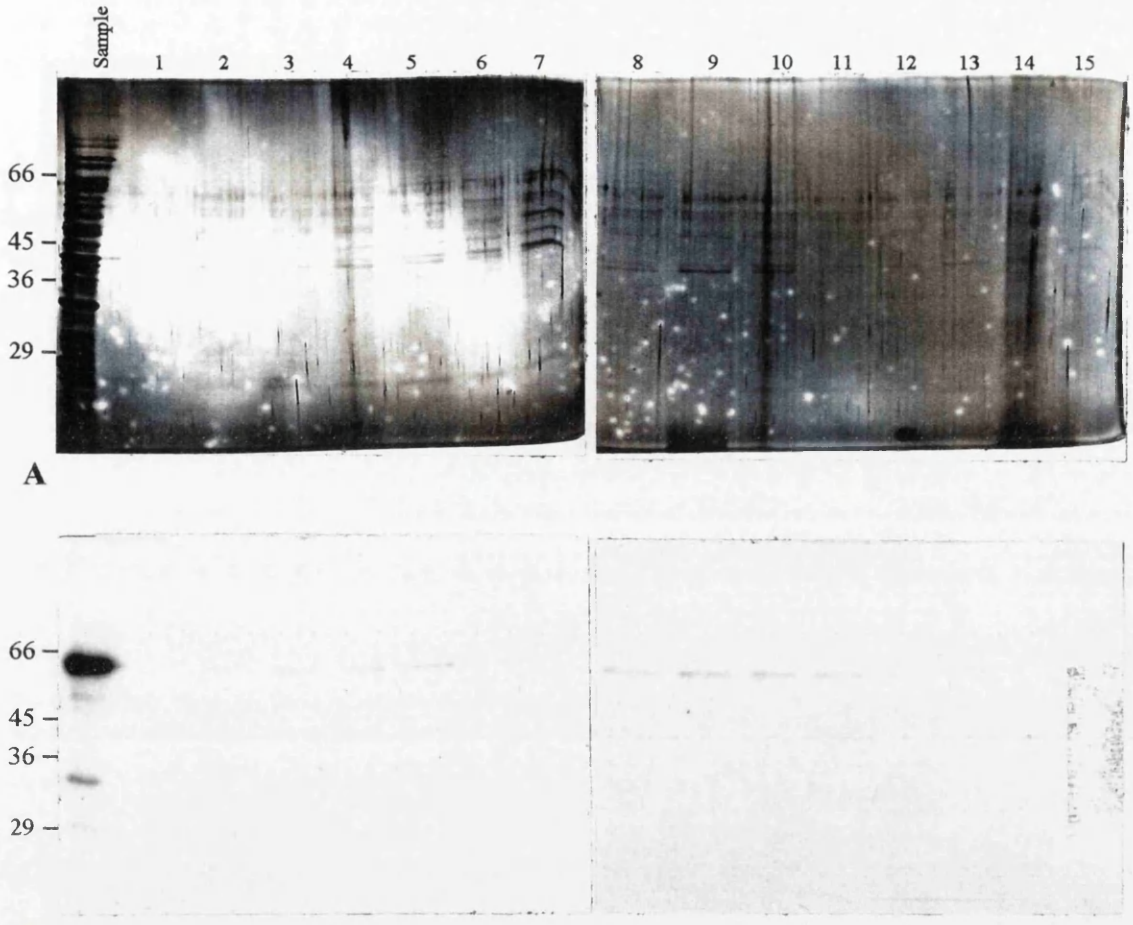



**Fig. 4.12**

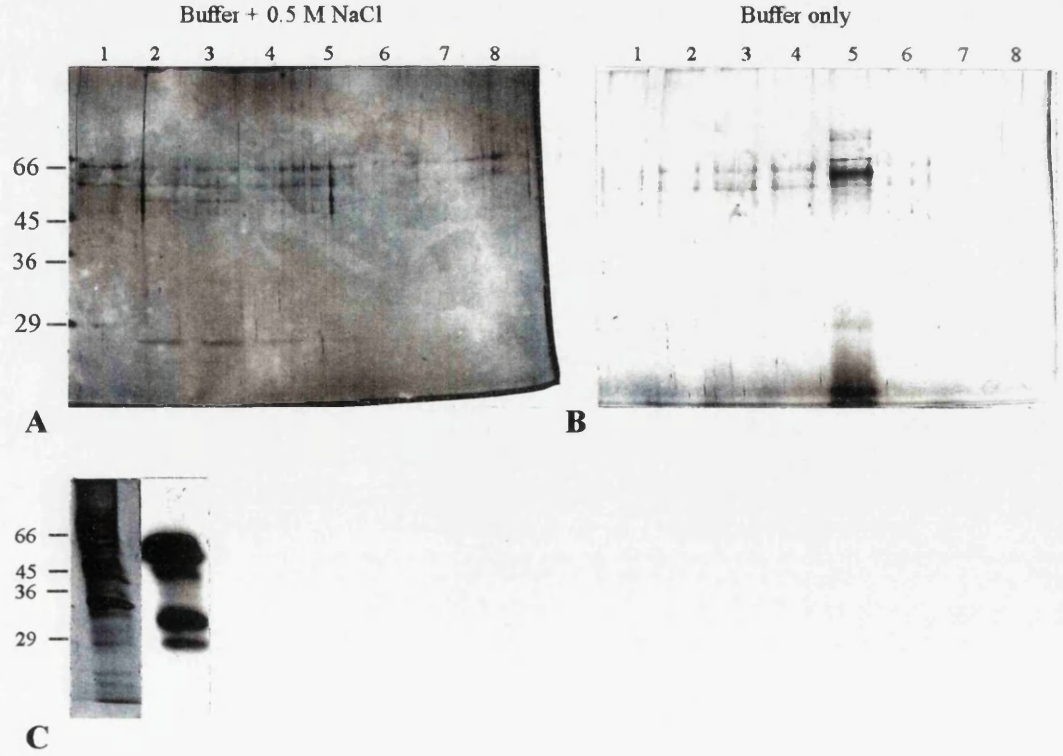
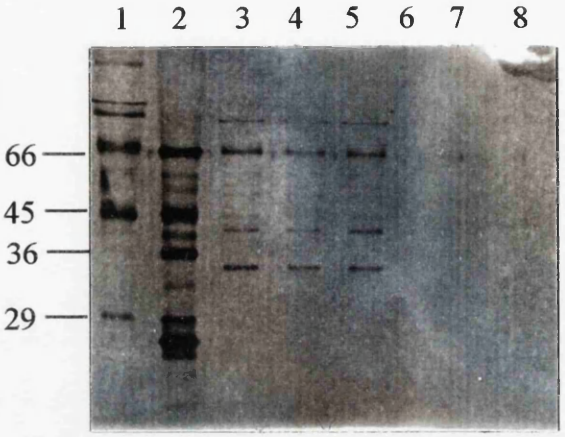
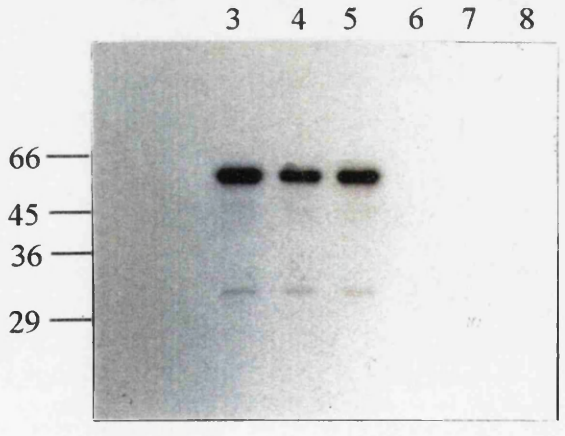



**Fig. 4.13**

**Fig. 4.14** Binding of the 60 kDa protein kinase to ATP agarose. Panel A shows a silver-stained SDS-PAGE gel used to analyse liquid kinase assays performed on supernatants from the binding and elution steps. Lanes 3, 4, and 5 contain the reactions performed on supernatants from binding incubations performed in buffers containing no NaCl, 10 mM NaCl, and 50 mM NaCl respectively while lanes 6, 7, and 8 contain the corresponding kinase reactions performed on the supernatants from the elution step. Panel B shows the autoradiograph of the gel. Lanes 1 and 2 contain protein molecular weight markers, the positions and sizes of which are indicated in panels A and B.




A




B

**Fig. 4.14**

**Fig. 4.15** Binding of pre-labelled 60 kDa protein kinase to H-9 Sepharose. In each block the left section shows a silver-stained SDS-PAGE gel lane and the right section shows the corresponding autoradiograph. Protein molecular weight markers are shown to the left in each panel. While it is not apparent in all of the photographs of the autoradiographs, examination of the originals shows clearly that in each case the strong band at 60 kDa is a doublet.

Panel A shows the SDS-PAGE analysis of material eluted from H-9 Sepharose by boiling in SDS-PAGE sample buffer after the Sepharose had been incubated with partially purified pre-labelled protein kinase in buffer and then washed five times in the same buffer. It can be seen that the 60 kDa protein kinase was readily detectable in the supernatant of the boiled Sepharose.

Panel B shows the SDS-PAGE analysis of an aliquot from a separate batch of pre-labelled protein kinase used in an experiment to compare kinase binding to H-9 Sepharose and control Sepharose (panels C, D and E). Panel C shows the SDS-PAGE analysis of supernatants from incubations of the pre-labelled protein kinase with H-9 Sepharose and control Sepharose in which the buffer was 30 mM Tris-HCl pH 7.0, 25 mM NaCl (first two blocks) while the latter two blocks show the analysis of the first washes (for which the buffer was 25 mM Tris-HCl pH 7.0). Following further washes the H-9 Sepharose was divided into two equal portions. One portion was incubated in buffer + 1% Triton-X100 and then in buffer + 2 mM ATP. The left block of panel D shows the SDS-PAGE analysis of the two supernatants. The 60 kDa protein kinase is clearly detectable in the ATP eluate but is barely visible in the Triton-X100 eluate. A larger volume of the ATP eluate (5x the volume) was also analysed in a large format SDS-PAGE gel (right block of panel D). Panel E shows the silver-stained gel lanes and autoradiographs for an equivalent analysis of the other portion of H-9 Sepharose in which incubation in 100 mM NaCl was used to elute the 60 kDa protein kinase.

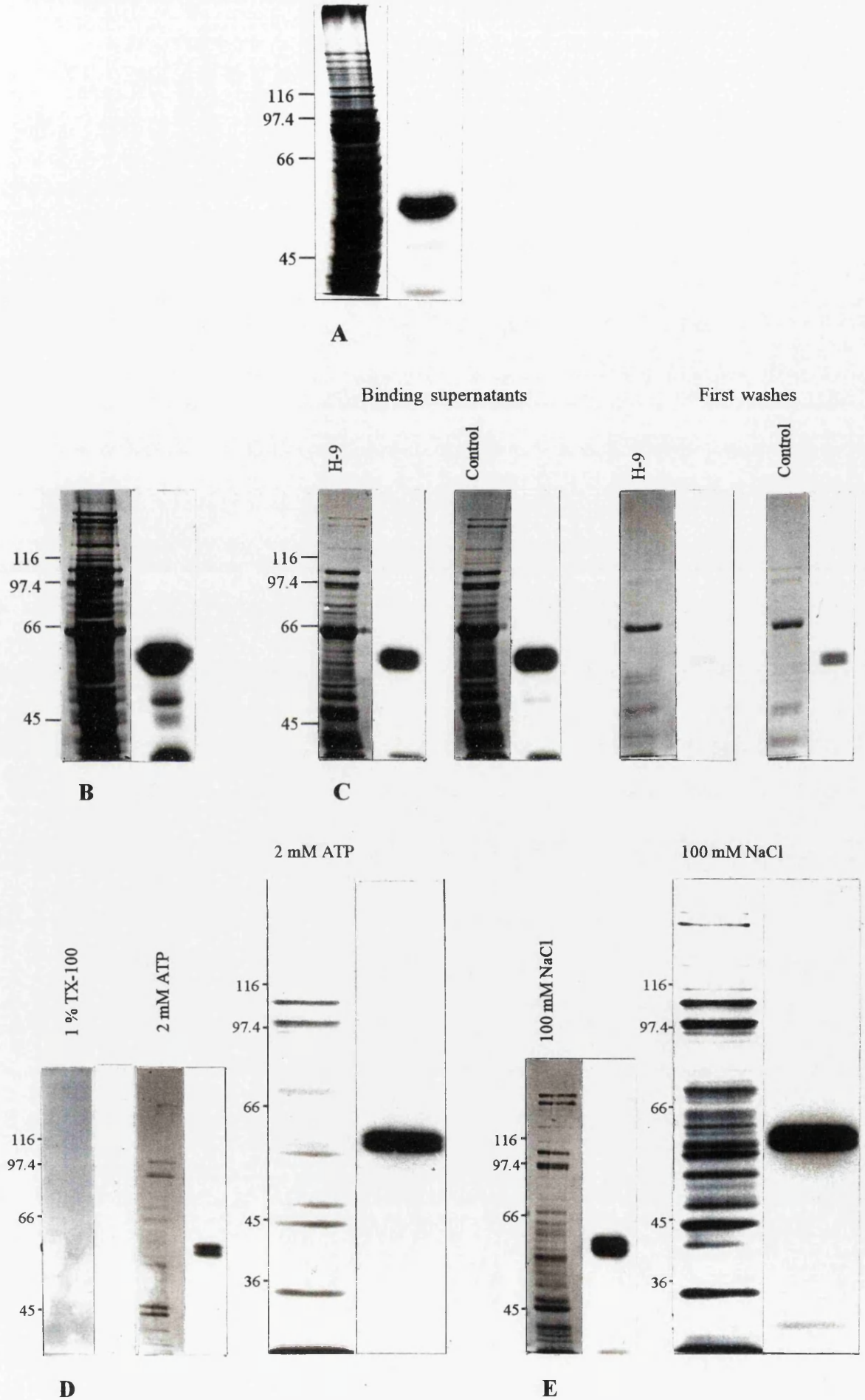
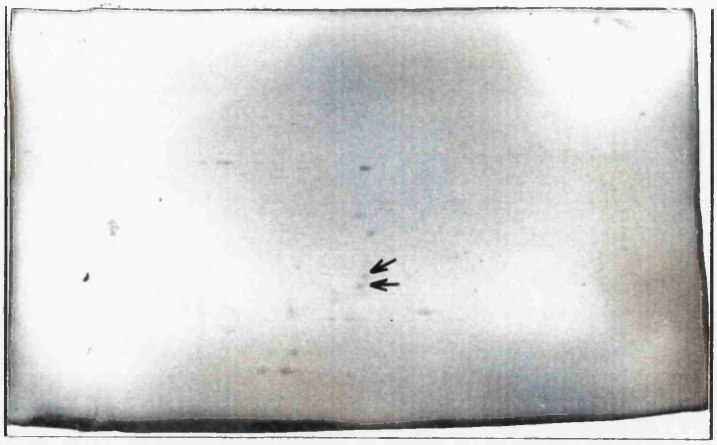

116
97.4
66
45
A
Binding supernatants
First washes
H-9
Control
H-9
Control
116
97.4
66
45
B
C
2 mM ATP
100 mM NaCl
1 % TX-100
2 mM ATP
100 mM NaCl
116
97.4
66
45
36
D
E

**Fig. 4.15**

**Fig. 4.16** Two-dimensional gel electrophoresis of protein eluted from H-9 Sepharose. Portions of the 2 mM ATP and 100 mM NaCl eluates shown in Fig. 4.15 panels C, E and D, F respectively were analysed by two-dimensional electrophoresis. Panel A shows the silver-stained second dimension gel used in the analysis of the NaCl eluate (first wash) while panel B shows its autoradiograph. The autoradiograph (not shown) of the gel used to analyse the ATP eluate showed the same two spots as seen in panel B (these were weaker in intensity compared to those from the NaCl eluate as seen in Fig. 4.15) but no protein was detected by silver-staining at any position in the gel. The two spots in panel B correspond exactly in position with the two silver-stained spots (arrowed) in panel A in the vertical dimension (SDS-PAGE), but do not quite match in position in the horizontal dimension (IEF). The lower molecular weight labelled spot (panel B) is closer in the horizontal dimension to the silver-stained spot at the same vertical position than the higher molecular weight labelled spot is to the higher molecular weight silver-stained spot.

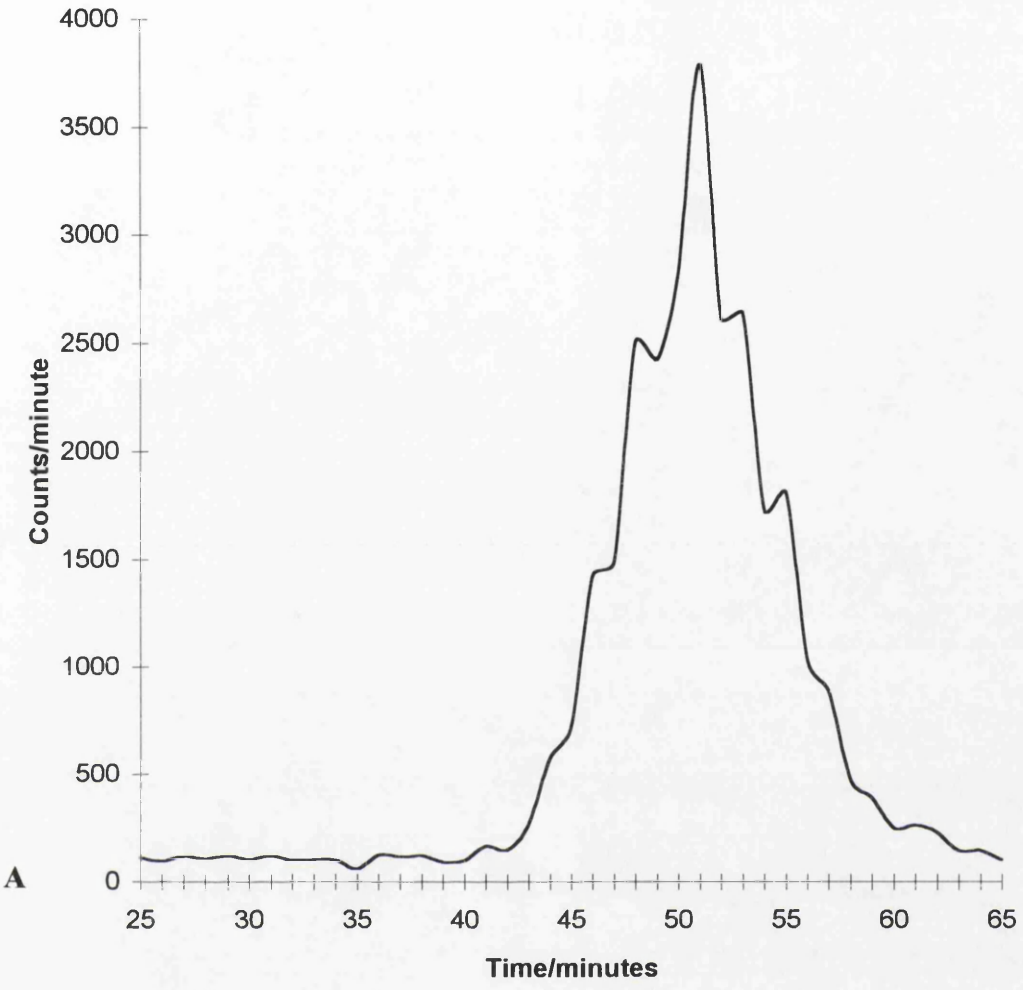


A

B

**Fig. 4.16**

**Fig. 4.17** Elution of pre-labelled 60 kDa protein kinase from the H-9 affinity column by a 0-10 mM ATP gradient. Panel A shows scintillation counts taken on whole fractions prior to analysis by SDS-PAGE (Fig. 4.18). Radioactive counts are shown on the vertical axis while time is shown on the horizontal. Panel B shows the run conditions used. Fractions were collected every minute throughout the experiment.




A

B

| Time/minutes | Flow ml/minute | % 25 mM Tris-HCl (pH 7.0) | % 25 mM Tris-HCl, 10 mM ATP (pH 7.0) | % 25 mM Tris-HCl (pH 7.0), 1 M NaCl |
|---|---|---|---|---|
| 0 | 0.5 | 100 | 0 | 0 |
| 21 | 0.4 | 100 | 0 | 0 |
| 23 | 0.3 | 100 | 0 | 0 |
| 24 | 0.2 | 100 | 0 | 0 |
| 25 | 0.2 | 100 | 0 | 0 |
| | Linear gradient 0-10 mM ATP between 25 and 45 minutes | | | |
| 45 | 0.2 | 0 | 100 | 0 |
| 55 | 0.2 | 0 | 0 | 100 |

**Fig. 4.17**

**Fig. 4.18** SDS-PAGE analysis of fractions collected from the H-9 affinity column during application of a 0-10 mM ATP gradient (Fig. 4.17). Panels A and C show silver-stained SDS-PAGE gels while panels B and D show the corresponding autoradiographs. The sizes of protein molecular weight markers are indicated to the left. Fractions containing >1000 cpm were analysed and the fraction numbers are indicated above each gel lane. Fractions 3, 4, and 5 (top left) were collected shortly after sample application under the starting buffer conditions, while fractions 45-57 were collected while the column was being developed with the ATP gradient (Fig. 4.17).

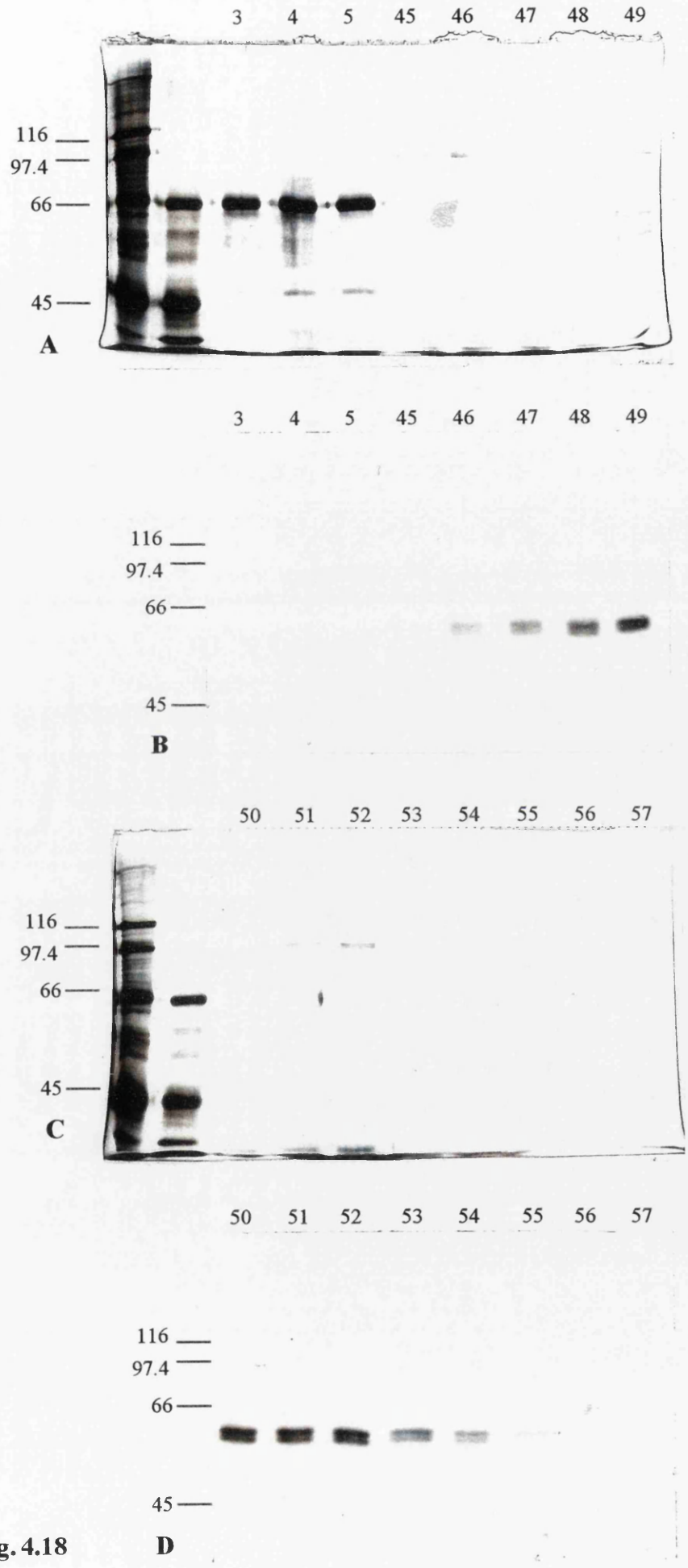
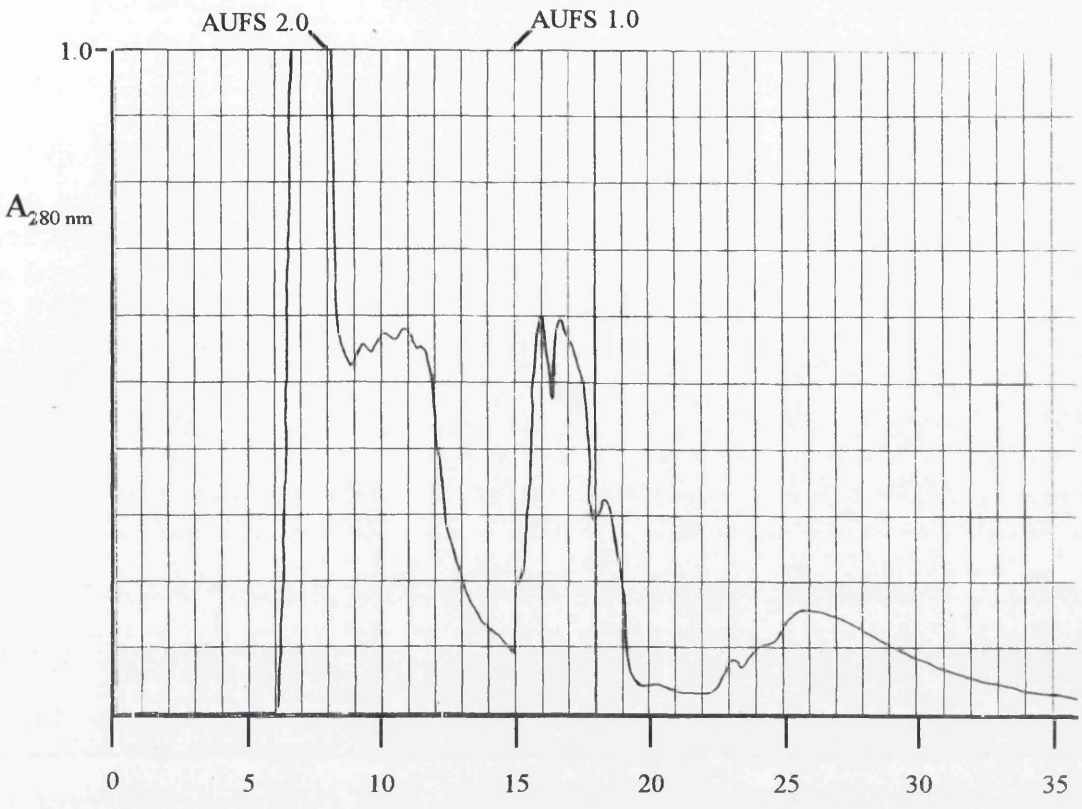
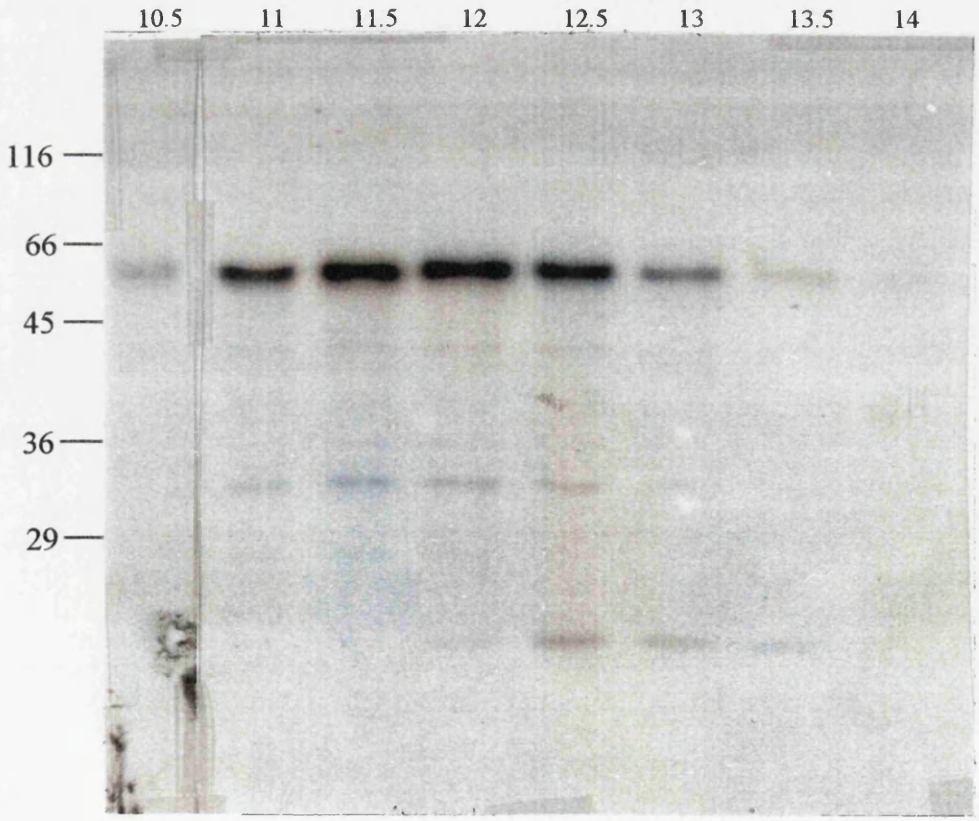

3 4 5 45 46 47 48 49
116
97.4
66
45
A
3 4 5 45 46 47 48 49
116
97.4
66
45
B
50 51 52 53 54 55 56 57
116
97.4
66
45
C
50 51 52 53 54 55 56 57
116
97.4
66
45
D

**Fig. 4.18**

**Fig. 4.19** Size-exclusion chromatography of the 100 000 g supernatant of the extract of *T. brucei* bloodstream forms assayed for the 60 kDa protein kinase. Panel A shows the absorbance trace from the size-exclusion chromatography experiment with absorbance (280 nm) shown on the vertical axis and time on the horizontal axis. The sample was loaded at time 0 and the mobile phase throughout the run was 20 mM Tris-HCl pH 5.7, 200 mM NaCl. The flow rate was 0.8 ml/minute and fractions were collected every half minute. Panel B shows the autoradiograph of an SDS-PAGE gel used to analyse liquid kinase assays of column fractions 10.5-14 (fraction numbers are indicated above the gel lanes). The positions and sizes of molecular weight markers are indicated the left of the figure.


AUFS 2.0
AUFS 1.0
1.0
$A_{280\,nm}$
0
5
10
15
20
25
30
35
Time (minutes)

A


10.5
11
11.5
12
12.5
13
13.5
14
116
66
45
36
29

B

**Fig. 4.19**

# Chapter Five

## Cloning of a *Trypanosoma brucei* cDNA encoding a protein kinase

## 5.1 Introduction

As an alternative strategy to the biochemical study of the 60 kDa autophosphorylating protein kinase for characterising protein kinases in *T. brucei*, PCR amplification of protein kinase cDNAs could be carried out using degenerate oligonucleotide primer pools designed from highly conserved peptide sequences from the catalytic domain. A number of potentially interesting partial sequences had been previously described from a study of this type (Hua and Wang, 1994). These sequences were examined to see if any of them were likely to be derived from genes for members of the protein kinase C family since the cloning of *T. brucei* genes for members of this protein kinase family would complement previous biochemical investigations of protein kinase C like activities (Keith *et al.*, 1990). The predicted peptide sequences from two of the amplified cDNA fragments (TbPK-A2 and TbPK-B1) were found to be similar to catalytic domain sequences of members of the protein kinase C family in database searches performed by Hua and Wang (1994). The predicted peptide sequence from TbPK-B1 showed 47% and 42% identity to catalytic domain sequences from a *Drosophila* protein kinase C (Schaeffer *et al.*, 1989) and the *Aspergillus* protein kinase NimA, respectively (Osmani *et al.*, 1988), while that from TbPK-A2 showed 56% and 58% identity to catalytic domain sequences from a rat protein kinase C (Ono *et al.*, 1988a+b) and the *Drosophila* polo kinase (Llamazares *et al.*, 1991). Based on these data TbPK-A2 (a product of primers based on catalytic subdomains VIb and IX) was chosen for further investigation. The similar level of sequence identity of this fragment to two protein kinases, only one of which is from the protein kinase C family, did not allow a firm prediction as to whether or not the short TbPK-A2 sequence was from a gene encoding a member of the PKC family. Two approaches were used to extend the TbPK-A2 sequence. The first of these was RACE (rapid amplification of cDNA ends) which is explained briefly in the next paragraph, while the second was to screen a λ cDNA library with a cloned PCR product similar to TbPK-A2. This was performed using standard procedures which are described in Chapter 2.

The basic strategy for 3′ RACE is shown in Fig. 5.1. First-strand cDNA synthesis is carried out using an oligo(dT) primer that anneals to the poly(A) tail of mature mRNAs. The oligo(dT) portion of the primer is tagged at its 5′ end by a sequence that is more suitable as a PCR primer. Such cDNA synthesis primers are sometimes called

adapter primers. Following cDNA synthesis PCR amplification is carried out using a sense gene-specific primer and a primer (the anchor primer) corresponding to the 5′ extension of the adapter primer. The specificity for the target cDNA in this method comes only from the gene-specific primer so it is often necessary to use a "nested" gene-specific primer in a second amplification using a small amount of the products of the initial reaction as the PCR template. To amplify 5′ cDNA ends from most organisms, 5′ RACE is carried out by engineering the addition of homopolymer tails to the pool of cDNAs using a terminal deoxynucleotide transferase. PCR of these tailed cDNAs is performed between antisense gene-specific primers and 5′ adapter primer complementary at its 3′ end to the homopolymer tail but with a 5′ anchor sequence. However, *T. brucei* mature mRNAs have a spliced leader sequence at their 5′ ends and, as an alternative to tailing of cDNAs, a primer based on the spliced leader can be used in combination with a gene-specific primer. As well as being technically easier than conventional 5′ RACE, the use of the spliced leader primer introduces a selection for amplification only of full length cDNAs derived from *T. brucei* mRNAs. However, the sequence of the *T. brucei* spliced leader has a considerably lower G+C content than is ideal for PCR. Fig. 5.2 shows the strategy for 5′ RACE using the spliced leader.

## 5.2 Results

### 5.2.1 Amplification of a cDNA fragment homologous to TbPK-A2

Non-degenerate gene specific primers were synthesised using sequence information from TbPK-A2. In order to simplify explanation of the PCR reactions carried out, the primers used have been given numbers. Fig. 5.3 shows the primer positions and their sequences. The 5′ ends of the two outermost primers 2 and 5 overlap the original degenerate primers by 7 and 5 bases respectively.

A PCR reaction, using primers 2 and 5 with a first strand cDNA template from bloodstream *T. brucei*, was performed to confirm the presence of TbPK-A2 in stock 869 and to generate a probe for library screens. Fig. 5.4 panel A shows an agarose gel containing the products of 30 cycles of PCR (annealing temperature 48 °C, extension time 1 minute) with primers 2 and 5 (lane 2) and DNA size markers (lane 1). The product band in lane 2 is of the size expected (171 bp) for an amplified cDNA fragment similar to TbPK-A2. To confirm the similarity between this fragment and that reported

by Hua and Wang (1994), DNA from the gel band was used as the template for PCR reactions with the different possible combinations of gene-specific primers. Fig. 5.4 panel B shows a gel containing aliquots of these reactions. All reactions gave single products of the expected sizes. The product of primers 2 and 5 was then cloned and sequenced completely in both strands. It was found to be identical in sequence to TbPK-A2 across the region of overlap except for a single base pair difference which led to a difference in the predicted peptide sequence.

### 5.2.2 Amplification of the 5′ cDNA sequence

To extend the sequence in the 5′ direction, *T. brucei* first-strand cDNA was amplified using primer 1 (spliced leader) and primer 5 (gene-specific). Fig. 5.5 panel A shows an agarose gel lane containing the products of 30 cycles of amplification (annealing temperature 47 °C, extension time 2 minutes) with these primers. The sizes of DNA markers are described to the left. Identical products were obtained with annealing temperatures of 38 °C and 50 °C. Individual DNA bands from the initial 5′ reaction were then re-amplified from gel slices. Since the desired 5′ products would overlap with the entire length of the cloned PCR product of primers 2 and 5 this fragment was used to probe a Southern blot of the second round 5′ RACE reactions. Fig. 5.5 panel B shows the gel used to separate the reaction products while Fig. 5.5 panel C shows the autoradiograph of the hybridised blot. The strongly hybridising band (lane 3) is from the DNA that was used to make the probe, included as a positive control. The high level of hybridisation suggests that the probe was labelled efficiently and was specific as a probe, although a rather large area of the blot was labelled due to overloading of the gel. Lanes 6, 8, and 10 contain products of PCR reactions using the spliced leader primer and gene-specific primer 5. Several of the bands hybridised to the probe but the signal was weak compared to that from the positive control. The intensity of the ethidium bromide staining of the test gel bands compared to that of the positive control band shows that the lower degree of hybridisation cannot be accounted for by lower quantities of DNA in the test samples if the hybridisation is to the visible gel bands. The most likely explanation of the hybridisation is that it is either entirely non-specific or is due only to hybridisation to the primers rather than to internal sequences. Therefore, no products likely to contain the 5′ end of the gene were obtained by 5′ RACE.

### 5.2.3 Amplification of the 3′ cDNA sequence

PCR amplification was carried out on *T. brucei* first-strand cDNA using primer 2 (gene-specific) and an oligo(dT) primer. Fig. 5.6 (panel A) shows an agarose gel used to analyse the products obtained. Lane 1 contains DNA size markers, lane 2 contains the products of a positive control reaction while lane 3 contains the products of the 3′ RACE reaction. It can be seen in lane 3 that no specific bands were obtained in the RACE reaction while the positive control template was successfully amplified (lane 2). A possible reason for the failure of the 3′ reaction to amplify products was that the lower annealing temperature used with an oligo(dT) primer was reducing the specificity of the reaction. To address this problem a new batch of cDNA was synthesised using an adapter primer so that the corresponding anchor primer could be used for PCR. Fig. 5.6 panel B depicts a gel containing products of the 3′ reaction using gene-specific primer 3 and an anchor primer (primer 6) in 33 cycles of PCR with annealing temperatures of 56 °C, 58 °C, and 60 °C (indicated on figure) and a 3 minute extension time. Two product bands were cloned from the 56 °C reaction (marked on the figure). Fig. 5.6 panel C shows duplicate lanes (2, 3, and 4, 5) containing products of reactions between the anchor primer and the gene specific primers 2 and 3 respectively. Lane 1 contains DNA size markers. The annealing temperatures were 50 °C for the primer 2 reaction and 46 °C for the primer 3 reaction. The extension time was 3 minutes in both cases. The ends of the two products cloned from the reaction between primers 3 and 6 were sequenced. The larger product had the anchor sequence at both ends. Furthermore, the failure of a sequencing reaction using the gene specific primer as the sequencing primer suggested that this primer was not present internally. The other clone had the gene-specific primer at one end and the anchor primer at the other but the sequence internal to the gene-specific primer was not similar to that of TbPK-A2. Thus the 3′ RACE also failed to extend the known sequence.

### 5.2.4 Isolation of a λgt11 cDNA clone containing the TbPK-A2 sequence

As an alternative strategy to isolating a cDNA containing the TbPK-A2 sequence, the cloned PCR product similar to TbPK-A2 (Section 5.2.1) was used as a probe to screen a *T. brucei* cDNA library constructed in λgt11. In the primary screen only a single plaque from a confluent plate appeared to show hybridisation to the probe and this was picked

for further screening. Fig. 5.7 shows autoradiographs of plaque lifts from secondary and tertiary screens enriched from the primary pick. The plaque chosen from the secondary screen (panel A) is marked with an arrow. In the tertiary screen (panel B) 22 plaques hybridised to the probe from a total of 78. One of these plaques (also marked) that was well separated from its neighbours was picked to provide a master stock for preparation of DNA. Purified phage DNA was restriction digested with Eco RI and the products run out in an agarose gel. The ethidium bromide stained gel was then examined by UV transillumination and, as well as the bands representing the λ vector, a single insert band of 2.8 kbp was seen. This DNA band was extracted from the gel and subcloned into the Eco RI site of pGEM7Zf(+) to give plasmid pTbplk1.

### 5.2.4.1 Subcloning and sequencing of pTbplk1

Plasmid pTbplk1 was digested with a variety of restriction enzymes and a physical map of the insert constructed (Fig. 5.8). Subclones were made of fragments A, B, C, D, and E as indicated in the figure. A deletion set was also made across subclone A from the 5′ end. The deletion set and subclones B and C were sequenced using standard procedures. The ends of subclone D were also sequenced and the overlap of the ends of subclone C with subclones B and D was checked by sequencing of pTbplk1 using gene-specific sequencing primers. Therefore, a continuous stretch (1554 bp) has been sequenced from the 5′ Eco RI site (nucleotide 1) to just past the Hinc II site in subclone D. Short sections have also been sequenced from the 3′ end of subclone D and from the 3′ end of the whole insert in pTbplk1. Fig. 5.9 panels A, B, and C shows the sequences of these three sections, arranged so that the start of each sequence corresponds to the end nearest the 5′ of the insert in pTbplk1.

The contiguous sequence from the 5′ end of pTbplk1 was translated using the Pepdata program of the Wisconsin GCG package which produces a single file with all possible translations linked together. A preliminary database search was then performed on the Pepdata output using the Fasta program of the GCG package. Predicted polypeptide sequences from two different sections of pTbplk1 (the open reading frames are also in opposite orientations) were significantly homologous to peptide sequences in the databases. As expected, one of these is similar to a range of protein kinase peptide sequences and includes the TbPK-A2 peptide sequence while the other is approximately

50% identical to a part of the ribosomal protein S4 (RPS4) homologues from mammals and yeast. Examination of the peptide sequence with homology to protein kinases shows a complete protein kinase catalytic domain according to the homology based definition of Hanks and Quinn (1991). They defined the amino terminal boundary at a position seven residues upstream from the first glycine in the glycine-rich loop of catalytic subdomain I (consensus G-x-G-x-x-G where x is any amino acid, see Section 1.2.2) and the carboxyl terminus at a hydrophobic residue 12-18 amino acids (for protein serine/threonine kinases) downstream of the conserved arginine of subdomain XI, with the last four amino acids commonly according with the consensus H-P-(aromatic amino acid)-(hydrophobic amino acid). The protein kinase catalytic domain predicted from pTbplk1 is labelled as block C in Fig. 5.9 (nucleotides 503-1267, sense strand in bold). The residues that define the boundaries of the catalytic domain are individually boxed, although the third glycine of the subdomain I consensus sequence is substituted by an alanine (see Section 5.2.4.3). The sequence that overlaps with TbPK-A2 is indicated by underlining (see Fig. 5.4 for more detailed information).

The section of sequence with homology to eukaryotic ribosomal protein S4 is labelled as block A in Fig. 5.9 (nucleotides 356-1, sense strand in bold). Alignments of the predicted peptide sequence with that of other RPS4 homologues strongly suggests that only a partial coding sequence is contained in pTbplk1 (Section 5.2.4.4). The coding sequence for the C-terminus of the partial protein sequence ends at the 5′ end (Eco RI site) of the insert in pTbplk1 while the boundary of the most N-terminal coding sequence has been set at nucleotide position 356 because extension of the open reading predicts peptide sequence unlike that of any known RPS4 homologue, although it is not at the end of the open reading frame.

The protein kinase catalytic domain from pTbplk1 was found to be most homologous to those of members of a distinct protein kinase family (polo-like kinases, see next section) and a 30 amino acid sequence encoded by nucleotides 127-216 of Block E (Fig. 5.9 panel B) shows a high level of homology to a C-terminal domain (PH2 domain) conserved among polo-like kinases (see Section 5.2.4.3). Thus the putative protein kinase from pTbplk1 appears to be a polo-like kinase. The predicted domain structure of the putative *T. brucei* protein kinase is compared with the domain structure of known polo-like kinases in Fig. 5.10. It has been assumed that both of the

unsequenced sections of pTbplk1 contain coding sequence for the protein kinase as both fall between sections of nucleotide sequence for which the predicted peptide sequences show significant homology to polo-like kinases (see Fig. 5.13). The different domains are described briefly in the legends to figures 5.9 and 5.10 and discussed further in Sections 5.2.4.2 and 5.2.4.3.

The finding that a region of pTbplk1 predicts a polypeptide with a high level of homology to a part of the sequence of RPS4 homologues was unexpected since the λ clone was obtained using a probe from the protein kinase coding sequence and a cDNA clone would be expected to contain coding sequence for only a single protein. The most likely explanation seemed to be that partial cDNAs for the protein kinase and RPS4 homologues had become artifactually ligated during library construction in which case there would be no expectation that the two genes need be close to each other in the genome. In order to check whether the protein kinase and RPS4 coding sequences were co-linear in the genome, a Southern blot of *T. brucei* genomic DNA was hybridised sequentially to probes from the protein kinase and RPS4 coding regions. Fig. 5.11 panel A shows the autoradiograph of the blot hybridised to the protein kinase probe (insert of subclone B) while panel B shows the autoradiograph of the same blot after it was stripped and hybridised to the RPS4 probe (insert of subclone E). The *T. brucei* genomic DNA was digested with Hind III or Eco RI (indicated on figure), neither of which cuts anywhere internally to pTbplk1. The hybridisations were performed under high stringency conditions. As seen in panel A the protein kinase probe hybridises to a single band in each lane (similar to the results of Hua and Wang, 1994) while the ribosomal probe hybridised to two bands in the Hind III lane but only one in the Eco RI lane. The single band hybridising to the protein kinase probe is not similar in size to any of the bands hybridising to the RPS4 probe. Therefore, while both the probe sequences used appear to be from *T. brucei*, they are not co-linear in the genome although are co-linear in the cDNA clone. This supports the simplest conclusion that the protein kinase and RPS4 coding sequences are artifactually ligated. Hybrid cDNAs that include ribosomal protein cDNAs are common contaminants of cDNA libraries, presumably because the messages are abundant. This has been shown to be the case in *T. brucei* by the high representation of ribosomal protein cDNAs in the randomly chosen *T. brucei* cDNAs sequenced by El-Sayed *et al.* (1995).

### 5.2.4.2 The putative protein kinase from pTbplk1 is a member of the polo-like kinase family

In order to identify protein kinases similar to the protein kinase predicted from pTbplk1, database comparisons with the catalytic domain peptide sequence (i.e. the peptide sequence of Block C in Fig 5.9) were performed using the BLAST network service at the NCBI to search the PDB, SwissProt, PIR, and GenPept databases and also the SwissProt and GenPept updates. Then, in order to determine percentage identities across the full catalytic domain for the closest matches (rather than the segments aligned by the BLAST program), the Fasta program of the Wisconsin GCG package was used. Table 5.1 lists the closest matches identified in the database searches. The highest level of identity is with Cdc5 of *Saccharomyces cerevisiae* to which the *T. brucei* catalytic domain peptide sequence is 53% identical over 257 amino acids. The *T. brucei* sequence also shows a very similar level of homology to plo1 of *Schizosaccharomyces pombe*. The next best match is with the human protein kinase Plk1 closely followed by the murine protein kinase STPK13 and a protein kinase from rat (rat plk). Three further sequences similar to Plk1 and two similar to STPK13 have been published but have not been included in this analysis because the degree of similarity between the reported sequences in both cases is so high that it seems likely that the differences represent polymorphisms in the same gene. The Plk1, STPK13 and rat plk proteins are very closely related to each other (> 94% identity over the entire proteins) and are probably functional homologues. The next two most related sequences are from two further murine protein kinases, Fnk and Snk, which are, respectively, 46.7% and 46.5% identical to the *T. brucei* sequence, but are both only about 52% identical to murine STPK13, while being 68.8% identical to each other. The last three entries in the table are a *C. elegans* protein kinase predicted from genomic DNA, a protein kinase from *Drosophila* named polo, and a fourth murine protein kinase Sak, for which there are two isoforms which have identical catalytic domains. The level of identity of the *T. brucei* sequence to the sequences listed in Table 5.1 compares with 25%-35% identity to many members of the cyclic nucleotide dependent and protein kinase C families. This lower level of identity suggests that the cloned sequence is not a member of the protein kinase C family.

Examination of the literature showed that Cdc5, plo1, Plk1, STPK13, rat plk, Fnk, Snk, and polo are all recognised members of a family of protein kinases (polo-like

kinases) for which the *Drosophila* protein kinase polo is the prototype. The common features which define members of this protein kinase family are closely related N-terminal catalytic domains and C-terminal extensions of about 300 amino acids that share a well-conserved 30 amino acid domain (termed the PH2 domain by Hamanaka *et al.*, 1994). Sak a and Sak b (Fode *et al.*, 1994) have N-terminal catalytic domains and significant C-terminal extensions but in neither case is the sequence C-terminal to the catalytic domain similar to that of polo-like kinases. The *C. elegans* protein kinase predicted from genomic sequence is not a recognised polo-like kinase but examination of the sequence suggests that it belongs to this protein kinase family (see Section 5.2.4.3). The high level of homology of the catalytic domain of the putative protein kinase from *T. brucei* to those of polo-like kinases suggested that it might be a family member. This hypothesis was strongly supported by the finding that an open reading frame, extending through the 3′ most sequenced section of subclone D (i.e. Block E in Figs. 5.9 and 5.10), predicts a peptide sequence that includes a domain homologous to the PH2 domain. Therefore, although the full coding sequence is not yet available, the putative protein kinase from *T. brucei* has been named TbPLK (*T. brucei* polo-like kinase).

### 5.2.4.3 The probable primary structure of TbPLK and alignments of the catalytic domain and C-terminal sequences with those of other polo-like kinases

It is expected that complete sequencing of pTbplk1 will show that the open reading frame for TbPLK extends at least from the 5′ end of Block C (nucleotide 503) to the 3′ end of the whole insert in pTbplk1. The open reading frame of Blocks C and D can be extended 5′ through Block B and this block of sequence probably encodes at least part of the N-terminus of TbPLK. No significant homology was found between the predicted peptide sequence from Block B and the peptide sequences N-terminal of the catalytic domain of other polo-like kinases, but the N-termini of polo-like kinases are poorly conserved except between the closely related Plk1, STPK13 and rat plk. Therefore the lack of homology between the *T. brucei* sequence and the other sequences is not evidence against Block B encoding the N-terminus of TbPLK. If Block B is part of the protein kinase cDNA it is unclear whether the N-terminal coding sequence is complete or not. There is a possible start codon (nucleotides 377-379) which, if used, would give an N-terminal domain of 42 amino acids but, if the open reading frame is extended slightly

5′ the predicted peptide sequence is KKKKRM similar to short strings of basic amino acids found in the N-termini of Cdc5 and Snk. These amino acid sequences could represent nuclear localisation signals (Dingwall and Laskey, 1986; Goldfarb *et al.*, 1986). However, not all the polo-like kinases have such sequences in their N-termini and similar sequences internal to the catalytic domain of polo-like kinases might fulfil a nuclear targeting function (Holtrich *et al.*, 1994). Whatever the true sequence for the N-terminus of TbPLK, it will clearly have to be confirmed from new library clones or PCR products.

It also appears that pTbplk1 does not contain the complete coding sequence for the C-terminal end of TbPLK since an open reading extending throughout Block F (which is from the very 3′ end of the insert in pTbplk1) predicts a peptide sequence that shows significant homology to the C-termini of polo-like kinases. However, there may be very little 3′ coding sequence for TbPLK not contained within pTbplk1 (see below). To summarise, pTbplk1 probably contains almost all of the coding sequence for the putative protein TbPLK but some 3′ coding sequence for TbPLK is not contained in pTbplk1 and, while the 5′ coding sequence may be complete, this sequence will also have to be confirmed from independent clones. The TbPLK message would also appear to have a considerable amount of untranslated sequence that is not contained in pTbplk1 since the size of the region within pTbplk1 that is likely to be protein kinase coding sequence is about 2 kb smaller than the size of the poly(A)$^+$ RNA that hybridised to TbPK-A2 in a Northern analysis performed by Hua and Wang (1994). The catalytic domain, junction domain, PH2 domain and C-terminal sequences are considered in more detail below.

Fig. 5.12 shows an alignment of the putative *T. brucei* protein kinase catalytic domain with the sequences from its closest homologues as listed in Table 5.1. For Plk1 and STPK13, for which further very closely related sequences have been published, a consensus of the available sequences was used. The *C. elegans* protein kinase (Wilson *et al.*, 1994) is a putative protein kinase predicted from genomic DNA by the program Genefinder, and the predicted peptide sequence in the databases ends at a stop codon shortly after the invariant arginine of subdomain XI. However, another predicted open reading frame that begins shortly downstream encodes polypeptide sequence homologous to the C-terminal extensions of polo-like kinases, including the PH2 domain. It therefore seems likely that the current prediction of two proteins from this region of genomic DNA is incorrect, and that the true message encodes a protein kinase

similar to the recognised polo-like kinases. Introduction of a small additional intron (position 21871 in cosmid K06H7 to 1040 in cosmid C14B9) gives a putative mRNA encoding a polo-like kinase. The last few amino acids of the catalytic domain predicted from this "corrected" putative mRNA are a closer match both to the corresponding positions in polo-like kinases and to the protein kinase consensus than the current prediction in the databases. The peptide sequences in the alignments are from the "corrected" putative mRNA.

Fig. 5.13 is a similar alignment of the polypeptide sequences of polo-like kinases from the junction domain to the C-termini. In both figures the Pileup program of the Wisconsin GCG package was used to make the alignments but in Fig. 5.13, because the full predicted peptide for the C-terminus of TbPLK is not yet available, the three partial sequences (from Blocks D, E and F; see Figs 5.9 and 5.10) were added to the alignment of the other complete sequences. Residues that are identical between two or more sequences are indicated by shading, while dots represent gaps introduced to align the sequences. In Fig. 5.12 the protein kinase catalytic subdomains are indicated by Roman numerals, and the PH2 domain is indicated in Fig. 5.13 by dashed underlining. The STPK13 and rat plk C-terminal sequences, both of which are very similar to that of Plk1 (H.s.), have been omitted in Fig. 5.13 for reasons of space.

Considering first the catalytic domain alignment in Fig. 5.12, subdomain I of protein kinases is a glycine-rich loop between β-strands that anchors the β-phosphate of the nucleotide substrate (reviewed by Bossemeyer, 1994). The consensus sequence for both protein serine/threonine and protein tyrosine kinases is GxGxxGxV, where x is any amino acid. The residue preceding the first glycine is hydrophobic in all the 117 protein kinases included in the alignment of Hanks and Quinn (1991) while the residue preceding the third glycine is very frequently tyrosine or phenylalanine. The sequence of TbPLK (GRGGFAKC) differs at two positions from this consensus; the valine replaced by cysteine being more heavily conserved than the third glycine replaced by alanine. Examination of Fig. 5.12 shows that these differences are conserved among polo-like kinases, with all known members of this family having subdomain I sequences corresponding to a consensus GxGGFAxC. The functional significance of the differences between the subdomain I consensus for polo-like kinase compared to the general protein kinase consensus is not known but it is thought that for some protein kinases the residues

between the glycines make important contacts with peptide substrates and pseudosubstrate inhibitors. Human CDK2 is negatively regulated by phosphorylation at two threonine residues between the second and third glycines: phosphorylation of T14 disrupts ATP binding while phosphorylation of either or both residues might also affect peptide substrate binding (De Bondt *et al.*, 1993). Therefore, while the glycine rich loop of protein kinases is a conserved motif necessary for nucleotide binding, the peptide sequence in this part of the catalytic domain can clearly have functional significance specific to particular protein kinases, and the conservation of an unusual subdomain I sequence among polo-like kinases suggests that this may be true for this protein kinase family.

Subdomain II has an invariant lysine and subdomain III an invariant glutamate (the subdomain labels have been placed directly below these residues) that are present in all the polo-like kinases including TbPLK. In PKA, the protein kinase which is best understood in structural terms, the lysine interacts with the ATP α- and β-phosphates and forms an ion-pair with the glutamate of subdomain III. Subdomain IV corresponds to the junction of the major and minor lobes and has no invariant residues. Subdomains V and VIb also lack invariant residues but there is still significant homology in small blocks between the polo-like kinases shown in the figure. Subdomain VIb is the catalytic loop in which the glutamate and asparagine are invariant among all protein kinases and the lysine is highly conserved among protein serine/threonine kinases. There is a high degree of sequence homology between the different polo-like kinases in this subdomain. Subdomains VII, with the highly conserved DFG motif, and VIII with a highly conserved PE doublet, also represent regions where there is a very high level of sequence conservation between the different polo-like kinases including TbPLK. All of the polo-like kinases correspond exactly to the protein kinase consensus DxWxxG in subdomain IX, while in subdomain X there is more divergence before the invariant arginine of subdomain XI present in all the sequences. To summarise, outside subdomain I, TbPLK has all of the most highly conserved residues of protein serine/threonine kinases while its subdomain I sequence corresponds to the consensus for polo-like kinases. Therefore, as far as can be judged from the catalytic domain peptide sequence alone, the putative protein TbPLK could be active as a polo-like protein kinase.

A feature common to many protein kinases is a phosphorylation site between subdomains VII and VIII. In PKA this site (T197) is phosphorylated shortly after protein synthesis, is phosphatase resistant and contributes to protein stability through electrostatic interactions with several other residues (Shoji *et al.*, 1979; Knighton *et al.*, 1991b). In ERK2 and CDK2 phosphorylation at equivalent positions has a regulatory role (reviewed by Taylor and Radzio-Andzelm., 1994). TbPLK and all other known polo-like kinases have a threonine at an equivalent position to T197 of PKA (indicated by an asterisk in Fig. 5.12) and it is possible that phosphorylation of this site is required for activity. Autophosphorylation has been demonstrated for Plk1 and STPK13 (see Section 5.3.3) and a labelled gel band that probably represents autophosphorylated polo kinase was detected in the products of protein kinase assays performed by Fenton and Glover (1993) but the phosphorylation sites have not been determined.

Golsteyn *et al.* (1994) noted that three glutamate residues in PKA that are implicated in binding to positively charged peptide substrates are replaced in human Plk1 by uncharged residues but at a further three positions, implicated in substrate binding in PKA, both protein kinases have hydrophobic residues. The positions where PKA has glutamates are indicated in Fig. 5.12 by filled boxes while the hydrophobic residues are indicated by open boxes. It can be seen that TbPLK and all the other polo-like kinases have uncharged residues at the former positions and hydrophobic residues at the latter as for Plk1. The preferential use of casein over histone H1 by polo, Cdc5 and Plk1/STPK13 (Fenton and Glover, 1993; Kitada *et al.*, 1993; Golsteyn *et al.*, 1995; Lee *et al.*, 1995) is consistent with the idea that the marked residues are involved in determining substrate specificity in polo-like kinases.

As mentioned above the complete sequence for the C-terminal extension of TbPLK is not yet available but there are three partial sequences likely to be from this region (peptide sequences of Block D, E and F) that have been added to the alignment of the other polo-like kinase sequences in Fig. 5.13. The peptide sequence of TbPLK immediately C-terminal to the catalytic domain (from Block D) could not be aligned convincingly with the other sequences and the TbPLK sequence in this region has therefore simply been presented in Fig. 5.13 without gaps. The next partial sequence (from Block E) includes the highly conserved PH2 domain that confirms the relatedness of TbPLK to polo-like kinases from other organisms. The PH2 domain was used to align

this section with the other sequences. It also proved possible to align the most C-terminal partial sequence (from Block F) with the alignment of the other polo-like kinase sequences. Examination of Fig. 5.13 shows that immediately C-terminal to the end of the catalytic domain there is a region of about 20 amino acids across which there is significant homology between all of the polo-like kinases from multicellular organisms, to which the Cdc5 and plo1 sequences show some homology while the TbPLK sequence is quite divergent. This is followed by a large region in which the Pileup program has introduced extensive gaps to align the sequences and there are only short sections of sequence identity. Assuming that the TbPLK open reading frame extends through subclone D, the putative protein kinase will have an insert of about 60 amino acids in the junction domain compared to Cdc5. In view of the divergence between the other polo-like kinases in this region in terms of both sequence and length of polypeptide chain, this does not seem particularly improbable. Following the large divergent junction domain is the highly conserved PH2 domain (dashed underlining). TbPLK has substitutions at five positions where all the other sequences in the alignment have the same amino acid, but three of these could be considered conservative. The TbPLK sequence is most similar to that of plo1 across the PH2 domain (53% amino acid identity). The alignment shows that, as well as a high degree of sequence conservation between polo-like kinases in the PH2 domain, the sequences C-terminal to this domain are more highly conserved than the junction domain and, consistent with this, the partial *T. brucei* sequences C-terminal to the PH2 domain show a degree of homology to the sequences of the other polo-like kinases.

#### 5.2.4.4 Similarity of the partial *T. brucei* ribosomal protein S4 sequence to sequences of ribosomal protein S4 (RPS4) homologues from other eukaryotes

The nucleotide sequence from pTbplk1 that predicts a peptide sequence similar to that of RPS4 homologues extends from nucleotides 1-356, translated in the reverse orientation (see Fig. 5.9). Fig. 5.14 depicts an alignment of the predicted peptide sequence from *T. brucei* with those of RPS4 homologues from *S. cerevisiae*, *Tetrahymena*, human (X isoform), *Drosophila* and potato. Residues common to two or more sequences at a particular position are indicated by shading. RPS4 sequences derived from two *T. brucei* expressed sequence tags (ESTs) are also included in the figure (El-Sayed *et al.*, 1995). In

both of these sequences there are ambiguities and only blocks of peptide sequence with a high degree of similarity to the other RPS4 homologues are shown. One of these EST sequences predicts peptide sequence similar to the N-termini of RPS4 homologues but does not overlap with the partial sequence from pTbplk1. The sequence of the other EST predicts a peptide sequence similar to the C-termini of other RPS4 homologues and overlaps with the sequence from pTbplk1. This EST sequence has an Eco RI site at a position corresponding exactly to the 5′ end of the insert from the λ clone, indicating that the 5′ end of the insert was ligated into the λ vector at a naturally occurring Eco RI site within the RPS4 coding region.

It can be seen from the alignment that there is very significant homology between the putative *T. brucei* RPS4 and the other RPS4 sequences. The *T. brucei* sequence from pTbplk1 is 51.3% identical to the *S. cerevisiae* RPS4 homologue over 115 amino acids and 50, 47.5, 47.5, and 44.1% identical to the human, *Drosophila*, potato, and *Tetrahymena* RPS4 homologues respectively over 118 amino acids. Unfortunately, little functional data is available for RPS4 proteins so it is not possible to comment on the significance of similarities and difference between the sequences but, because of the high degree of similarity of the *T. brucei* sequence to that of RPS4s from other organisms, the putative protein has been named TbS4lp (*T. brucei* S4-like protein).

### 5.3 Discussion

To complement earlier biochemical studies of protein kinase C from *T. brucei* it was decided to isolate a full length cDNA corresponding to a PCR fragment (TbPK-A2) for which the predicted peptide sequence showed homology to some members of the protein kinase C family. Various approaches were used in attempts to isolate a full length cDNA including 5′ and 3′ RACE and library screening. A λgt11 clone was obtained by screening a library with a PCR fragment similar to TbPK-A2. Sequencing the insert from this clone showed that it encodes a complete protein kinase catalytic domain that, as expected, includes the TbPK-A2 sequence. However, analysis of the predicted protein sequence showed that the kinase belongs to a distinct family of protein kinases implicated in cell cycle control in a number of organisms (polo-like kinases) rather than to protein kinase C. In the next section the results of the 5′ and 3′ RACE are discussed. The *T. brucei* RPS4 homologue is discussed briefly in Section 5.3.2. The available information

about polo-like kinases in higher eukaryotes is reviewed in Section 5.3.3 while possible future work into the cellular function of TbPLK is discussed in Section 5.3.4 in relation to what is known about cell cycle control in trypanosomes.

### 5.3.1 5′ and 3′ RACE

The 5′ and 3′ RACE proved unsuccessful as approaches for obtaining the complete protein kinase gene sequence. This lack of success might have been overcome by using a number of improvements to the procedure. In the 5′ RACE, a gene-specific primer could have been used in cDNA synthesis. This would bias cDNA synthesis towards producing only the desired products, and reduce the length of mRNA that has to be reverse-transcribed thereby reducing the chances of premature termination. Instead, oligo(dT) primers (with and without anchor sequences) were used in all cDNA syntheses. The specific product of primers 2 and 5 could, however, be readily amplified from both batches of cDNA that were made. Another possible problem in the 5′ RACE was the relatively low annealing temperatures that were used with the 22 bp primer based on the *T. brucei* spliced leader (primer 1). The sequence of the 39 bp spliced leader has a low G+C content and to compensate for this a longer primer could have been synthesised which would have allowed higher annealing temperatures in the PCR reactions with possible improvements in reaction specificity.

The 3′ RACE reactions had not been well characterised at the time the decision was made to abandon PCR in favour of library screening. The preliminary reactions appeared more promising than those from the 5′ RACE as products of a significant size were amplified in reactions with annealing temperatures much higher than could be used in the 5′ RACE. The major products were cloned and the ends sequenced but these proved not to be the specific products desired. It is possible that one or more further rounds of PCR with nested primers might have given the desired product as a single band but the results obtained show that an improvement to the procedure that was clearly needed was the use of single primer controls. One of the two products that was cloned had the anchor primer at both ends while a number of the product bands from the 3′ RACE with the two different gene-specific primers 2 and 3 and the 3′ anchor primer appeared to be similar (Fig. 5.6 panel C), suggesting that these too could be products of only one primer.

The results of Hua and Wang (1994) showed that the messenger RNA hybridising to TbPK-A2 was large (5.4 kb), while the closest matches to the peptide sequence predicted by TbPK-A2 were protein kinases with the catalytic domain in very different positions within the polypeptide chain. Therefore, it was not possible to judge the likely sizes of specific products for products of either the 5′ or 3′ RACE reactions. The largest likely sizes for the 5′ and 3′ product were also towards the limit of what might be considered reasonable for RACE, at least using standard PCR reagents and conditions. Therefore, RACE was not an ideal approach compared to screening of a cDNA library but only a small amount of a cDNA library was available. The clone obtained from the library screen is not full length and 5′ and 3′ RACE with the improvements suggested above and with new primers from sequence nearer the 5′ and 3′ ends respectively would be one way to isolate amplified cDNAs from which to complete the sequence. Another approach would be to re-screen the library with subclones from the 5′ and 3′ ends of pTbplk1.

### 5.3.2 A ribosomal protein S4 homologue from *T. brucei*

Part of pTbplk1 appears to be a truncated coding sequence for a *T. brucei* RPS4 homologue. Ribosomal proteins in general are highly conserved among eukaryotes and the degree of homology of the putative *T. brucei* protein and other RPS4 homologues is high. Its *T. brucei* origin was confirmed by high stringency genomic Southern hybridisation (Section 5.2.4.1). The putative protein has been named *T. brucei* S4-like protein (TbS4lp).

Ribosomes, as large molecular assemblies of many components, are difficult to study and only very limited data is available on the functional role of RPS4 homologues. One way of probing the proximity of different ribosomal components to each other in the assembled ribosomes is through the use of cross-linking agents. In studies of this type RPS4 has been found to contact eIF-3 in the 40S subunit near to the interface with the 60S subunit and to be situated close to the 28S rRNA (Nygård and Westermann; 1982 Nygård and Nika, 1982). Cross links identified between RPS4 and ribosomal protein L5 also place RPS4 at the subunit interface and ribosomal protein L5, in association with the 5S RNA, may be important in placing the AUG start codon into the correct position for initiation (Uchiumi *et al.*, 1986).

Although obtained as an artifact, the *T. brucei* RPS4 sequence might be useful for phylogenetic studies if a full length sequence was obtained. *T. brucei* ribosomes do not appear to have been studied in any detail, but it is interesting that in *Tetrahymena* the RPS4 homologue is phosphorylated and its phosphorylation state correlates with the nutritional state of the cell, the thermal stability of its ribosomes and the sensitivity of the ribosomes to cycloheximide (Hallberg *et al.*, 1981; Palm *et al.*, 1995).

### 5.3.3 Polo-like kinases in higher eukaryotes

The polo-like kinases are a recently discovered protein kinase family (the sequence of the *Drosophila* gene *polo* was published in 1991 although the phenotypes of mutants in both *polo* and *CDC5* were examined before the genes were sequenced) but work on the cellular function of these kinases has progressed rapidly. The mitotic defects of *polo* mutants in *Drosophila*, *CDC5* mutants in *S. cerevisiae* and *plo1*$^{+}$ disruptants in *S. pombe* implicate the encoded protein kinases in regulation of the cell cycle and, while no genetic or reverse genetic experiments have been reported for Plk1 and STPK13, there is considerable evidence to suggest that polo, Cdc5, plo1, Plk1, STPK13 are functional homologues. However, Snk and Fnk are expressed as early growth response genes and thus differ in expression pattern to the other polo-like kinases (Simmons *et al.*, 1992; Donohue *et al.*, 1995). Therefore there are likely to be different roles for different polo-like kinases, at least in mammals.

The phenotypes of mutants in *polo* have been studied in some detail (Sunkel and Glover, 1988; Llamazares *et al.*, 1991). The *polo*$^{1}$ allele was originally discovered by its female sterile phenotype (the stronger *polo*$^{2}$ allele is lethal at the larval stage in homozygotes). A proportion of female *polo*$^{1}$ homozygotes of heterozygote mothers survive to adulthood but the embryos of these homozygotes (which lack maternally provided wild-type protein) accumulate defects during the syncytial stage in which multiple nuclear divisions occur in a single cytoplasm and fail before the fourteenth syncytial cell cycle when the embryo would normally become cellular. The nuclei of these embryos become polyploid and the chromatin is disorganised despite undergoing synchronous rounds of condensation and decondensation. Spindle structures were observed that did not appear to be nucleated by a centrosome. The mitoses of other developmental stages were also examined in these studies. Spindle morphologies of

larval neuroblasts of *polo*$^{1}$ homozygotes included condensed chromosomes that had failed to align on a metaphase plate, overcondensed chromosomes and monopolar spindles nucleated from a single centrosome, consistent with the defects found in the syncytial stage. The meiotic defects of *polo* mutants also point to a role for the polo kinase in regulating centrosome and spindle structure (Sunkel and Glover, 1988). Chromosomal non-disjunction occurs in spermatids of *polo*$^{1}$ males and mitochondrial segregation, which is also mediated by the meiotic spindle, was abnormal suggesting that the defect is with the spindle rather than the chromosomal kinetochores. The second meiotic division (similar to mitosis) is also more severely affected than the first division in which the centromeres do not split apart.

*CDC5* was shown to be an essential gene in *S. cerevisiae* in an experiment in which one copy of the *CDC5* gene was replaced in diploid cells by the *LEU2* gene and the disruptant cells then sporulated. In 24 tetrads analysed only 2 spores were viable in each and these were all Leu$^{-}$ and therefore contained the wild type *CDC5* gene. Spores possessing the disrupted gene could be rescued by a multicopy plasmid carrying the *CDC5* gene (Kitada *et al.*, 1993). However, a plasmid carrying the gene for a mutant Cdc5 with a single substitution at an essential protein kinase catalytic residue was unable to rescue disruptant spores, strongly suggesting that protein kinase activity is required for Cdc5 function. Wild-type cells transfected with a plasmid expressing *CDC5* grew but with increased doubling times, suggesting that overexpression of *CDC5* is harmful. "Switch-off" experiments were also performed in *CDC5* disruptants carrying a plasmid expressing *CDC5* under control of the *GAL1* promoter. When glucose was added to the growth medium to repress expression of the *CDC5* gene from the plasmid, cell number stopped increasing by about 2 hours after the addition of glucose but cell mass continued to increase for up to 6 hours before stopping. At 4 hours after the addition of glucose 82% of the cell population consisted of enlarged cells which each had a large bud while 15% of the population consisted of two enlarged connected cells with large buds in which the nuclei had almost completed division but were connected by a thin chromatin bridge. The DNA content of the cells was analysed by flow cytometry. Before the addition of glucose, 25% of cells had a 1C DNA content ($G_1$ phase cells) and 61% of cells had a 2C DNA content ($G_2$/M phase cells) while at the 6 hour time point (after addition of glucose) only 3% of cells had a 1C DNA content, 71% of cells had a 2C

DNA content and 23% of cells had a 4C DNA content (presumably due to the enlarged connected cells each with a large bud observed by microscopy). The cellular effects of disruption and overexpression of *plo1*$^+$ in *S. pombe* have also been examined (Ohkura *et al.*, 1995). In tetrad analysis experiments in which one copy of the *plo1*$^+$ gene was replaced in diploids by the *ura4*$^+$ gene and the cells then sporulated and grown in medium containing uracil to allow only *ura4*$^+$ spores to grow, the disruptant haploids were inviable, showing that *plo1*$^+$ is an essential gene. These spores showed three types of defect. Some cells had one cluster of overcondensed chromosomes, others had two interphase nuclei and were unseptated, while a third set had two clusters of overcondensed chromosomes and were also unseptated. Analysis over a time course suggested that the first two types of defect resulted from arrest in the anaphase of the first mitosis and that the third type resulted from a blockage in the second mitosis after failure to produce a septum in the first mitosis. Loss of *plo1*$^+$ function was also examined in experiments in which haploid *plo1*$^+$ disruptants were maintained by episomal expression of *plo1*$^+$ under the control of a thiamine repressible promoter. In the absence of thiamine, the cells grew apparently normally but, following the addition of thiamine to the growth medium, the cells accumulated both mitotic and septation defects as seen in the spore germination experiments. In the spore germination experiments, it appeared that most spores were blocked in their first mitosis with the remainder blocked in the second mitosis. However, in the switch-off experiments, cells were observed which had multiple nuclei but no septum, suggesting that one or two nuclear divisions had occurred after the failure of septum formation before mitosis also became blocked. Since the loss of *plo1*$^+$ function would be more gradual in the switch-off experiments, it was suggested that a lower level of plo1 activity might be required for mitosis than for septum formation. Immunochemical experiments showed that, in both spore germination and switch-off experiments, over-condensed chromosomes could be observed that were associated with monopolar spindles (similar to defects seen in *polo* mutants in *Drosophila*) and that actin ring formation and the deposition of septal material were abnormal. Episomal expression of *plo1*$^+$ was also used to overexpress the protein in wild type cells leading to cells with over-condensed chromosomes and cells with one or more septa and an interphase nucleus. It was also shown, in a temperature-sensitive mutant of *cdc25* that arrests in late $G_2$, that overexpression of *plo1*$^+$ could drive septum formation

at the restrictive temperature and a similar result was also obtained with a *cdc2* mutant, suggesting that *plo1*$^{+}$ can drive septum formation without p34$^{cdc2}$ activation being required. Furthermore, *plo1*$^{+}$ overexpression could also drive septum formation in temperature-sensitive mutants of *cdc10* that arrest prior to Start in $G_1$. Therefore it appears that in *S. pombe*, plo1 is the key inducer of septum formation while also being a regulator of spindle formation. Based on the similarity of the mitotic defects in *plo1*$^{+}$ disruptants to those in *cut7*$^{+}$ mutants (Hagan and Yanagida, 1990; Hagan and Yanagida, 1992; cut7 is a member of the kinesin superfamily) it was suggested that cut7 might be a substrate of plo1.

No gene disruption experiments have been performed for any of the mammalian polo-like kinases but Hamanaka *et al.* (1994) found that microinjection of *in vitro* transcribed STPK13 mRNA into NIH 3T3 cells maintained in low serum medium induced thymidine incorporation while microinjection of antisense RNA into growing NIH 3T3 cells blocked thymidine incorporation.

The biochemical properties of polo, Cdc5, Plk1 and STPK13 have also been examined. Kitada *et al.* (1993) were able to use an antiserum against Cdc5 to immunoprecipitate a casein kinase activity from *S. cerevisiae* overexpressing *CDC5* but not from cells overexpressing the mutant of *CDC5* (encoding a protein substituted at a single catalytic domain residue) that had been shown to be unable to complement *CDC5* disruptants. However, they were unable to assay Cdc5 activity from wild-type cells. Fenton and Glover (1993) measured the activity of polo with respect to progress through the *Drosophila* syncytial embryo cell cycle again using casein as a model substrate. Polo kinase activity was maximal at a time corresponding to late anaphase-telophase when histone H1 kinase activity, immunoprecipitated by a cyclin B antibody, was at a low level. A phosphorylated band was detected in the assays that probably represents autophosphorylated polo, but it was not determined whether autophosphorylation affects casein kinase activity. The most detailed biochemical studies of polo-like kinases have been performed for Plk1 and STPK13 (Golsteyn *et al.*, 1994; Golsteyn *et al.*, 1995; Hamanaka *et al.*, 1995; Lee *et al.*, 1995). Golsteyn *et al.* (1994) examined Plk1 protein levels in populations of HeLa cells sorted by centrifugal elutriation. Plk1 protein was at a very low level in cell populations enriched for $G_1$ cells, was at a higher level in populations enriched for S phase cells and was at maximal levels in populations with the

highest proportion of cells in $G_2$/M. As expected $p34^{cdc2}$ levels did not vary significantly between the populations at different cell cycle stages. A more precise examination of Plk1 levels during mitosis was performed using HeLa cells arrested in prometaphase by a nocodazole block and then sampled at various time points following release. Consistent with the results from the cells sorted by elutriation Plk1 levels were high in the cells blocked in M phase and remained high 30 minutes after release. Plk1 levels were detectably lower 1 hour after release and by 4 hours, when the great majority of the cells had re-entered $G_1$ (as assessed by re-attachment to the culture dish), Plk1 was almost undetectable. Thus it appears that Plk1 protein levels fall during late mitosis/early $G_1$. This analysis was extended in the later study (Golsteyn *et al.*, 1995) to include assay of Plk1 protein kinase activity. HeLa cells were synchronised at $G_1$/S by a thymidine/aphidicolin double block and then collected at regular intervals after release from the block. Plk1 activity (assayed with casein) and $p34^{cdc2}$ activity (assayed with histone H1) were low between 0 and 8 hours after release from the block but increased dramatically as the cells passed through mitosis between 12 and 16 hours after release. By 18 hours when most cells had completed mitosis and re-entered $G_1$ both Plk1 and $p34^{cdc2}$ activities had decreased again. Similar experiments, either performed using cells released from a nocodazole-induced prometaphase block or cells from an asynchronous culture fractionated by centrifugal elutriation, confirmed that Plk1 activity was maximal during mitosis. Plk1 protein levels were shown to fluctuate during the cell cycle as found in the previous study but it was reported that Plk1 from mitotically arrested cells was 4-6 times more active in phosphorylating casein than the equivalent amount of Plk1 from interphase cells. Similar results showing maximal levels of Plk1 at $G_2$/M with a 4-6 fold increase in specific activity in Plk1/STPK13 isolated from mitotic cells compared to that isolated from S phase cells were obtained by Hamanaka *et al.* (1995) and Lee *et al.* (1995). Hamanaka *et al.* (1995) also found that Plk1 isolated from mitotic cells migrates slower in SDS-PAGE than Plk1/STPK13 from S phase cells and immunoprecipitation of STPK13 from NIH 3T3 cells cultured in medium containing [$^{32}$P]orthophosphate showed that STPK13 from mitotic (non-adherent) cells was phosphorylated while that from adherent cells was not. As expected, the phosphorylated form had reduced mobility in SDS-PAGE and also showed increased protein kinase activity against casein. Phosphoamino acid analysis showed that STPK13 isolated from mitotic NIH 3T3 cells

labelled with orthophosphate *in vivo* was phosphorylated only on serine, while *in vitro* autophosphorylation (also activated in mitotic Plk1/STPK13) with [$\gamma$-$^{32}$P]ATP was on both serine and threonine, but did not cause the mobility shift associated with Plk1/STPK13 isolated from mitotic cells (autophosphorylation of Plk1/STPK13 in kinase assays was also detected by Golsteyn *et al.*, 1995 and Lee *et al.*, 1995). Treatment of STPK13 immunoprecipitated from mitotic cells with protein phosphatase 2A converted most of the slower migrating protein to the faster migrating form and reduced the protein kinase activity 5-10 fold. It seems likely that the lower mobility form of Plk1/STPK13, phosphorylated on at least one serine residue, is an activated form. However, it is also possible that Plk1/STPK13 is phosphorylated at multiple sites and that an intermediately phosphorylated form is the activated form. Since *in vitro* autophosphorylation of Plk1/STPK13 did not activate the enzyme one or more other protein kinases are probably responsible for the activating phosphorylation(s). The possibility that p34$^{cdc2}$ was the activating enzyme was investigated by Hamanaka *et al.* (1995). Examination of the Plk1/STPK13 peptide sequence showed no consensus p34$^{cdc2}$ phosphorylation sites containing serine and, although p34$^{cdc2}$/cyclin B was found to phosphorylate Plk1 *in vitro*, this did not lead to a mobility shift and caused only a small activation of Plk1 (<2-fold). Phosphorylation of Plk1 by p34$^{cdc2}$ followed by autophosphorylation of Plk1 also failed to lead to the mobility shift associated with Plk1/STPK13 isolated from mitotic cells. In contrast, Lee *et al.* (1995) were unable to show any phosphorylation of Plk1 by p34$^{cdc2}$/cyclin B and also found that the MAP kinase ERK1 was unable to phosphorylate Plk1.

Having found that Plk1/STPK13 from mitotic cells was phosphorylated and that this phosphorylation caused a mobility shift in SDS-PAGE, Hamanaka *et al.* (1995) re-examined the kinetics of phosphorylation and activation of Plk1 in cells progressing from $G_1$/S to $G_2$/M as a comparison with those of p34$^{cdc2}$ dephosphorylation and activation at the $G_2$/M transition. Cells were sampled at various time points from a culture synchronised at the $G_1$/S transition by a double aphidicolin block and then released into nocodazole (to trap the cells at prometaphase). At 7.5 hours after nocodazole release when the slower migrating form of Plk1 was first detected 90% of the cells were in $G_2$/M and 40% had entered M phase. The amount of the slower migrating (phosphorylated) form of Plk1 increased at later time points as more cells entered mitosis. The higher

mobility (dephosphorylated) form of $p34^{cdc2}$ appeared at approximately the same time as the lower mobility form of Plk1 and, as expected, these changes coincided with activation of both enzymes. In order to test whether Plk1 phosphorylation and activation are under the control of a $G_2$ checkpoint response to DNA damage, a similar experiment was performed except that after release from the aphidicolin block and before addition of nocodazole the cells were treated with nitrogen mustard for 30 minutes to cause DNA damage and delay at the $G_2$ checkpoint. In this experiment 17.5 hours after release, 90% of cells were in $G_2$/M, but less than 15% were in mitosis and no mobility shifts were observed in either Plk1 or $p34^{cdc2}$. There was gradual increase in Plk1 casein kinase activity over the time course which paralleled the increase in Plk1 protein; no abrupt increase in Plk1 casein kinase activity (which is presumably due to the activating phosphorylation) was observed. Thus Plk1 activation appears to be associated with the $G_2$/M transition and is blocked by activation of the $G_2$ checkpoint response to DNA damage. The sharp reduction in Plk1/STPK13 activity in cells exiting mitosis could be due to post-translational modification of the protein (perhaps by removal of the activating phosphate(s)) but Lee *et al.* (1995) found that Plk1/STPK13 activity correlated closely with the protein level in cells exiting mitosis and re-entering $G_1$ and suggested that post-translational modifications were probably not responsible for the decline in Plk1/STPK13 activity at this stage. A possible mechanism for the loss of Plk1/STPK13 protein would be through its localisation to the spindle microtubule overlap zone during anaphase and then to the midbody of the postmitotic bridge between the dividing cells (see below), a region that may be discarded as the dividing cells split apart.

The expression of the transcripts encoding polo-like kinases has also been examined. The abundance of the mRNA for Cdc5 was found to vary at least 20-fold in a cell cycle dependent manner in synchronised populations of *S. cerevisiae* (Kitada *et al.*, 1993). The mRNA was at a very low level in $G_1$ phase cells and increased during cell cycle progression to peak at $G_2$/M coincidentally with the peak in abundance of the mRNA for the Clb1 mitotic cyclin. The expression of STPK13 mRNA during the cell cycle was examined in NIH 3T3 cells by Lake and Jelinek (1993). In cell populations trapped in $G_0$ by serum starvation and then sampled at various time points after serum replenishment, STPK13 was first detected from cells at approximately the middle of S

phase and reached maximal levels at the end of S phase. The variation in STPK13 levels between $G_0$ and late S phase cells was at least 15-fold. Since the levels of STPK13 mRNA in the first cell cycle following serum replenishment might not be typical of those in subsequent cell cycles, STPK13 mRNA levels were also measured in cells sampled from a population synchronised by serum starvation followed by serum replenishment and subsequent addition of nocodazole to block the cells in prometaphase of the first mitosis. Non-adherent (mitotic) cells were then placed in fresh medium lacking nocodazole and Northern analysis was performed on samples of the blocked cells and cells collected at late $G_1$ and at S phase. STPK13 mRNA was detected in mitotic cells, was undetectable in $G_1$ phase cells, but could be detected in S phase cells confirming the results of the earlier experiment. A further experiment was also performed to test whether completion of S phase might be required for accumulation of STPK13 mRNA using the DNA polymerase inhibitor aphidicolin to block DNA synthesis. It was found that although cells blocked in S phase showed reduced STPK13 mRNA accumulation compared to untreated controls, STPK13 mRNA was nevertheless readily detectable. Therefore, accumulation of STPK13 mRNA does not appear to be obligatorily linked to completion of DNA synthesis. Cell cycle fluctuations in STPK13 mRNA abundance could be due to cell cycle regulation of transcription or to post-transcriptional processes. In order to investigate this, Lake and Jelinek (1993) isolated nuclei from serum starved ($G_0$) cells and from cells collected 6 and 24 hours after serum replenishment (corresponding to $G_1$ and $G_2$/M respectively) and analysed the production of STPK13 transcripts. The three sets of nuclei produced STPK13 RNA at similar rates indicating that the differences in abundance of the mRNA at different stages of the cell cycle must be due to post-transcriptional processes. Lee *et al.* (1995) also found that STPK13 mRNA was undetectable during G1 and early S phase but became detectable at late S phase. Consistent with the fact that STPK13 mRNA could not be detected in $G_0$ phase cells, mRNAs for STPK13/Plk1 have been found to be most abundant in tissues with a high proliferative index (including tissues from embryonic and newborn mice and from human cancers) but at a low or undetectable level in other tissues (Clay *et al.*, 1993; Lake and Jelinek, 1993; Golsteyn *et al.*, 1994; Hamanaka *et al.*, 1994; Holtrich *et al.*, 1994). In *Drosophila*, *polo* transcripts were also found to be abundant in tissues and in developmental stages in which there is a high level of mitotic activity. Particularly high

levels were detected in early embryos and in adult females presumably reflecting the requirements for cell cycle proteins in the growth of early embryos and for oogenesis. The expression of mRNA for mammalian polo-like kinases has also been examined in cell lines induced to differentiate. Lake and Jelinek (1993) examined STPK13 expression in murine erythroleukaemia cells (MELC) induced to differentiate and in a mutant MELC line that fails to differentiate. STPK13 poly(A)$^+$ mRNA levels decreased 6-fold in normal MELC induced for 12 hours, but were only slightly reduced in the mutant MELC. This decrease in poly(A)$^+$ mRNA levels was not mirrored by the total STPK13 RNA level which was at similar levels in cells induced for 12, 24 and 48 hours (total STPK13 RNA was measured in the cell cycle experiments described above). The explanation favoured by Lake and Jelinek (1993) for the loss of oligo(dT) binding was that it was due to shortening of the poly(A) tail. These authors were not able to measure STPK13 protein levels but shortening of the poly(A) tail is known to reduce the efficiency of translation of a number of mRNAs. After 12 hours of induction MELC cells are not committed to terminal differentiation so the decline in STPK13 poly(A)$^+$ mRNA abundance is a precommitment event. Hamanaka *et al.* (1994) found that when rat PC12 overexpressing the NGF receptor were induced to differentiate by treatment with NGF, rat plk mRNA expression was greatly reduced. Holtrich *et al.* (1994) found that in human T lymphocytes stimulated for 1 day with phytohaemagglutinin, Plk1 mRNA was induced between days 2 and 3 and that this induction could be blocked by cycloheximide. On this basis the *Plk1* gene is not a early growth response gene (see below).

The cellular localisation of polo and of Plk1/STPK13 has been examined by immunochemistry. The staining observed with antibodies to the polo protein kinase in *Drosophila* syncytial embryos was found to undergo cyclical changes (Llamazares *et al.*, 1991). Staining of interphase cells was mainly cytoplasmic but this became weaker in prophase and the antigen began to be associated with condensing chromatin. The association of polo with chromatin was strongest in metaphase and anaphase with cytoplasmic staining beginning to return in telophase. The subcellular localisation of Plk1 and STPK13 has been examined in HeLa cells and NIH 3T3 cells (Golsteyn *et al.*, 1994; Golsteyn *et al.*, 1995; Lee *et al.*, 1995). Golsteyn *et al.* (1995) performed a detailed analysis of the localisation of Plk1 at different stages of the cell cycle. Plk1 was diffusely distributed throughout interphase cells although the centrosomes were most strongly

stained. The spindle poles remained strongly stained throughout prophase and metaphase, however, during anaphase, the staining of the spindle poles disappeared but Plk1 could be detected at the equatorial plane in the spindle microtubule overlap zone. Plk1 staining was detected in the region of the cleavage plane during telophase and then near to the midbody of the postmitotic bridge. In contrast to the results of Llamazares *et al.* (1991) in *Drosophila*, no association of Plk1 with chromatin was observed. Lee *et al.* (1995) first examined the localisation of STPK13 in NIH 3T3 cells and obtained similar results to those of Golsteyn *et al.* (1995). They also noted that the localisation of STPK13 during M phase was similar to that of MKLP-1 (mitotic kinesin-like protein), a member of the kinesin superfamily (Nislow *et al.*, 1992). Microinjection of the CHO1 antibody which recognises MKLP-1 into $PtK_1$ fibroblast cells was shown to cause mitotic arrest in metaphase, the mitotic defects including failure of one or more chromosomes to migrate fully to the spindle equator and failure of some spindle microtubules to focus at the spindle poles (Nislow *et al.*, 1992). Immunoprecipitations were performed on S150 fractions from HeLa cells arrested in mitosis by nocodazole in order to investigate whether co-localisation of Plk1 and MKLP-1 might reflect interaction between the two proteins. MKLP-1 was found to co-precipitate with Plk1 and Plk1 could phosphorylate MKLP-1 in kinase reactions performed *in vitro* on the immunoprecipitates. It was also found that, after denaturation of the immunocomplexes and subsequent immunoprecipitations, the Plk1 antibody only precipitated Plk1 and did not co-precipitate MKLP-1 while the CHO1 antibody only precipitated MKLP-1. Therefore, co-precipitation of the two proteins is not due to cross-reactivity of the Plk1 antibody. Thus it appears that MKLP-1 associates with Plk1 and may well be an endogenous substrate for this protein kinase.

The murine polo-like kinases Snk and Fnk (Simmons *et al.*, 1992; Donohue *et al.*, 1995) have not been as well studied as STPK13 and, in particular, no work has been performed with the proteins. The gene for Snk was identified via cloning of a partial Snk cDNA artifactually ligated to another cDNA under investigation. The section of the chimaera which was later used to isolate full length Snk clones was examined further because it was found to be responsible for the hybridisation of the chimaeric cDNA to an RNA strongly induced in mitogen-stimulated murine F-2 cells. Fnk was identified in a search for genes induced by FGF in fibroblasts using targeted differential display with

one PCR primer based on protein kinase catalytic domain VIb and the other based on zinc-finger domains (Fnk does not in fact contain a zinc finger domain). Snk mRNA was found to be strongly induced (10 to 20-fold) within 1 hour after stimulation of serum-starved NIH 3T3 cells with serum or with phorbol myristate acetate. This induction was transient, the Snk mRNA returned to basal levels within 6 hours following mitogen stimulation and the level of induction was increased when the translational inhibitor, cycloheximide, was added to the cell culture at the same time as the serum or phorbol ester. Nuclear run-on analysis showed that serum induction of Snk mRNA occurred at the transcriptional level. The synthetic glucocorticoid dexamethasone prevented induction of Snk mRNA in living cells and in run-on analysis but did not affect basal expression. The mRNA for Fnk was found to be rapidly induced following exposure of serum-starved NIH 3T3 cells to FGF-1 with maximum levels at 1 hour after stimulation. Expression had returned to basal levels by 8 hours. FGF-2 which is related to FGF-1 had similar effects on Fnk expression as did PDGF-BB, serum and phorbol myristate acetate, while TGF-β1, EGF and IGF-1 had little effect. The RNA synthesis inhibitor actinomycin D prevented induction of Fnk by FGF-1 showing that induction of Fnk mRNA expression by FGF-1 is at least partly due to transcriptional activation of the *fnk* gene. As for Snk, addition of a protein synthesis inhibitor (cycloheximide) with the mitogen increased the level of induction of Fnk mRNA expression. On the basis of these results both Snk and Fnk can be characterised as early-growth-response genes (also known as immediate-early genes) based on the fact that their transcription is rapidly and transiently upregulated following mitogen stimulation. The lack of dependence of this increase in transcription on protein synthesis is a common feature of many of these genes. Snk mRNA could be detected at fairly low levels in lung, heart and brain tissues from an adult male mouse but not in thymus, spleen, liver, kidney and intestine while Fnk mRNA was found to be expressed at a low level in the heart of newborn animals but at significant levels in intestine, kidney, liver, lung and skin. In adult tissues only skin showed a high level of expression.

To summarise, genes encoding protein kinases related to polo have recently been isolated from organisms ranging from mammals to yeasts and, in this work, from the evolutionarily distinct kinetoplastid *T. brucei*. There is significant evidence to suggest that polo, Cdc5, plo1, Plk1, STPK13 and rat plk are functional homologues but a

number of differences should also be noted. Casein kinase associated with polo immunoprecipitated from *Drosophila* syncytial embryos was maximal at late anaphase and telophase at which point cdc2 kinase activity was low (Fenton and Glover, 1993), while Plk1/STPK13 was found to be activated at the $G_2$/M transition (Golsteyn *et al.*, 1995; Hamanaka *et al.*, 1995; Lee *et al.*, 1995). The polo kinase was also reported to be associated with condensed chromosomes at certain points in the cell cycle (in both syncytial and cellular embryos) while no association with chromosomes has been reported for Plk1. At present not sufficient is known about polo and Plk1/STPK13 for these data to be reconciled but it would not be surprising for cell cycle events to be regulated differently in syncytial embryos compared to mammalian cells grown in culture. Another difference is found between the cellular phenotypes observed in the *S. cerevisiae* and *S. pombe* switch-off experiments. The gradual removal of Cdc5 in these experiments led to failure in late nuclear division with the arrested cells having an intact spindle. Similar experiments in *S. pombe* showed that bipolar spindle formation was defective and, as described above, polo and Plk1/STPK13 are also implicated in spindle regulation. Ohkura *et al.* (1995) suggested that the different phenotypes observed could reflect the fact that in *S. cerevisiae* bipolar spindle formation begins during S phase while in *S. pombe* S phase is completed before spindle assembly. Another difference between the cell cycles of the two yeasts is that actin ring formation in *S. cerevisiae* occurs after spindle elongation while in *S. pombe* the two events are approximately simultaneous (see references in Ohkura *et al.*, 1995) although actin ring formation was not specifically examined in the *CDC5* switch off experiments. It seems likely that polo-like kinases will have multiple substrates and while some are probably shared between divergent eukaryotes, others are probably not; this is likely to relate to differences in the organisation of the cell cycle in different organisms.

Future work with polo-like kinases is likely to involve the identification of substrates and more detailed studies into the regulation of protein kinase activity. The only known substrate for a polo-like kinase is the human kinesin-related protein (KRP) MKLP-1. However, another KRP, cut7, has been suggested as a likely substrate for plo1 in *S. pombe* and given that a number of KRPs are known to be involved in spindle function (see reviews by Moore and Endow, 1996 and Barton and Goldstein, 1996) it would not be surprising if KRPs are among the substrates of other members of the polo-

like kinase family. Plk1/STPK13 appear to be substrates for a separate protein kinase, and identification of proteins other than substrates with which polo-like kinases interact will also be an important area of work. It seems possible that the highly conserved PH2 domain may be involved in an interaction between polo-like kinases and another conserved protein. Another question which awaits investigation is the cellular role of the Snk and Fnk polo-like kinases which have not yet been examined at the protein level.

### 5.3.4 Future studies into the cellular function of TbPLK

The available data on polo-like kinases in higher eukaryotes (Section 5.3.3) suggests that TbPLK will be an interesting protein to study in *T. brucei*. Of the known polo-like kinases there is evidence implicating polo, plo1, Plk1, STPK13 (and by implication rat plk) in regulation of the centrosome and mitotic spindle during mitosis. Cdc5 is also implicated in control of mitotic progression (*CDC5* disruptants failed in late nuclear division in switch off experiments) but may not act on the mitotic spindle. Therefore it seems likely that TbPLK will be involved in mitotic regulation in *T. brucei*. As described briefly in Section 1.1.4, *T. brucei* is one of the best-studied examples of a lower eukaryote in which cell division is asymmetrical and occurs without extensive disassembly of the cytoskeleton. The mitotic spindle has only been visualised in electron micrographs (Sherwin and Gull, 1989a, Robinson *et al.*, 1995) but it was shown in these studies that the *T. brucei* spindle is unusual in that it persists after the daughter nuclei have separated, perhaps acting to place the daughter nucleus into the gap left between the basal bodies after they have moved apart (see Fig. 1.3 panel A and Robinson *et al.*, 1995). The study of the *T. brucei* spindle would also be interesting with respect to the question of how the numerous minichromosomes are segregated. It seems unlikely that each minichromosome could have its own microtubule and so minichromosomes may not be segregated by the mitotic spindle. While it seems probable that TbPLK will be involved in regulating spindle function, there is evidence for the involvement of polo-like kinases in regulating other processes involved in cell division; in particular, plo1 in *S. pombe* appears to be the key regulator of septum formation (both formation of the actin ring and deposition of septal material). At present there is no evidence for an actin ring in *T. brucei* although actin genes are transcribed (BenAmar *et al.*, 1988). The microtubular structures of *T. brucei* are quite different to those of higher eukaryotes and are potential

drug targets; differences in sensitivity between *T. brucei* and mammalian cells to drugs that disrupt microtubules have been reported (Seebeck and Gehr, 1983). The further study of proteins that may be involved in regulating microtubule structures such as the mitotic spindle is therefore interesting in this respect. In summary, although no data have been obtained regarding the cellular function of TbPLK, most of its homologues from higher eukaryotes appear to be involved in control of mitosis and its further characterisation could complement earlier morphological studies of the *T. brucei* cell cycle.

One of the most immediate aims in any future work would be confirm the coding sequence for TbPLK as the 3′ coding sequence is incomplete and, although the 5′ coding sequence could be complete (with the open reading frame starting at position 377 in Block B (see Fig. 5.9)), it is equally likely that there is further 5′ coding sequence not contained within pTbplk1 (see Section 5.2.4.3). Ideally, a full length cDNA would be obtained by library screening or alternatively 3′ and 5′ RACE might be used with the improvements suggested in Section 5.3.1. PCR amplification could be used with a library as template with gene-specific and vector primers. The isolation of genomic clones for TbPLK would be another possibility. Many of the experiments that might be performed to investigate the cellular function of TbPLK would depend on the availability of antibodies to the protein and the raising and characterisation of antisera to TbPLK would be another priority.

Methods have recently been developed for the transfection of *T. brucei* that would allow reverse genetics to be used in investigating the cellular function of TbPLK. One type of experiment would be to examine the cellular effects of loss of function of *tbplk* and of overexpression of *tbplk*. It seems likely that *tbplk* will be an essential gene in *T. brucei* and so it would probably be necessary to knockout one copy of *tbplk*, then express the gene under the control of an inducible promoter from an expression construct either integrated into the genome or as an episome followed by knocking out the second genomic copy before removing the inducer and examining cellular phenotype. Wirtz and Clayton (1995) have described a system that would appear to be suitable. A procyclic cell line expressing the *E. coli* tetracycline repressor from a construct introduced into the tubulin locus is used as the genetic background. The gene of interest is ligated into an expression construct that includes an upstream PARP promoter

(mutated to include the TetR operator sequences) and downstream selectable marker (phleomycin resistance) with the ends of the construct coming from an rDNA nontranscribed spacer region to target the construct to this (transcriptionally inactive) region of the genome (stable transformation of *T. brucei* occurs by homologous recombination). Using a luciferase reporter gene Wirtz and Clayton (1995) found that the level of expression could be controlled over a range of four orders of magnitude in response to tetracycline concentration but there is a possible problem with the present system for genes that express toxic products. As the gene of interest and the selectable marker gene are under control of the same promoter, selection of transfectants has to be carried out in the presence of tetracycline leading to the expression of the gene under study. In view of the work performed with polo-like kinases in other systems, it seems likely that loss of function of *tbplk* would lead to defects in the mitotic spindle and the results of an experiment performed by Robinson *et al.* (1995) using the antimicrotubule agent rhizoxin (which had been shown to affect mitosis and subpellicular microtubule events more severely than flagellum growth) suggest a possible phenotype for disruptants in *tbplk.* In this experiment, exponentially growing procyclic cells were treated with a low concentration of the drug and the phenotypes of cells examined at various time points. It was found that, in such cultures, about 30% of the cells had a kinetoplast but were anucleate (1K0N, these cytoplasts were termed zoids) while the proportion of cells with one kinetoplast and one nucleus (1K1N) was greatly reduced compared to control cultures; in addition, a second unusual cell type with one kinetoplast and two nuclei (1K2N) was also observed. The 1K2N and 1K0N cells were assumed to be siblings but, after approximately one cell cycle period, the number of 1K0N zoids was about twice the number of 1K2N cells suggesting that more than one type of cell division was leading to zoid production. Closer examination of the 1K1N cells from the drug treated culture showed that the nuclei of these cells had a higher DNA content than those of 1K1N cells from the control culture. The interpretation of these results was that in the divisions producing one 1K2N cell and one 1K0N zoid, mitosis occurred producing two nuclei but that nuclear segregation was defective. The second type of division producing 1K0N zoids was assumed to involve the 1K1N cells with elevated nuclear DNA content as the sibling to the zoid. This type of division presumably results from failure of mitosis at an earlier stage (i.e. before two nuclei can be observed by staining) but nevertheless after

nuclear DNA replication. Thus the two types of division postulated are essentially similar and could be explained by defects of differing severity in the mitotic spindle. It seems possible that *tbplk* disruptants would show similar morphologies although this would depend on whether TbPLK mainly or exclusively acts on the mitotic spindle or whether it has a wider role. If the latter was the case, then cytokinesis might also be blocked in TbPLK disruptants and there might be other morphological defects. Given the large range over which Wirtz and Clayton (1995) were able to control the level of expression of a reporter gene it should be possible to examine the cellular effect of different degrees of underexpression and overexpression of TbPLK. "Dominant negative" strategies in which a mutant protein that may interfere with the function of the wild-type protein is expressed in a wild-type background might also be useful in examining TbPLK function.

Using antibodies, it would be possible to determine the subcellular localisation of TbPLK. As described in Section 5.3.3, the localisation of polo and of Plk1/STPK13 has been shown to change during the cell cycle and, since the procyclic cell cycle has been well characterised by microscopy, a similar analysis could be performed for TbPLK. An association of TbPLK with the mitotic spindle seems likely but TbPLK might also be associated with other microtubule structures and, in particular, an association of TbPLK with the basal bodies or associated microtubules seems possible. Once the subcellular localisation of the wild-type protein during the cell cycle had been determined, it would be interesting to compare this with the localisation of mutant forms of TbPLK. For instance, if the conserved PH2 domain is involved in an interaction between TbPLK and another protein then mutation or deletion of this domain might lead to the mutant protein having a different subcellular localisation to the wild-type protein. Since it would be desirable to be able to distinguish directly between mutant and wild-type proteins in experiments of this type it would be useful to have developed a system for expression of an epitope-tagged form of TbPLK and to shown similar localisation of the tagged and wild-type proteins so that two different antibodies could be used to detect wild-type and mutant proteins independently.

Antibodies could also be used together with nucleic acid probes, in an analysis of the level of expression of TbPLK protein and mRNA respectively. It would clearly be of interest to examine the expression of TbPLK and its mRNA both between life cycle stages and between cells from the same life cycle stage at different points in the cell

cycle. The expression of TbPLK protein and message could be examined in slender bloodstream, stumpy bloodstream, procyclic and metacyclic forms by Western and Northern analysis respectively. Since TbPLK is likely to be cell cycle regulated both proliferating and stationary phase cultured procyclic forms should be included in this analysis. It would be expected that TbPLK protein would be detectable in slender bloodstream and proliferating procyclic forms but at a low or undetectable level in the non-proliferating stumpy, metacyclic and stationary phase procyclic forms. In general *T. brucei* mRNA levels correlate with the level of expression of the protein and it would seem likely that the level of TbPLK message would be lower in non-proliferating forms than in proliferating forms. A 5.4 kb poly(A)$^+$ RNA hybridising to TbPK-A2 was found to be at similar levels in slender bloodstream and procyclic forms in a Northern analysis performed by Hua and Wang (1994) but the level of expression of the RNA has not yet been examined in any non-proliferating cell type. Ideally the levels of both hybridising poly(A)$^+$ and total hybridising RNA would be measured in future analyses.

Analysis of the expression of TbPLK mRNA and protein during the cell cycle would be more technically demanding than a comparison between life cycle stages because of the need to obtain populations of cells at similar stages of the cell cycle using either cell synchronisation or cell sorting methods. Gale *et al.* (1994a) found that stationary phase cultured procyclic forms, stimulated to proliferate by addition of fresh medium, underwent the first cell cycle in a semi-synchronous manner as assessed by flow cytometric analysis of cellular DNA content. This synchronisation method could be used to assess the cell cycle expression of TbPLK protein and mRNA at different cell cycle stages, however, as noted by Gale *et al.* (1994a), it is possible that the first cell cycle in previously quiescent cells may not be typical of later cell cycles. Nutrient deprivation has also been used to synchronise culture-adapted bloodstream forms but the yield obtained was low (Morgan *et al.*, 1993). Affranchino *et al.* (1993) used hydroxyurea treatment to synchronise procyclic cells at the $G_1$/S transition and this would be another possible approach to obtaining synchronous cultures. The alternative to the production of synchronous cultures of *T. brucei* is the sorting of asynchronous populations by flow cytometry (density centrifugation methods are not applicable due to the high motility of *T. brucei*). Although this method allows the use of a broader range of cell types, at present the only satisfactory method for sorting cell populations on the basis of cell cycle

stage is by analysis of DNA content using propidium iodide staining. The method for propidium iodide staining requires methanol fixation of the cells and RNase treatment. The latter treatment precludes Northern analysis and although Western analysis could be performed on cells sorted using this method, other analyses that might be desired, such as assay of protein kinase activity would probably not be possible due to protein denaturation (although Gale and Parsons, 1994a were able to assay the activity of renaturable protein kinases from cells stained in this way). It would be interesting to compare the expression of TbPLK with that of other putative cell cycle proteins such as the *T. brucei* cdc2-related protein kinases (Mottram and Smith, 1995) and the mitotic-like cyclin homologue CYC1 (Affranchino *et al.*, 1993) although none of these are well characterised with respect to cellular function. The expression of TbPLK protein and mRNA could also be examined during the stumpy to procyclic transition performed *in vitro* a process that occurs coincidentally and synchronously with re-entry into a dividing cell cycle (Matthews and Gull, 1994b). It might be expected that TbPLK would be at a low or undetectable level in stumpy forms but that levels would increase at some point in the first cell cycle.

The biochemical properties of TbPLK could be investigated using immunoprecipitated protein from *T. brucei* cells, some experiments might also be performed with TbPLK overexpressed in another system. All of the polo-like kinases studied to date use casein as substrate and, given that TbPLK has similar residues to those of other polo-like kinases at positions likely to be important for substrate binding (Section 5.2.4.3), it seems likely that TbPLK will be active as a casein kinase. The results of Hamanaka *et al.* (1995) suggest that a protein kinase phosphorylates and activates Plk1/STPK13 at the $G_2$/M transition and it would be of interest to determine whether TbPLK is also phosphorylated *in vivo* and at what stages of the cell cycle by immunoprecipitating TbPLK from cultures grown in medium containing [$^{32}$P]orthophosphate. At least some of the polo-like kinases also autophosphorylate. It would be desirable to determine the sites of any such phosphorylations. If large amounts of active TbPLK could be obtained from an expression system such as Sf9 insect cells (as used for Plk1 by Golsteyn *et al.*, 1995 and Lee *et al.*, 1995) it should be possible to determine the site(s) of any *in vitro* autophosphorylation by phosphopeptide mapping, however the determination of sites phosphorylated *in vivo* would probably be difficult

due to smaller amounts of protein. An alternative approach would be transfection of *T. brucei* with constructs expressing mutant forms of TbPLK with substitutions at the putative phosphorylated residues and an epitope tag to allow isolation separately from wild type TbPLK. It is likely that phosphorylated sites in polo-like kinases from higher eukaryotes will be identified in the future and this knowledge should aid the investigation of the positions of any phosphorylated sites in TbPLK.

While reverse genetic and immunolocalisation experiments may well give an indication of the cellular role of TbPLK, in the longer term it will be important to identify proteins (particularly substrates) that TbPLK interacts with. One way of achieving this would be through co-immunoprecipitation of TbPLK and interacting proteins followed by identification of co-immunoprecipitated proteins by partial protein sequencing. However, preparation of sufficient amounts of protein(s) interacting with TbPLK for protein sequencing is likely to be very difficult. An alternative approach to the identification of proteins interacting with TbPLK would be to use the yeast two-hybrid system but this is also technically difficult. However, as described in Section 5.3.3, Lee *et al.* (1995) have reported strong evidence that a member of the kinesin superfamily of motor proteins, MKLP-1, is an endogenous substrate for Plk1 while Ohkura *et al.* (1995) found that disruptants of *plo1*$^+$ showed similar phenotypes to those of mutants in *cut7*$^+$ which encodes a kinesin-related protein (KRP) as well as reporting a genetic interaction between *plo1*$^+$ and *cut7*$^+$. Therefore, it would not be unreasonable to perform experiments to test whether TbPLK showed any interaction with KRPs. One way in which this might be performed would be to immunoprecipitate TbPLK and then perform Western blot analysis with antibodies against KRPs to test for co-immunoprecipitation. No genes encoding *T. brucei* KRPs have been isolated although a gene for a KRP has been cloned from *Leishmania chagasi* (Burns *et al.*, 1993b). The motor domain of KRPs (about 340 amino acids) includes some highly conserved subdomains and antibodies against peptides corresponding to two motor subdomains (which are reasonably well conserved in the *Leishmania* KRP) have been used to identify proteins that showed biochemical properties similar to those of KRPs from organisms as divergent as Xenopus (Sawin *et al.*, 1992) and Chlamydomonas (Fox *et al.*, 1994) as well as being used to isolate a Xenopus cDNA encoding a KRP from an expression library (Walczak *et al.*, 1996). It should be possible to use such antibodies to probe Western blots of protein

immunoprecipitated by antibodies to TbPLK. If the results of such analysis suggested that TbPLK did bind to one or more KRPs, it is less clear how such a protein would be identified more specifically. The only identified polo-like kinase substrate, human MKLP-1, cannot be grouped with any of the known KRP subfamilies (see review by Moore and Endow, 1996) and the only other known member of the outgroup is from another mammal (hamster). On this basis, it cannot at present be predicted whether lower eukaryotes are likely to have KRPs related to MKLP-1, although the other KRP currently suggested as a polo-like kinase substrate, cut7 (*S. pombe*), is from a defined subfamily of KRPs (the BimC subfamily) that has members from both multicellular and unicellular organisms. If evidence was obtained for an interaction of TbPLK with one or more KRPs, it would be possible to clone the genes for *T. brucei* members of this superfamily. This could be performed by PCR or library screening with degenerate oligonucleotides based on the peptide sequences of conserved motor subdomains or by expression library screening with antibodies to these subdomains.

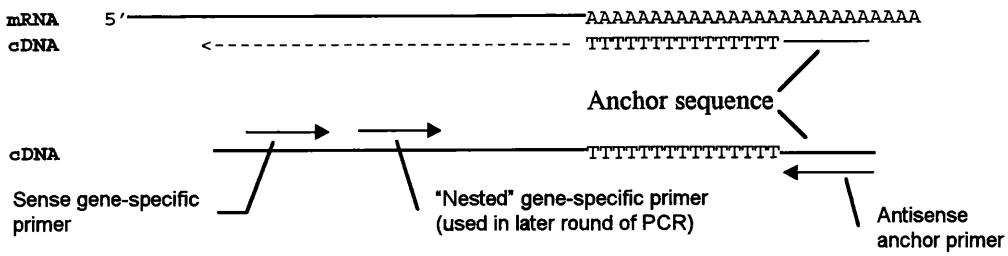
In summary, on the basis of homology to polo-like kinases in higher eukaryotes TbPLK is likely to be involved in regulating the mitotic spindle in *T. brucei* and perhaps in other aspects of cell cycle control. Cell division in this organism is very different to that in both multicellular eukaryotes and in the yeasts *S. cerevisiae* and *S. pombe* but has been relatively well studied with respect to changes in the subpellicular cytoskeleton. Examination of the cellular function of Tbplk should complement both these studies and recent studies into cdc2-related kinases in *T. brucei* (see Section 1.3.2.2). *T. brucei* is at present the most evolutionarily divergent organism from which a polo-like kinase homologue has been identified and the further study of TbPLK would also be interesting in terms of determining the extent of conservation of function among the polo-like kinases.

**Fig. 5.1** Schematic diagram showing the main features of the 3' RACE method. **(A)** cDNA synthesis; **(B)** PCR.

**Fig. 5.2** Strategy for 5' RACE using a 5' primer based on the *T. brucei* spliced leader **(A)** cDNA synthesis; **(B)** PCR.

**Fig. 5.3** Primers used in PCR reactions. The figure shows schematically the placement of the primers on a cDNA that includes TbPK-A2. The boxed section of the line indicates TbPK-A2 (not to scale), the two arrows in bolder type represent the original degenerate primers at the ends of TbPK-A2.

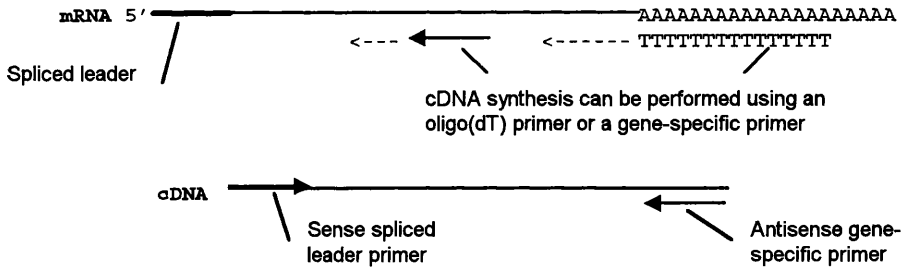
A




B

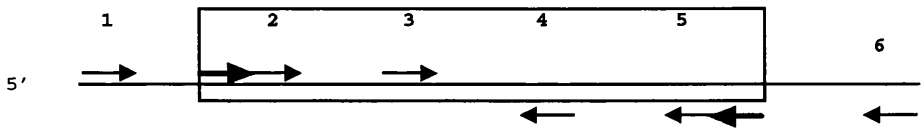
**Fig. 5.1**

A




B

**Fig. 5.2**

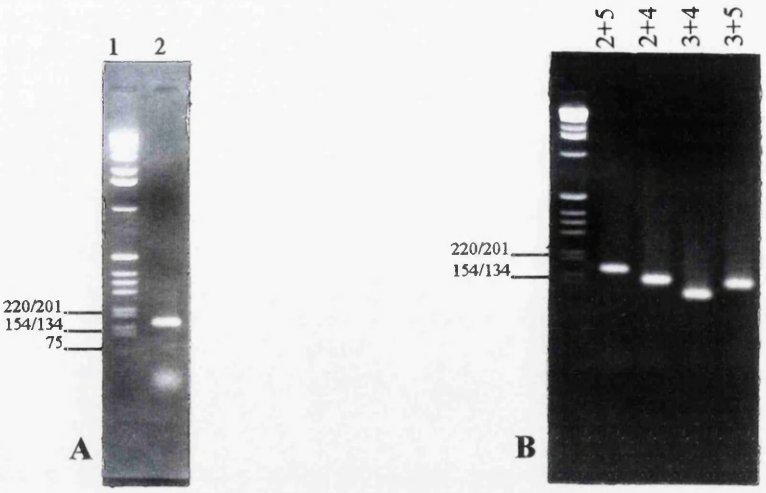



| | |
|---|---|
| Primer 1: sense primer based on the spliced leader | 5' TTAGAACAGTTTCTGTACTATA 3' |
| Primer 2: sense gene-specific primer | 5' TGACAACATAATGATGGATGCA 3' |
| Primer 3: "nested" sense gene-specific primer | 5' AAGATTAGGGACTTCGGTTTGG 3' |
| Primer 4: "nested" antisense gene-specific primer | 5' GCGAACCTTCAATGATCTCCGG 3' |
| Primer 5: antisense gene-specific primer | 5' ACGTCCACTTCGTAACTATGTC 3' |
| Primer 6: anchor antisense primer | 5' CGAGAATTCGGTGGCAGCAGCCAACT 3' |

**Fig. 5.3**

**Fig. 5.4** Amplification of a cDNA fragment homologous to TbPK-A2. Panel A shows a 2% agarose gel stained with ethidium bromide containing products of 30 cycles of PCR using primers 2+5 with a cDNA template. The strongly stained band in Lane 2 is of the size expected (171 bp) for a fragment similar to TbPK-A2. Panel B shows an agarose gel containing products of PCR reactions with aliquots of the primer 2+5 product as template. The primer combinations used are indicated above each lane. The major products are of the sizes predicted (171, 143, 109 and 137 bp for the products generated using primers 2+5, 2+4, 3+4 and 3+5 respectively). The major product generated using primers 2+5 was cloned and sequenced. Panel C shows the nucleotide and predicted peptide sequences of this PCR product and the original clone TbPK-A2 compared with a similar sequence (labelled Tbplk1) from a λgt11 cDNA clone (Section 5.2.4.1). Dashes indicate identity with the sequence from the λgt11 clone while PCR primers are underlined. Note that the 3' primers (bottom left) are shown as their complements (sense strand) but are actually antisense primers. The primer 2+5 product and TbPK-A2 differ at a single nucleotide position internal to the primer regions at which the sequence of the λgt11 clone agrees with that of the primer 2+5 product. As might be expected there are further differences between the sequence of the λgt11 clone and TbPK-A2 within the region of the degenerate primers used to amplify TbPK-A2.

**Fig. 5.5** Amplification of the 5' cDNA sequence. Panel A shows a gel loaded with the products of 30 cycles of PCR using the modified 5' RACE procedure shown in Fig. 5.2. The primers were 1 (spliced leader primer) and 5 (gene specific) and a cDNA template was used. The annealing temperature in the PCR reactions was 47 °C and the extension time was 2 minutes. Repeating the reaction with annealing temperatures between 38 °C and 50 °C made little difference to the results (not shown). Products of the 5' RACE reaction shown in panel A were re-amplified from agarose gel slices with primers 1 and 5. In order to test whether any of these reactions contained the specific product desired aliquots were run out in an agarose gel (panel B) which was blotted, and the blot was probed with the cloned fragment similar to TbPK-A2 (panel C shows the autoradiograph of the blot). Lanes 1, 2, 4, 5, 7, 9, and 11 are empty, lane 3 contains a portion of the DNA that was used to make the probe as a positive control, while lanes 6, 8, and 10 contain aliquots of the different PCR reactions performed with products of the original 5' RACE reaction as template.




```
PRIMER 2+5                  --AC----------------------------------
Tbplk1      CGTGACCTGAAACTTGGTAACATAATGATGGATGCAAATATGAACGTGAA
TbPK-A2     --G-----T--G----AC----------------------------------
Tbplk1/2+5  R  D  L  K  L  G  N  I  M  M  D  A  N  M  N  V  K
TbPK-A2     -  -  -  -  -  D  -  -  -  -  -  -  -  -  -  -  -

PRIMER 2+5  ----G-----------------------------------------------
Tbplk1      GATTGGGGACTTCGGTTTGGCTGCTGAGTTGCAGTACGACGGGGAGCGGA
TbPK-A2     ----A-----------------------------------------------
Tbplk1/2+5   I  G  D  F  G  L  A  A  E  L  Q  Y  D  G  E  R  K
TbPK-A2      -  R  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -

PRIMER 2+5  -------------------------------------------------
Tbplk1      AACGCACTATTTGTGGCACGCCCAATTATATCGCACCGGAGATCATTGAA
TbPK-A2     -------------------------------------------------
Tbplk1/2+5    R  T  I  C  G  T  P  N  Y  I  A  P  E  I  I  E
TbPK-A2       -  -  -  -  -  -  -  -  -  -  -  -  -  -  -  -

PRIMER 2+5  ----------------------------C--
Tbplk1      GGTTCGCGCGAGGGACATAGTTACGAAGTGGATGTGTGGTCGCTCGGTGT
TbPK-A2     -------------------------------C--A---G-CT-------
Tbplk1/2+5  G  S  R  E  G  H  S  Y  E  V  D  V  W  S  L  G  V
TbPK-A2     -  -  -  -  -  -  -  -  -  -  -  -  -  A  P  -  -
```

C

**Fig. 5.4**

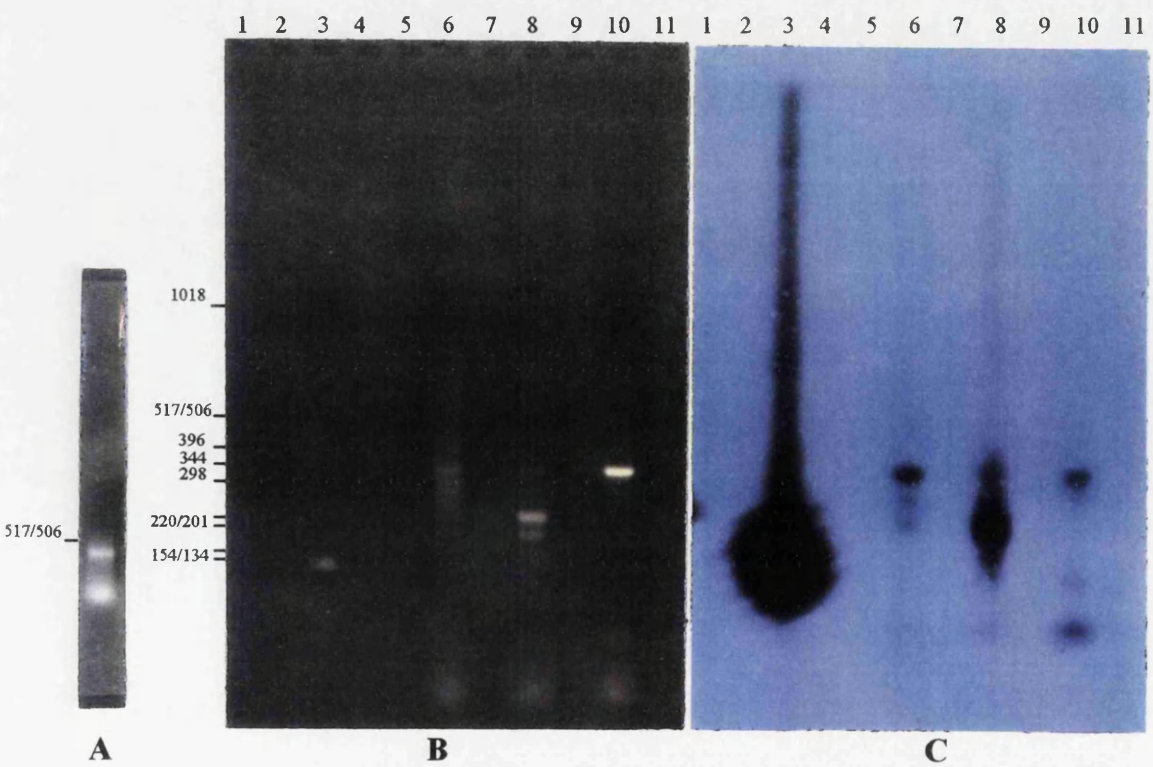



**Fig. 5.5**

**Fig. 5.6** Amplification of the 3' cDNA sequence. Panel A shows an agarose gel containing products of a 3' RACE reaction between primer 2 and an oligo(dT) primer. Lane 1 contains DNA size markers, lane 2 contains the products of a control reaction, while lane 3 contains the products of the 3' RACE reaction. It can be seen that there are no visible product bands of significant size, a possible reason for the failure of the reaction being that the lower annealing temperatures that have to be used with an oligo(dT) primer were reducing reaction specificity. In order to address this a new batch of cDNA was made using an adapter primer so that 3' RACE could be performed with a 3' anchor primer (primer 6) and gene specific primers 2 or 3 at higher annealing temperatures. Panel B shows the products of 3' RACE reactions using primers 3 and 6 at annealing temperatures of 56 °C, 58 °C, and 60 °C. The annealing temperatures used are indicated above each lane while the positions and sizes of DNA markers are indicated to the left. Two of the products from the 56 °C reaction were cloned (marked on figure). Panel C shows products of 3' RACE reactions between primer 2 and primer 6, and between primer 3 and primer 6. Lane 1 contains DNA size markers, lanes 2 and 3 are duplicate lanes containing products of primers 2+6, while lanes 4 and 5 are duplicate lanes containing products of primers 3+6.

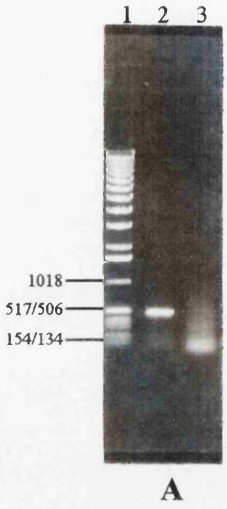
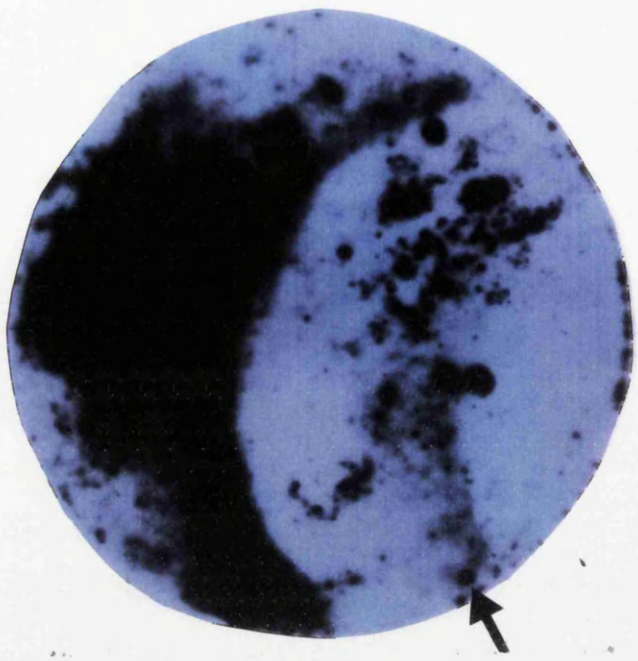
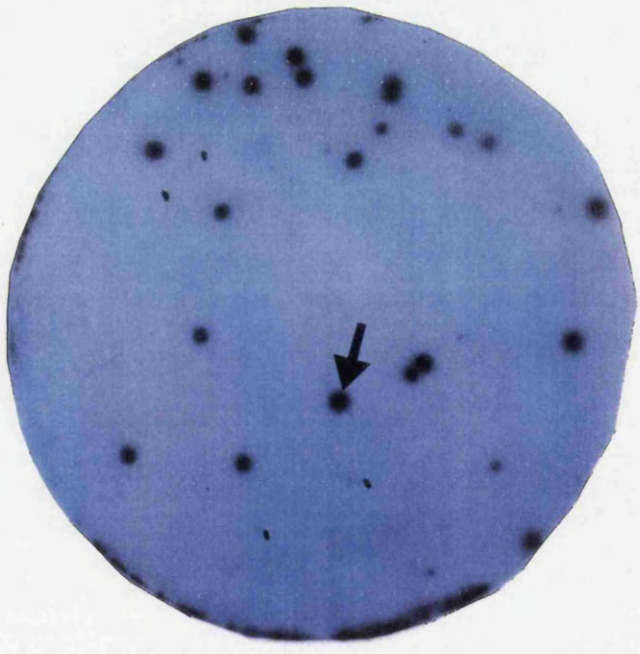

1 2 3
1018
517/506
154/134
A
60 °C
58 °C
56 °C
1636
1018
517/506
154/134
Reaction products cloned
B
1 2 3 4 5
1636
1018
517/506
396
298
220/201
154/134
75
C

**Fig. 5.6**

**Fig. 5.7** Screening of a λgt11 cDNA library with the PCR product homologous to TbPK-A2. In the primary screen only a single plaque hybridised to the probe on a confluent plate and was picked for further rounds of screening (not shown). Panel A shows the autoradiograph of a plate from the secondary screen from which the plaque indicated by the arrow was picked for further screening. In the tertiary screen (panel B) 22 plaques hybridised to the probe from a total of 78. The plaque that was picked for the production of larger quantities of the phage, isolation of DNA, and subcloning of the insert into a bacterial plasmid is arrowed.



A



B

Fig. 5.7

**Fig. 5.8** Restriction mapping and subcloning of pTbplk1. The insert (2.8 kb) from the λ clone was cut out with Eco RI and cloned into the Eco RI site of pGEM-7Zf(+) to give plasmid pTbplk1. A restriction map was made of pTbplk1 and subclones were constructed as shown. The sequencing strategy used is explained in the text.

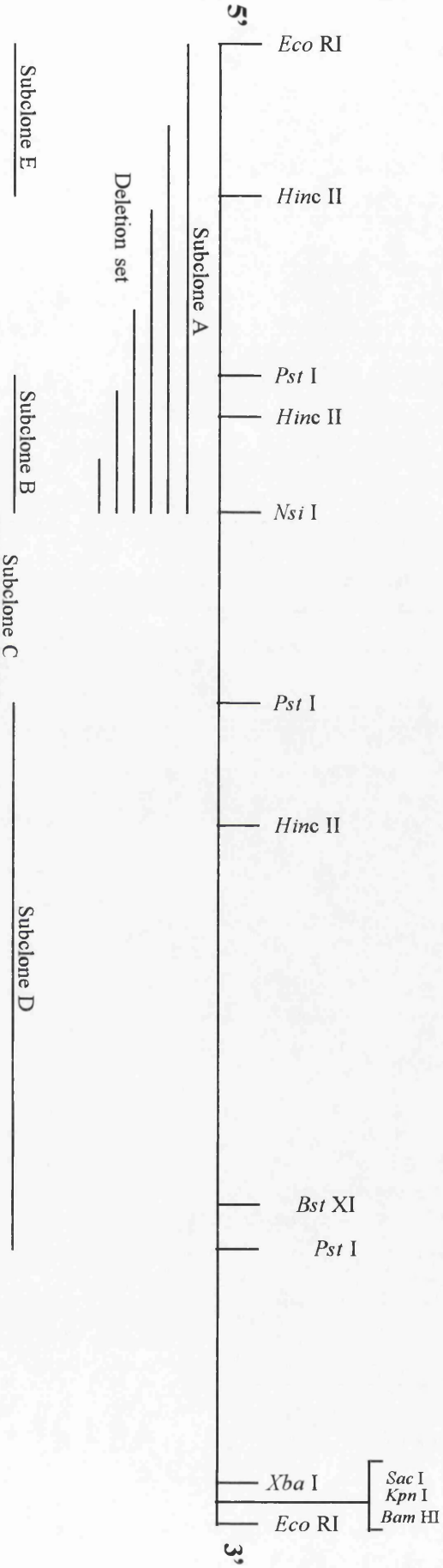
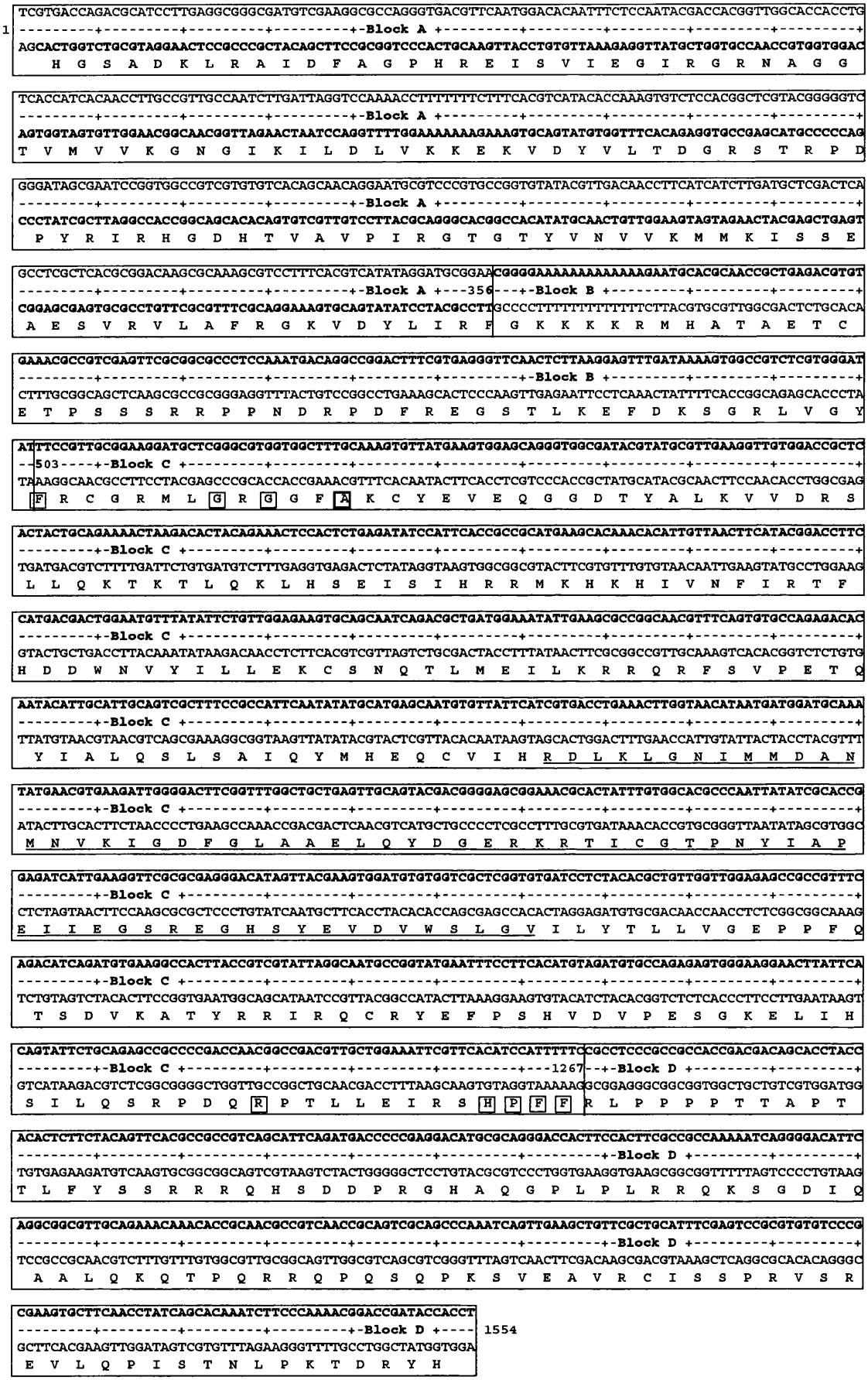
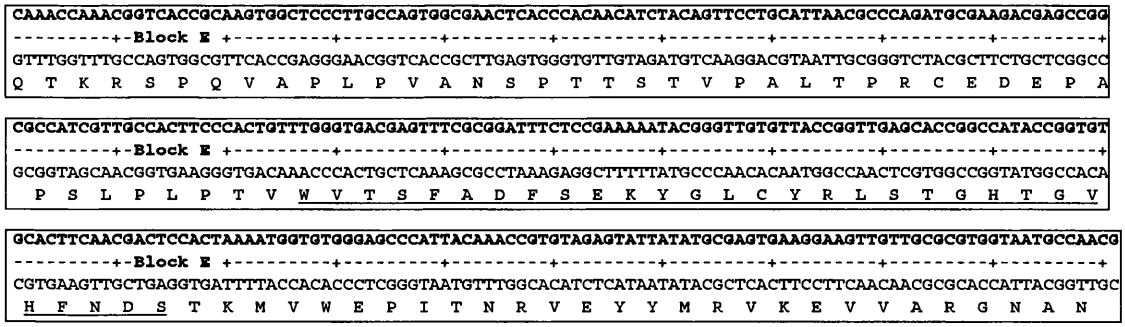

5'
Eco RI
Hinc II
Pst I
Hinc II
Nsi I
Pst I
Hinc II
Bst XI
Pst I
Xba I
Sac I
Kpn I
Bam HI
Eco RI
3'
Subclone A
Deletion set
Subclone E
Subclone B
Subclone C
Subclone D

**Fig. 5.8**

**A**


TCGTGACCAGACGCATCCTTGAGGCGGGCGATGTCGAAGGCGCCAGGGTGACGTTCAATGGACACAATTTCTCCAATACGACCACGGTTGGCACCACCTG
Block A
AGCACTGGTCTGCGTAGGAACTCCGCCCGCTACAGCTTCCGCGGTCCCACTGCAAGTTACCTGTGTTAAAGAGGTTATGCTGGTGCCAACCGTGGTGGAC
H G S A D K L R A I D F A G P H R E I S V I E G I R G R N A G G
TCACCATCACAACCTTGCCGTTGCCAATCTTGATTAGGTCCAAAACCTTTTTTTTCTTTCACGTCATACACCAAAGTGTCTCCACGGCTCGTACGGGGGTC
Block A
AGTGGTAGTGTTGGAACGGCAACGGTTAGAACTAATCCAGGTTTTGGAAAAAAAGAAAGTGCAGTATGTGGTTTCACAGAGGTGCCGAGCATGCCCCCAG
T V M V V K G N G I K I L D L V K K E K V D Y V L T D G R S T R P D
GGGATAGCGAATCCGGTGGCCGTCGTGTGTCACAGCAACAGGAATGCGTCCCGTGCCGGTGTATACGTTGACAACCTTCATCATCTTGATGCTCGACTCA
Block A
CCCTATCGCTTAGGCCACCGGCAGCACACAGTGTCGTTGTCCTTACGCAGGGCACGGCCACATATGCAACTGTTGGAAGTAGTAGAACTACGAGCTGAGT
P Y R I R H G D H T V A V P I R G T G T Y V N V V K M M K I S S E
GCCTCGCTCACGCGGACAAGCGCAAAGCGTCCTTTCACGTCATATAGGATGCGGAACGGGGGAAAAAAAAAAAAAAAGAATGCACGCAACCGCTGAGACGTGT
Block A +---356---+ Block B
CGGAGCGAGTGCGCCTGTTCGCGTTTCGCAGGAAAGTGCAGTATATCCTACGCCTTGCCCCTTTTTTTTTTTTTCTTACGTGCGTTGGCGACTCTGCACA
A E S V R V L A F R G K V D Y L I R F G K K K K R M H A T A E T C
GAAACGCCGTCGAGTTCGCGGCGCCCTCCAAATGACAGGCCGGACTTTCGTGAGGGTTCAACTCTTAAGGAGTTTGATAAAAGTGGCCGTCTCGTGGGAT
Block B
CTTTGCGGCAGCTCAAGCGCCGCGGGAGGTTTACTGTCCGGCCTGAAAGCACTCCCAAGTTGAGAATTCCTCAAACTATTTTCACCGGCAGAGCACCCTA
E T P S S S R R P P N D R P D F R E G S T L K E F D K S G R L V G Y
ATTTCCGTTGCGGAAGGATGCTCGGGCGTGGTGGCTTTGCAAAGTGTTATGAAGTGGAGCAGGGTGGCGATACGTATGCGTTGAAGGTTGTGGACCGCTC
503 Block C
TAAAGGCAACGCCTTCCTACGAGCCCGCACCACCGAAACGTTTCACAATACTTCACCTCGTCCCACCGCTATGCATACGCAACTTCCAACACCTGGCGAG
F R C G R M L G R G G F A K C Y E V E Q G G D T Y A L K V V D R S
ACTACTGCAGAAAACTAAGACACTACAGAAACTCCACTCTGAGATATCCATTCACCGCCGCATGAAGCACAAACACATTGTTAACTTCATACGGACCTTC
Block C
TGATGACGTCTTTTGATTCTGTGATGTCTTTGAGGTGAGACTCTATAGGTAAGTGGCGGCGTACTTCGTGTTTGTGTAACAATTGAAGTATGCCTGGAAG
L L Q K T K T L Q K L H S E I S I H R R M K H K H I V N F I R T F
CATGACGACTGGAATGTTTATATTCTGTTGGAGAAGTGCAGCAATCAGACGCTGATGGAAATATTGAAGCGCCGGCAACGTTTCAGTGTGCCAGAGACAC
Block C
GTACTGCTGACCTTACAAATATAAGACAACCTCTTCACGTCGTTAGTCTGCGACTACCTTTATAACTTCGCGGCCGTTGCAAAGTCACACGGTCTCTGTG
H D D W N V Y I L L E K C S N Q T L M E I L K R R Q R F S V P E T Q
AATACATTGCATTGCAGTCGCTTTCCGCCATTCAATATATGCATGAGCAATGTGTTATTCATCGTGACCTGAAACTTGGTAACATAATGATGGATGCAAA
Block C
TTATGTAACGTAACGTCAGCGAAAGGCGGTAAGTTATATACGTACTCGTTACACAATAAGTAGCACTGGACTTTGAACCATTGTATTACTACCTACGTTT
Y I A L Q S L S A I Q Y M H E Q C V I H R D L K L G N I M M D A N
TATGAACGTGAAGATTGGGGACTTCGGTTTGGCTGCTGAGTTGCAGTACGACGGGGAGCGGAAACGCACTATTTGTGGCACGCCCAATTATATCGCACCG
Block C
ATACTTGCACTTCTAACCCCTGAAGCCAAACCGACGACTCAACGTCATGCTGCCCCTCGCCTTTGCGTGATAAACACCGTGCGGGTTAATATAGCGTGGC
M N V K I G D F G L A A E L Q Y D G E R K R T I C G T P N Y I A P
GAGATCATTGAAGGTTCGCGCGAGGGACATAGTTACGAAGTGGATGTGTGGTCGCTCGGTGTGATCCTCTACACGCTGTTGGTTGGAGAGCCGCCGTTTC
Block C
CTCTAGTAACTTCCAAGCGCGCTCCCTGTATCAATGCTTCACCTACACACCAGCGAGCCACACTAGGAGATGTGCGACAACCAACCTCTCGGCGGCAAAG
E I I E G S R E G H S Y E V D V W S L G V I L Y T L L V G E P P F Q
AGACATCAGATGTGAAGGCCACTTACCGTCGTATTAGGCAATGCCGGTATGAATTTCCTTCACATGTAGATGTGCCAGAGAGTGGGAAGGAACTTATTCA
Block C
TCTGTAGTCTACACTTCCGGTGAATGGCAGCATAATCCGTTACGGCCATACTTAAAGGAAGTGTACATCTACACGGTCTCTCACCCTTCCTTGAATAAGT
T S D V K A T Y R R I R Q C R Y E F P S H V D V P E S G K E L I H
CAGTATTCTGCAGAGCCGCCCCGACCAACGGCCGACGTTGCTGGAAATTCGTTCACATCCATTTTTGCGCCTCCCGCCGCCACCGACGACAGCACCTACC
Block C +---1267---+ Block D
GTCATAAGACGTCTCGGCGGGGCTGGTTGCCGGCTGCAACGACCTTTAAGCAAGTGTAGGTAAAAAGGCGGAGGGCGGCGGTGGCTGCTGTCGTGGATGG
S I L Q S R P D Q R P T L L E I R S H P F F R L P P P P T T A P T
ACACTCTTCTACAGTTCACGCCGCCGTCAGCATTCAGATGACCCCCGAGGACATGCGCAGGGACCACTTCCACTTCGCCGCCAAAAATCAGGGGACATTC
Block D
TGTGAGAAGATGTCAAGTGCGGCGGCAGTCGTAAGTCTACTGGGGGCTCCTGTACGCGTCCCTGGTGAAGGTGAAGCGGCGGTTTTTAGTCCCCTGTAAG
T L F Y S S R R R Q H S D D P R G H A Q G P L P L R R Q K S G D I Q
AGGCGGCGTTGCAGAAACAAACACCGCAACGCCGTCAACCGCAGTCGCAGCCCAAATCAGTTGAAGCTGTTCGCTGCATTTCGAGTCCGCGTGTGTCCCG
Block D
TCCGCCGCAACGTCTTTGTTTGTGGCGTTGCGGCAGTTGGCGTCAGCGTCGGGTTTAGTCAACTTCGACAAGCGACGTAAAGCTCAGGCGCACACAGGGC
A A L Q K Q T P Q R R Q P Q S Q P K S V E A V R C I S S P R V S R
CGAAGTGCTTCAACCTATCAGCACAAATCTTCCCAAAACGGACCGATACCACCT
Block D +----1554
GCTTCACGAAGTTGGATAGTCGTGTTTAGAAGGGTTTTGCCTGGCTATGGTGGA
E V L Q P I S T N L P K T D R Y H

**B**




**C**

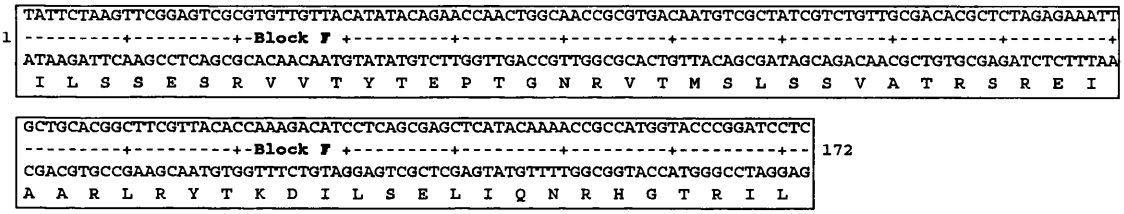



**Fig. 5.9** Nucleotide sequences from pTbplk1 and predicted peptide sequences (sense strands in bold). Panel A: sequence from the 5' end of pTbplk1 (5' Eco RI site to approximately 100 bp 3' of the Hinc II site in subclone D); Panel B: sequence from the 3' end of subclone D; Panel C: sequence from the 3' end of the insert in pTbplk1. The sequences have been arranged so that nucleotide 1 in A, B, C corresponds to the end nearest the 5' end of pTbplk1 (see Fig. 5.8). The boxed blocks refer to the blocks shown in the schematic diagram in Fig. 5.10. The peptide sequence from Block A shows homology to eukaryotic ribosomal protein S4 while that from Block C is homologous to protein kinase catalytic domains and includes the TbPK-A2 sequence (underlined). The boundaries of the protein kinase catalytic domain have been set according to the definition of Hanks and Quinn (1991) and the residues that they used in making this definition are boxed (see text). Note that the alanine residue in the box with the heavier outline is at a position at which most protein kinases have a glycine (see Section 5.2.4.3). The open reading frame of Block C extends through all of Blocks B and D while that of Block A is in the opposite orientation. However, the ends of Blocks A and B at their junction are not the ends of open reading frames (see text). The underlined section of Block E (panel B) is homologous to a highly conserved C-terminal domain of polo-like protein kinases (PH2 domain).

**Fig. 5.10** Schematic diagram showing the predicted polypeptide domains from pTbplk1 compared with polo-like kinases and with *S. cerevisiae* ribosomal protein S4 (RPS4). The shading indicates domains across which there is a high level of homology. The *T. brucei* polypeptide is labelled in the blocks defined in Fig. 5.9. The double slash indicates the likely junction of the ribosomal protein and protein kinase sequences. The polo-like kinases are shown in five domains. The polypeptide from Block B is shown aligned with the N-termini of the polo-like kinases but has yet to be formally shown that it does represent the N-terminus of a *T. brucei* polo-like kinase, and the coding sequence may be incomplete. The polypeptide between the catalytic domain and the highly conserved PH2 domain is labelled as the junction domain and the more 5' of the two unsequenced sections of pTbplk1 is likely to encode part of this domain. Nucleotides 214-301 of Block E, the more 3' of the two unsequenced regions and Block F seem likely to be coding sequence for the C-terminus of the *T. brucei* protein kinase (see Fig. 5.13 and text). While both the junction domain and the C-termini are shown unshaded there is a degree of homology between different polo-like kinases over these domains, the C-terminal domain being more highly conserved than the junction domain. Note also that ribosomal protein S4 homologues from divergent eukaryotes are similar to *S. cerevisiae* ribosomal protein S4 (Fig. 5.14), and that Plk1, STPK13 and rat plk are highly homologous to each other across the entire polypeptides.

**Fig. 5.11** Hybridisation of probes from the protein kinase and RPS4 coding regions of pTbplk1 to a Southern blot of *T. brucei* genomic DNA. *T. brucei* genomic DNA was digested with Hind III (H) or Eco RI (E) and the products run out in an agarose gel which was then blotted. The blot was hybridised first to a probe made from the insert of subclone B (protein kinase coding region, panel A), then stripped and re-hybridised to a probe made from the insert of subclone E (RPS4 coding region, panel B). The positions and sizes of DNA markers are indicated to the left of each autoradiograph, but due to the close spacing of the markers in the relevant region of the gel, the sizes of the hybridising bands can only be estimated very approximately.

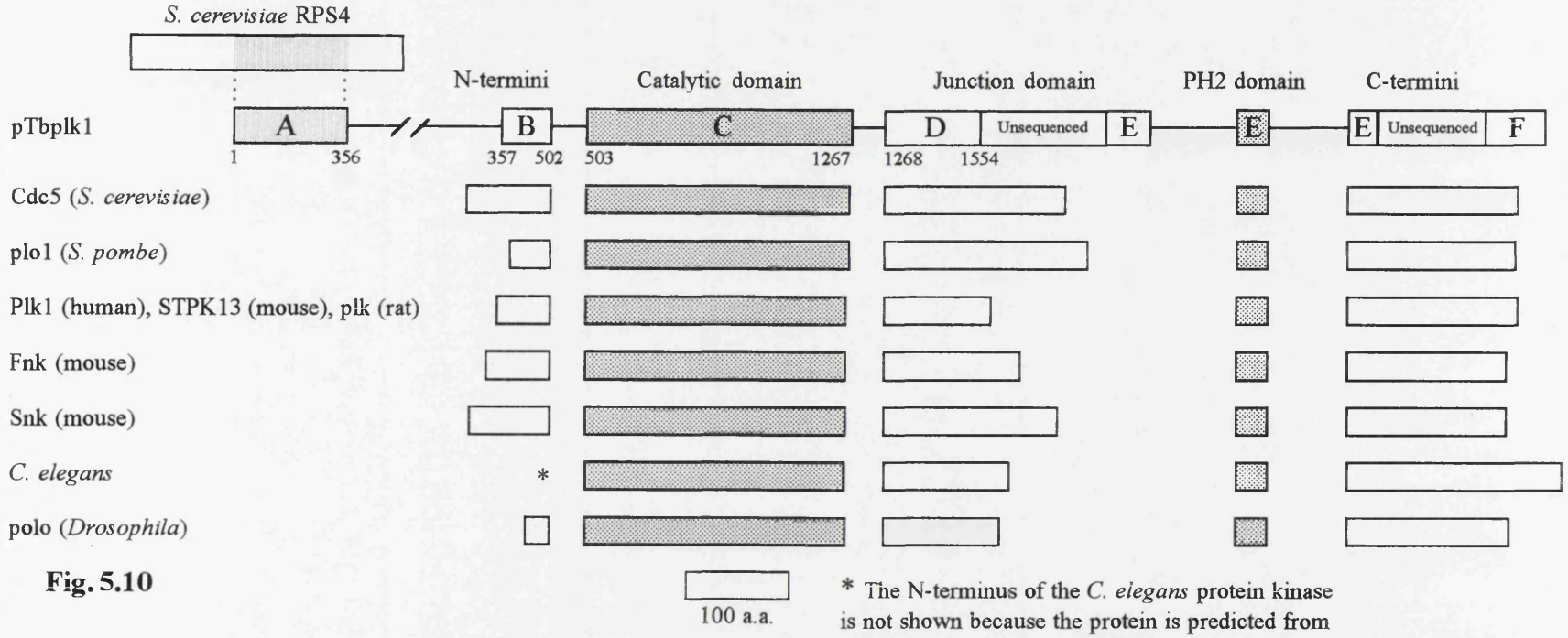



**Fig. 5.10**

* The N-terminus of the *C. elegans* protein kinase is not shown because the protein is predicted from genomic DNA and the 5' boundary of the open reading frame is uncertain

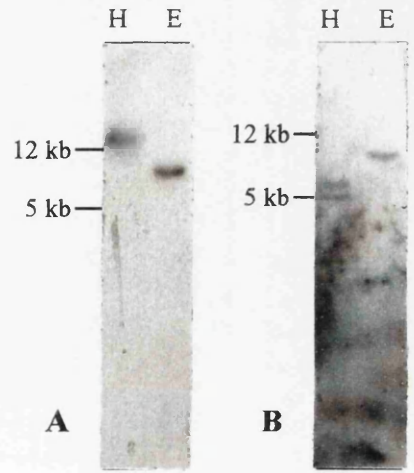



**Fig. 5.11**

**Table 5.1** Similarity of the protein kinase catalytic domain predicted from pTbplk1 to sequences in the protein databases. The method used for obtaining the homology scores is explained in the text. Three human sequences (Lake and Jelinek, 1993; Holtrich *et al.*, 1994; Hamanaka *et al.*, 1994) and two murine sequences (Clay *et al.*, 1993; Hamanaka *et al.*, 1994) are so similar to Plk1 and murine STPK13 respectively that the differences probably represent polymorphisms and these sequences have been excluded from the table. Note also that the % identity score for the *C. elegans* protein kinase is for a sequence slightly different than that in the databases (see Section 5.2.4.3)

| Name | Organism | Catalytic domain identity to TbPLK (%)/ length of overlap | Calculated molecular weight /Da | Reference |
|---|---|---|---|---|
| Cdc5 | *S. cerevisiae* | 53 / 257 | 81,030 | Kitada *et al.*, 1993 |
| Plo1 | *S. pombe* | 53.2/250 | 77,301 | Ohkura *et al.*, 1995 |
| Plk1 | *H. sapiens* | 49 / 257 | 68,254 | Golsteyn *et al.*, 1994 |
| STPK13 | *M. musculus* | 49 / 257 | 68,300 | Lake and Jelinek, 1993 |
| plk | *R. norvegicus* | 47.5 / 257 | 68,313 | Amstrup *et al.* (unpublished) |
| Fnk | *M. musculus* | 46.7/257 | 69,995 | Donohue *et al.*, 1995 |
| Snk | *M. musculus* | 46.5 / 254 | 77,811 | Simmons *et al.*, 1992 |
| K06H7.1 | *C. elegans* | 44 / 257 | * | Wilson *et al.*, 1994 |
| polo | *D. melanogaster* | 43.6 / 257 | 66,947 | Llamazares *et al.*, 1992 |
| Sak | *M. musculus* | 38.9 / 257 | 103,685 (Sak a)<br>52,657 (Sak b) | Fode *et al.*, 1994 |

* this protein kinase is predicted from genomic sequence and the boundaries of the open reading frame are not certain (see Section 5.2.4.3)

**Table 5.1**

**Fig. 5.12** Alignment of catalytic domain peptide sequences of polo-like protein kinases. The catalytic domain sequences of the known polo-like kinases were aligned with the pileup program of the GCG package. The shading indicates residues common to two or more of the sequences, dots indicate gaps introduced by the pileup program to align the sequences. The protein kinase catalytic subdomains are labelled underneath the sequences. Filled boxes indicate positions at which the cAMP-dependent protein kinase (PKA) has glutamate (acidic) residues that are implicated in binding positively charged substrates while open boxes indicate positions where PKA has hydrophobic substrate binding residues (see text). The position marked by a star (between subdomains VII and VIII) is constitutively phosphorylated in PKA and contributes to protein stability while, in many other protein kinases, phosphorylation at the equivalent position has a regulatory role. Note that the *C. elegans* protein kinase is predicted from genomic sequence and that the last ten amino acids in the figure differ from those currently predicted in the databases.

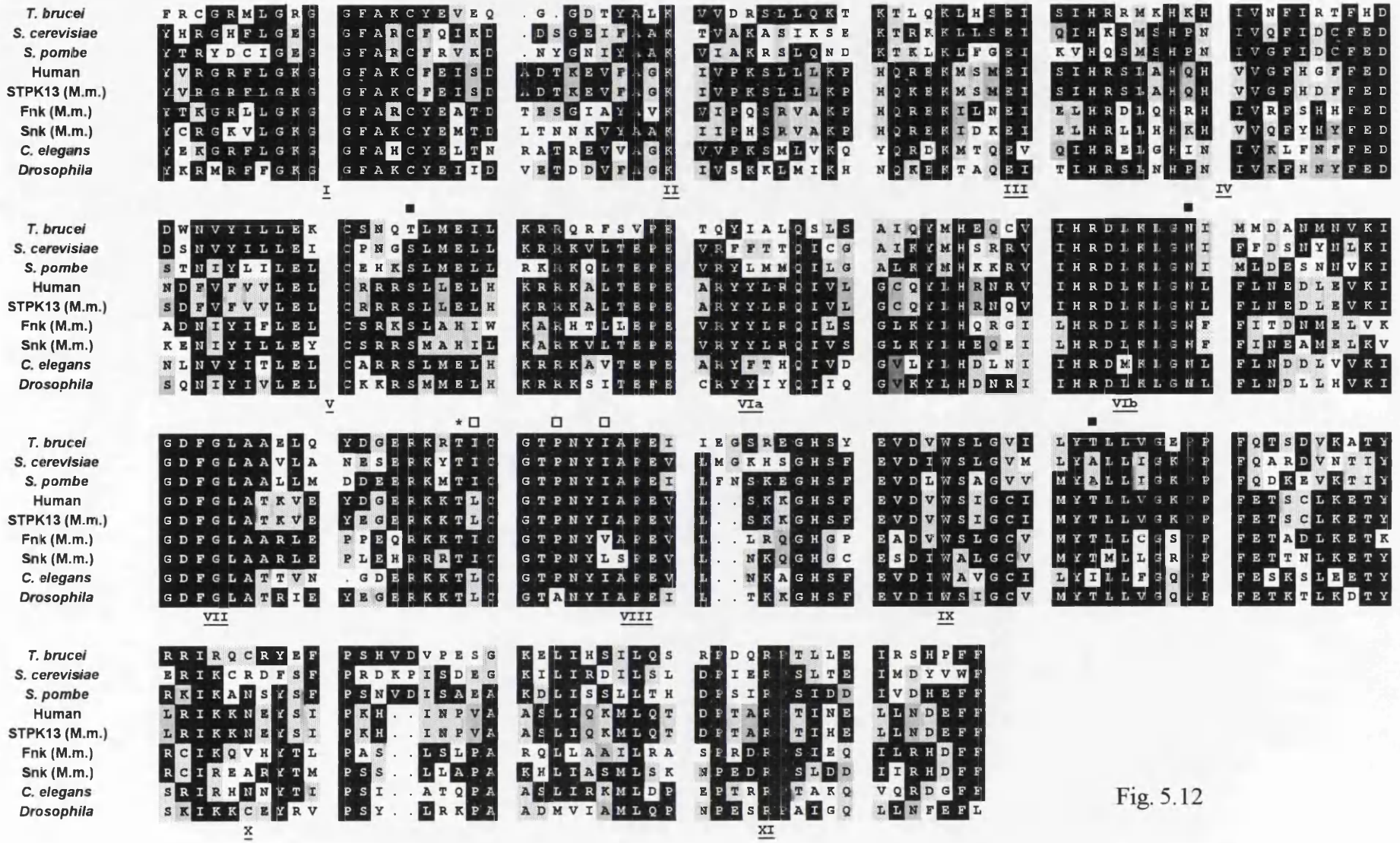

T. brucei
S. cerevisiae
S. pombe
Human
STPK13 (M.m.)
Fnk (M.m.)
Snk (M.m.)
C. elegans
Drosophila
I
II
III
IV
V
VIa
VIb
VII
VIII
IX
X
XI

Fig. 5.12

**Fig. 5.13** Alignment of the C-terminal extensions of polo-like kinases. The pileup program of the GCG package was used to align the *S. cerevisiae* (S.c.), *S. pombe* (S.p.), human (H.s.), Fnk (mouse), Snk (also murine), *Drosophila* (D.m.) and *C. elegans* (C.e.). Details of the three *T. brucei* (T.b.) sequences shown in the figure (from Blocks D, E and F) are given in Section 5.2.4.3. The shading indicates residues common to two or more of the sequences, while dots represent gaps introduced by the pileup program to align the sequences. The highly conserved PH2 domain is underlined.

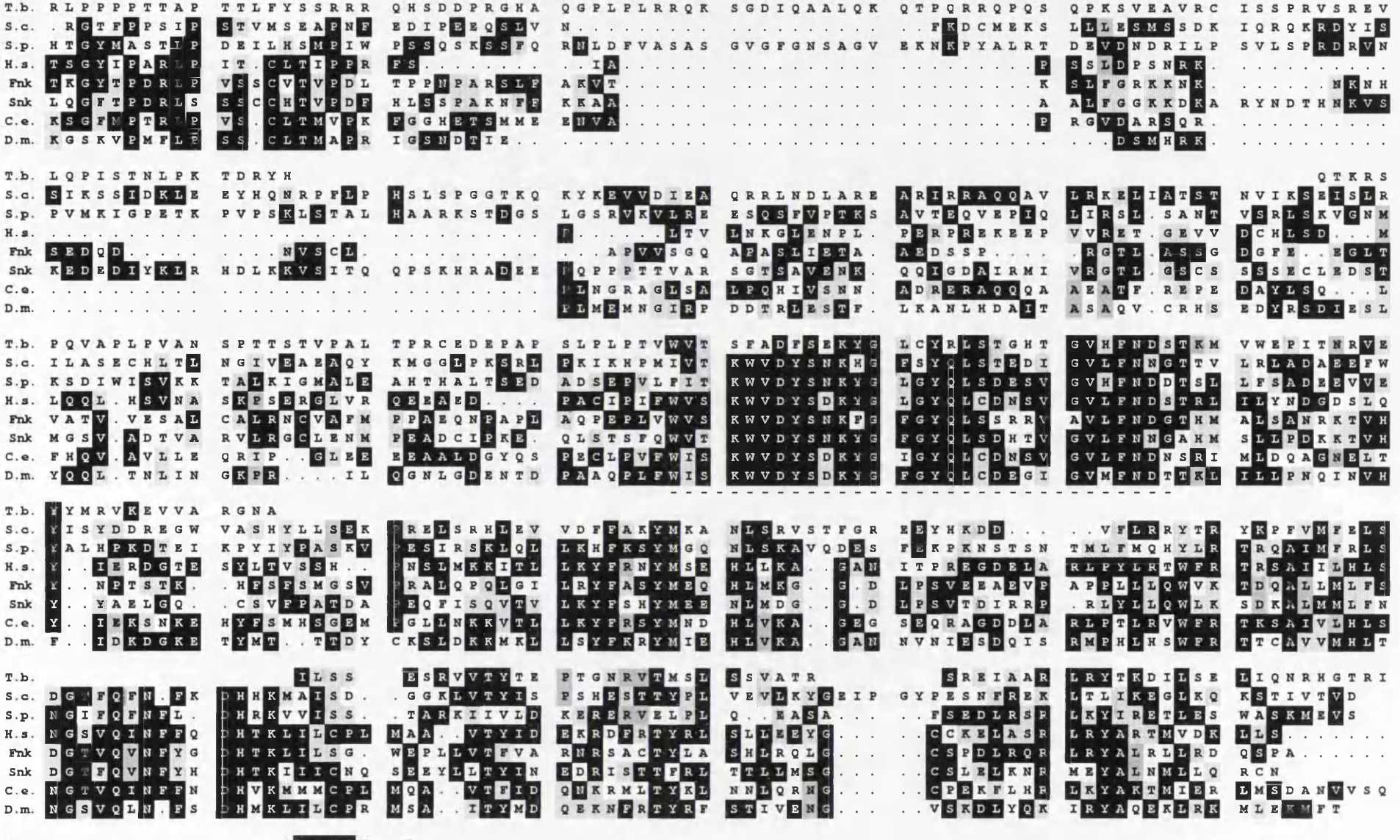



Fig. 5.13

**Fig. 5.14** Alignment of the predicted peptide sequence of the *T. brucei* S4 like protein (TbS4lp) with those of ribosomal protein S4 homologues from *S. cerevisiae*, *Tetrahymena*, human, *Drosophila*, and potato. Residues common to two or more sequences are indicated by shading. The sequences marked as *T. brucei** are from two expressed sequence tags (El-Sayed *et al.*, 1995, accession numbers T26789 and T26790). The peptide sequence in the first two blocks is predicted from T26789 while that in the bottom block is from T26790. The nucleotide sequences from these expressed sequence tags have sequencing errors and so peptide sequences have been added only where there is homology to the other RPS4 sequences. Comparison of the very 5' end of pTbplk1 and the sequence of T26790 shows that the partial RPS4 coding sequence is ligated to the λ vector at a naturally occurring Eco RI site. At the other end of the apparent RPS4 coding sequence in pTbplk1 there is no overlap with T26789.

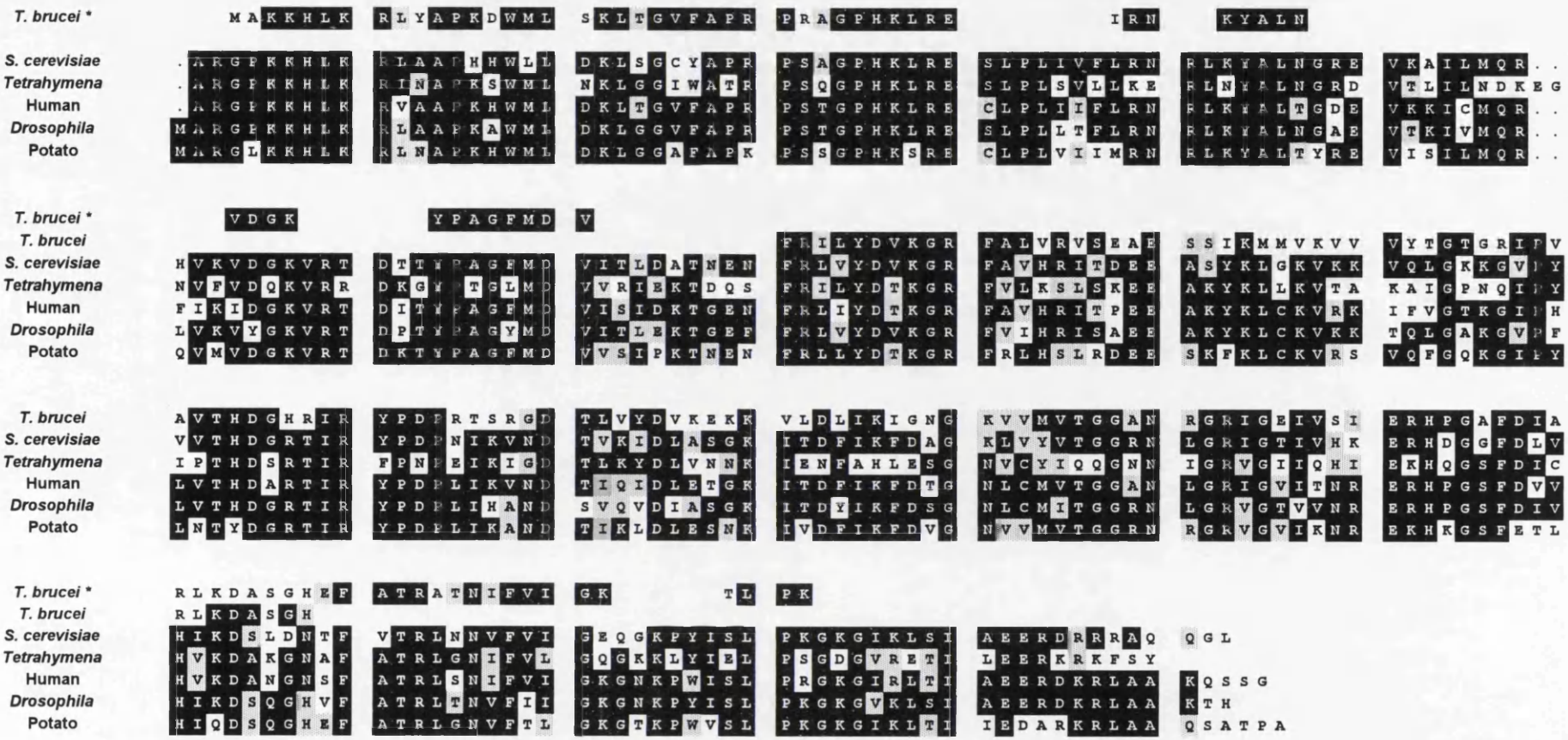

T. brucei *
S. cerevisiae
Tetrahymena
Human
Drosophila
Potato
T. brucei

Fig. 5.14

# Chapter 6

# Future Perspectives

Protein kinases have been studied in *T. brucei* on the basis that they might play similar cellular roles to protein kinases studied in higher eukaryotes. Biochemical approaches to the study of *T. brucei* protein kinases have provided little information about their functional role. However, the future prospects for a functional analysis of *T. brucei* protein kinases in studies based on gene cloning seem much better particularly as transfection technologies for *T. brucei*, suitable for use in examining protein function, have recently been developed. Reverse genetics experiments, such as gene knockouts, can give clues to the cellular function of a particular protein kinase. However, as both protein kinases and protein phosphatases are often organised into networks, the isolated study of an individual protein kinase may be of limited interest and a priority in much protein kinase research is the identification of proteins with which the catalytic subunit of the protein kinase under investigation interacts. These proteins will include substrates and might also include regulatory and targeting subunits and other signalling proteins such as other protein kinases. Unfortunately, *T. brucei* remains a difficult organism to study with respect to the identification of proteins interacting with a known protein because approaches such as co-immunoprecipitation followed by protein microsequencing are hampered by the difficulty of obtaining large quantities of *T. brucei* cells. The identification of proteins interacting with the known proteins based on informed guesses about their identity would be limited by the fact that only a small proportion of *T. brucei* protein coding genes have been sequenced and specific antibodies are only available for a few proteins. A technique that can be used to overcome these problems is known as the yeast two-hybrid system. In this screen for protein-protein interactions the gene for a known protein of interest is fused to that for the DNA-binding domain of a yeast transcriptional activator that binds to a specific DNA sequence but cannot activate transcription except when in the close proximity of an activation domain. The construct is expressed in a yeast strain that has a binding site for the transcriptional activator upstream of a reporter gene. A library of cDNAs or genomic fragments is then created using an expression vector encoding the activation domain and the yeast strain expressing the hybrid "bait" protein is transfected with this library. If any of the hybrid proteins including the transcriptional activator domain bind to the "bait" protein the DNA-binding and transcriptional activator domains of the two hybrids will be brought into close proximity and the reporter gene will be expressed. Cells expressing

the reporter gene are selected and the vector recovered to allow sequencing of the gene for the protein that interacts with the known protein (see review by Evangelista *et al.*, 1996).

On the other hand research into protein kinases in biological model systems including unicellular eukaryotes such as *S. cerevisiae* is progressing rapidly. In *S. cerevisiae*, for example, the investigation of protein function is greatly simplified by fact that classical and reverse genetics are well developed and the entire genome has been sequenced (see reviews by Dujon, 1996 and Johnston, 1996). The future results from research into protein kinases in these organisms would be certain to simplify the study of protein kinases in *T. brucei* by providing a more complete set of characterised protein kinases to which *T. brucei* protein kinases could be compared first for sequence homology and, if chosen for further study, for functional homology. This would be particularly true if the complete sequence of the *T. brucei* genome or a more extensive database of *T. brucei* EST sequences was available. Therefore, while *T. brucei* protein kinases are surely involved in the regulation of processes that are of great interest with respect to its biology (such as differentiation between life cycle stages), there does not at present seem to be justification for the study of large numbers of *T. brucei* protein kinases in comparison to other research that might be undertaken. Research into a particular *T. brucei* protein kinase is most likely to be justified when homology to protein kinases characterised in other organisms suggests that study of such a kinase will complement existing work into an aspect of *T. brucei* biology.

At present the only putative signalling network in *T. brucei* for which a number of potential components have been identified is one utilising cAMP as second messenger. Pays *et al.* (1989) identified an expression site associated gene (named ESAG 4) encoding a putative transmembrane protein predicted to have an extracellular N-terminal domain and a smaller C-terminal domain with adenylate cyclase activity (based on sequence homology to the catalytic domain of *S. cerevisiae* adenylate cyclase). In further studies at least 3 GRESAGs (Genes-Related to ESAGs) related to ESAG 4 were identified (GRESAGs 4.1, 4.2 and 4.3). These genes were shown to be transcribed in both bloodstream and procyclic forms and ESAG 4 and GRESAG 4.1 were able to complement a *S. cerevisiae* adenylate cyclase mutant. Immunolocalisation experiments using antibodies raised against the ESAG 4 protein (which cross-reacted with the

GRESAG 4.1 protein and possibly also the GRESAG 4.2 and 4.3 proteins) showed that in both bloodstream and procyclic forms the proteins are located along the length of the flagellum but could not be detected elsewhere (Alexandre *et al.*, 1990; Paindavoine *et al.*, 1992). Rolin *et al.* (1993) investigated whether adenylate cyclase activity varied during the *in vitro* differentiation of stumpy bloodstream forms to procyclic forms and showed a transient peak of activity at about 4 hours after differentiation was induced. As dividing cells began to appear, adenylate cyclase activity rose gradually to a second peak at about 40 hours (although maximal activity was considerably lower than that for the first peak) and then declined so that by 48 hours activity was similar to that in established procyclic forms. There is also evidence for a possible role for cAMP signalling in other differentiation events in *T. brucei*. Mancini and Patton (1981) analysed parasite intracellular cAMP levels during infections of rats with a pleomorphic *T. brucei* stock and found that levels were low early in infection but rose during the first peak of parasitaemia and were highest just before intermediate and stumpy forms began to appear but then fell dramatically as intermediate and stumpy forms became predominant. The levels of cAMP per parasite cell then rose again as the infection progressed to the second wave of parasitaemia; this increase preceding the appearance of the new population of slender forms. Van den Abbeele *et al.* (1995) found that a crude membrane fraction prepared from tsetse proventriculus/oesophagus stimulated adenylate cyclase activity in procyclic forms while soluble and membrane fractions from other tsetse tissues were inactive. This fraction also stimulated activity in bloodstream forms but while a number of different treatments of bloodstream form cells have this effect (see references in Paindavoine *et al.*, 1992) the proventriculus/oesophagus membrane fraction is the only stimulus known for adenylate cyclase in procyclic forms suggesting possible signalling from the vector to the parasite through one or more of the GRESAG 4 proteins.

The family of adenylate cyclases related to the ESAG 4 protein are the only components of a putative cAMP signal transduction cascade characterised in *T. brucei*, however, three genes encoding putative *T. brucei* PKA homologues and a gene encoding a putative *T. brucei* cAMP phosphodiesterase have recently been isolated (M. Boshart, personal communication) and the proteins seem likely to be involved in cAMP signal transduction. The predicted amino acid sequence of ESAG 8 was found to contain leucine-rich repeats and similar repeats have been found in a regulatory domain of *S.*

*cerevisiae* adenylate cyclase which mediate interaction with Ras homologues. On this basis, it was suggested that the ESAG 4 and ESAG 8 proteins might interact to form an adenylate cyclase complex that might be regulated by *T. brucei* Ras homologues (Ross *et al.*, 1991), however, there is currently no experimental evidence to support this. Given the good evidence for involvement of cAMP signalling in several differentiation events in *T. brucei* and the fact that several genes encoding proteins known or likely to be involved in cAMP signalling have already been cloned, further study of this putative signal transduction cascade would be worthwhile. Among the questions that need to be addressed are the nature of the ligands that bind to the extracellular domain of the adenylate cyclases and the substrates for the *T. brucei* PKA enzymes.

In addition to the putative cAMP signalling pathway a number of other studies have been published dealing with possible pathways through which *T. brucei* might sense and respond to extracellular events. There is also some evidence for signalling from the parasite to the host. Among these Hide *et al.* (1989) showed an effect of EGF on the growth of procyclic *T. brucei* in culture and suggested that *T. brucei* might have an EGF receptor homologue. Sztein and Kierszenbaum (1991) found that a soluble factor from *T. b. rhodesiense* was able to prevent activated human T cells progressing through the cell cycle while Keku *et al.* (1993) found that a soluble factor from *T. brucei* was able to inhibit the growth of HL-60 cells. Kardami *et al.* (1992) identified molecules from *T. brucei* and from *Leishmania* that had properties similar to those of basic FGF. A *T. brucei* factor that has been studied in more detail is trypanosome-derived lymphocyte-triggering factor (TLTF) which is released by *T. b. brucei* and stimulates $CD8^+$ T cells to divide and to secrete IFN-γ, probably through binding directly to CD8. It was also shown that IFN-γ and a neuronally-produced molecule related to IFN-γ stimulate parasite growth suggesting that there may be bi-directional signalling between *T. b. brucei* and its mammalian hosts (Olsson *et al.*, 1991; Olsson *et al.*, 1993; Olsson *et al.*, 1994). It has also been suggested that the production of stumpy forms in the mammalian bloodstream is induced by an exogenous soluble factor(s) that accumulates during infection (Seed and Sechelski, 1989; Hesse *et al.*, 1995, M. Boshart, personal communication) and the results from the latter two studies suggest that this is parasite produced. The identification of *T. brucei* receptors and their ligands would be particularly interesting with respect to understanding the parasite's biology and cell surface receptors would also

present possible targets for immune intervention in African trypanosomiasis. However, in none of these studies has a *T. brucei* receptor been identified by gene cloning.

Putative proteins defined by gene cloning that might be involved in the intracellular aspects of *T. brucei* signal transduction cascades include the two *T. brucei* MAP kinases, Nrk, and TbZFK (see Table 1.2) as well as the *T. brucei* PKA homologues (the TbCRKs and TbPLK seem more likely to be involved in regulation of the cell cycle than in signal transduction) and small GTP-binding proteins (Field *et al.*, 1995; M. Sowa, personal communication). At present no signal transduction cascades linking any of these proteins together have been mapped. There is also very little known about downstream targets of signal transduction cascades in *T. brucei* but cellular processes such as differentiation require major changes in gene expression. In many other eukaryotes, transcription factors are targets of signal transduction cascades but the equivalent targets in *T. brucei* might be proteins involved in controlling RNA maturation and mRNA stability as there is limited transcriptional regulation of gene expression. No genes encoding proteins likely to be involved in these processes have been isolated.

With respect to future work, the identification of *T. brucei* cell surface receptors by biochemical studies seems unlikely to be a profitable approach in view of the fact that cell surface receptors involved in signal transduction tend to be low abundance molecules and *T. brucei* cells are difficult to obtain in large numbers. However, biochemical studies might provide evidence for a particular type of receptor for which the gene could then be isolated using a homology based approach. Such an approach might also be adopted based on information about receptors in other organisms. For instance, it seems possible that *T. brucei* will have members of the superfamily of proteins known as "two-component systems" (see Section 1.2.1). In the bloodstream stages the proteins encoded by ESAGs (and any related GRESAGs) of currently unknown function might be investigated as potential receptors (as is already known to be the case for ESAG 4) since a number of these appear to be cell surface associated.

*T. brucei* may be able to monitor its environment without utilising cell surface receptors dedicated exclusively to a sensory function. For example, a stimulus for the initiation of the stumpy to procyclic transformation (at least *in vitro* but quite possibly also *in vivo*) is a temperature shift which would obviously not require a cell surface receptor (the mechanism through which this signal acts is unknown). In addition to a

temperature shift, citrate and cis-aconitate induce stumpy to procyclic transformation *in vitro*. These are intermediates in the TCA cycle and it is possible that a change in the flux through a particular metabolic pathway caused by differences in the profile of metabolic substrates, available in the tsetse compared to the mammal, signals differentiation *in vivo*. In view of this possibility, some *T. brucei* signal transduction cascades might only be identified through the initial characterisation of intracellular signalling molecules such as protein kinases and members of the Ras superfamily (for which genes can be cloned using homology-based techniques) followed by searches for interacting proteins. The downstream targets of signal transduction cascades in *T. brucei* are likely, in many cases, to be more divergent from downstream targets of signal transduction cascades already characterised in higher eukaryotes compared to proteins playing an intermediate role in signal transduction cascades. The difficulty of characterising such molecules would be another reason to begin mapping a *T. brucei* signal transduction cascade from initial studies of a protein such as a protein kinase. In this respect MAP kinase cascades which are known as central mediators of signalling in both multicellular and unicellular eukaryotes would be an obvious choice for investigation and two genes encoding putative MAP kinase homologues have been isolated from *T. brucei* (see Table 1.2).

The control of the cell cycle is known to involve protein phosphorylation in both multicellular and unicellular eukaryotes. Of the *T. brucei* protein kinases predicted from DNA sequence all of the TbCRKs and TbPLK are probably involved in cell cycle regulation. It has not been determined whether any of the *T. brucei* CRKs acts a functional homologue of cdc2 which is the key regulator of entry into mitosis, however, identification and characterisation of a *T. brucei* functional homologue of cdc2 (assuming there is one) would be of great interest. The existence of a family of CRKs in *T. brucei* (Mottram and Smith, 1995 suggested that there might be at least 6 members of this family in *T. brucei*) might reflect the use of different CRKs to control different cell cycle checkpoints, the use of different CRKs in different life cycle stages or, possibly, the use of CRKs for roles other than controlling cell cycle checkpoints. The characterisation of the *T. brucei* CRKs as putative cell cycle kinases would complement other research into the cell cycle in *T. brucei* and research into differentiation. In order to answer questions about cell cycle regulation in *T. brucei* the identification of at least some of the proteins with which the TbCRKs interact will be required; the putative cyclin partners being the

most obvious of these. The identification of proteins interacting with the TbCRKs would almost certainly be easier if the sequence of the *T. brucei* genome or a more extensive set of EST sequences were available. However, information from future studies into the function of cdc2-related kinases in higher eukaryotes might not be as important with respect to functional characterisation of the TbCRKs as for some other *T. brucei* protein kinases. There is already a large background of information available about cdc2 related kinases from both unicellular and multicellular eukaryotes but there are no unicellular eukaryotes currently being studied as model systems which are known to possess such a large a family of CRKs as *T. brucei*.

Possible future work into the cellular function of TbPLK is discussed in Chapter 5. Briefly, the further study of TbPLK should complement previous studies of the procyclic cell cycle as the kinase is likely to be involved in regulation of the mitotic spindle. This is an interesting area of research since mitosis in *T. brucei* is very different to that in other eukaryotes in which it has been examined. Examination of the phenotype of mutants in *tbplk* may provide information about function. It seems likely that *tbplk* will be an essential gene in *T. brucei* but by expressing *tbplk* under the control of an inducible promoter it should be possible to perform switch-off and overexpression experiments and, assuming TbPLK is involved in regulating the mitotic spindle, the loss of function of *tbplk* should cause defects such as complete or partial failure of nuclear genome segregation (perhaps with the production of zoids) which could be observed by microscopy.

The possible future research on the TbPKAs, TbCRKs, TbPLK and the two *T. brucei* MAP kinase homologues has been considered above but the possible research on Nrk has not. This protein kinase and the genes encoding it have actually been studied in most detail. Nrk is an example of a protein kinase for which peptide sequence homology to known protein kinases gives little indication about likely cellular function. Studies have shown that Nrk protein levels are strongly upregulated in stumpy forms compared to slender bloodstream and procyclic forms, that this occurs early in the differentiation of slender to stumpy forms and that the regulation of expression is at the translational level (this would be worth investigating in its own right). Thus Nrk might have a cellular role specific to stumpy forms or to the slender to stumpy transformation and this would warrant further study. However, as homology gives no indication about what proteins

are likely to interact with Nrk, the identification of these is likely to be difficult unless a technique such as the yeast two-hybrid screen can be applied successfully. The other *T. brucei* protein kinase identified by gene cloning is TbZFK but no information has been published regarding its function. Although it is a member of the PKC family, this is a large protein kinase family and the homology of TbZFK to PKCι, although perhaps suggesting that TbZFK may have particular biochemical properties, gives no real indication of likely cellular function. The remaining *T. brucei* protein kinases identified by gene cloning are those corresponding to the amplified cDNA fragments TbPK-A3-A5 and TbPK-B1-2 (Hua and Wang, 1994; Section 1.3.2.4). The predicted peptide sequences from these fragments, although clearly related to protein kinase catalytic domain sequences, are too short to give a reliable indication of the relatedness of each putative protein kinase to protein kinases from other eukaryotes. For instance, one of the two families of protein kinases to which the predicted peptide sequence of TbPK-A3 is most related, the Snf1 family, includes *S. cerevisiae* SNF1, the mammalian AMP-activated protein kinase (AMPK) and the plant 3-hydroxy-3-methylglutaryl-CoA reductase kinase (HMG-CoA reductase kinase) which share acetyl-CoA carboxylase as substrate. The mammalian and plant enzymes also share HMG-CoA reductase as a substrate (see references in Stapleton *et al.*, 1996). A *T. brucei* HMG-CoA reductase activity has recently been characterised (Coppens *et al.*, 1995; Coppens and Courtoy, 1995) and therefore the study of the protein kinase corresponding to TbPK-A3 might complement this research. However, without a full sequence it is not clear that the protein kinase corresponding to TbPK-A3 is most closely related to the AMPK/HMG-CoA reductase kinase enzymes. The isolation and sequencing of full length genes or cDNAs corresponding to these fragments of sequence has not been undertaken but would certainly be justified in terms of the overall study of protein kinases in *T. brucei*.

It seems unlikely that the purification approach to gene cloning taken with the 60 kDa autophosphorylating protein kinase (Chapter 4) will be used in future studies into protein kinases in *T. brucei*. However, the 60 kDa protein kinase is undoubtedly suitable for the purification approach based on its biochemical properties: the activity is stable to both freezing and high salt concentrations as well as being readily assayed even in quite impure preparations. The property of autophosphorylation can be exploited to specifically label the protein kinase with $^{32}P$ to allow its identification even after

denaturing purification steps. Possible future approaches to completing the purification of the 60 kDa protein kinase are discussed in Chapter 4 but an alternative or complement to any such work would be the further biochemical characterisation of the protein kinase activity. If, on the basis of such characterisation, the 60 kDa protein kinase could be placed into a particular protein kinase family then PCR with degenerate primers based on peptide sequences conserved in this protein kinase family could be used to try to isolate the gene encoding this protein kinase. This analysis would also allow an assessment of whether any protein kinase genes identified in other studies (including genome and EST sequencing projects) were likely to encode the 60 kDa protein kinase. A number of protein kinase inhibitors specific for members of particular protein kinase families were tested against the autophosphorylating activity of the 60 kDa protein kinase but the data obtained did not allow identification of the 60 kDa protein kinase as a member of any particular protein kinase family (data not shown). As the 60 kDa protein kinase activity was not characterised in detail, further analysis of other inhibitors, its substrate specificity and its activity with a wide range of divalent cations might allow its identification as a member of a particular protein kinase family.

In summary, the outlook for research into protein phosphorylation in *T. brucei* has changed a great deal since this project was initiated and the availability of the sequences of a number of genes encoding protein kinase catalytic subunits and newly-developed techniques for transfection of *T. brucei* should be of great value in analysis of the function of these proteins. The detailed characterisation of *T. brucei* protein phosphorylation cascades involved in signal transduction and control of the cell cycle would undoubtedly be of interest with respect to the biology of this parasite. However, as discussed at the beginning of this chapter, protein phosphorylation research in a number of biological model systems is progressing rapidly and information about protein kinases from these systems would be very useful both with respect to deciding whether a particular *T. brucei* protein kinase should be studied further and in actually conducting research into *T. brucei* kinases. The sequence of the *T. brucei* genome or a very complete set of EST sequences may be available in the relatively near future and this would also greatly simplify research into *T. brucei* protein kinases. Therefore, protein kinase research in *T. brucei* should probably be limited at present to protein kinases for which further study seems likely to complement existing work. In addition, the study of

*T. brucei* protein kinases may have relevance to the development of therapies for African trypanosomiasis. The ability of *T. brucei* to differentiate is fundamental to its complex life cycle and the identification of *T. brucei* signalling pathways through research into proteins such as the protein kinases may lead to the identification of key targets for chemotherapeutic intervention.

# References

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*. Biochemical Journal **275**: 7-14.

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

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*. Molecular and Biochemical Parasitology **43**: 279-288.

Allende, J. E. and Allende, C. C. (1995). Protein kinase CK2 - An enzyme with multiple substrates and a puzzling regulation. FASEB Journal 9: 313-323.

Azzi, L., Meijer, L., Ostvold, A. C., Lew, J. and Wang, J. H. (1994). Purification of a 15-kDa cdk4-binding and cdk5-binding protein. Journal of Biological Chemistry **269**: 13279-13288.

Balber, A. E. (1972). *Trypanosoma brucei*: fluxes of the morphological variants in intact and X-irradiated mice. Experimental Parasitology **31**: 307-319.

Ball, K. L., Dale, S., Weekes, J. and Hardie, D. G. (1994). Biochemical characterization of two forms of 3-hydroxy-3-methylglutaryl-CoA reductase kinase from cauliflower (*Brassica oleracia*). European Journal of Biochemistry **219**: 743-750.

Barford, D., Flint, A. J. and Tonks, N. K. (1994). Crystal structure of human protein tyrosine phosphatase 1B. Science **263**: 1397-1404.

Barford, D. and Johnson, L. N. (1989). The allosteric transition of glycogen phosphorylase. Nature **340**: 609-616.

Barry, J. D. and Turner, C. M. R. (1991). The dynamics of antigenic variation and growth of African trypanosomes. Parasitology Today **7**: 207-211.

Barton, N. R. and Goldstein, L. S. B. (1996). Going mobile: microtubule motors and chromosome segregation. Proceedings of the National Academy of Sciences, USA **93**: 1735-1742.

Bass, K. E. and Wang, C. C. (1991). The in vitro differentiation of pleomorphic *Trypanosoma brucei* from bloodstream into procyclic form requires neither intermediary nor short-stumpy stage. Molecular and Biochemical Parasitology **44**: 261-270.

Bastian, L. S., Lodes, M. J., Myler, P. J. and Stuart, K. (1995). Long-range restriction mapping of megabase-sized chromosomes that may be homologs in *Trypanosoma brucei*. Journal of Eukaryotic Microbiology **42**: 278-282.

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*. Molecular and Biochemical Parasitology **63**: 193-202.

BenAmar, M. F., Pays, A., Tebabi, P., Dero, B., Seebeck, T., Steinert, M. and Pays, E. (1988). Structure and transcription of the actin gene of *Trypanosoma brucei*. Molecular and Cellular Biology **8**: 2166-2176.

Berger, B. J., Carter, N. S. and Fairlamb, A. H. (1995). Characterisation of pentamidine-resistant *Trypanosoma brucei brucei*. Molecular and Biochemical Parasitology **69**: 289-298.

Biellmann, J. F., Samama, J. P., Bränden, C. I. and Eklund, H. (1979). X-ray studies of the binding of Cibacron Blue F3GA to liver alcohol dehydrogenase. European Journal of Biochemistry **102**: 107-110

Bienen, E. J., Saric, M., Pollakis, G., Grady, R. W. and Clarkson, A. B., Jr. (1991). Mitochondrial development in *Trypanosoma brucei brucei* transitional bloodstream forms. Molecular and Biochemical Parasitology **45**: 185-192.

Blum, M. L., Down, J. A., Gurnett, A. M., Carrington, M., Turner, M. J. and Wiley, D. C. (1993). A structural motif in the variant surface glycoproteins of *Trypanosoma brucei*. Nature **362**: 603-609.

Borst, P. (1991). Transferrin receptor, antigenic variation and the prospect of a trypanosome vaccine. Trends in Genetics **7**: 307-309.

Borst, P., Fase-Fowler, F., Frasch, A. C. C., Hoeijmakers, J. H. J. and Weijers, P. J. (1980). Characterisation of DNA from *Trypanosoma brucei* and related trypanosomes by restriction endonuclease digestion. Molecular and Biochemical Parasitology **1**: 221-246.

Borst, P., Van der Ploeg, M., Van Hoek, J. F. M., Tas, J. and James, J. (1982). On the DNA content and ploidy of trypanosomes. Molecular and Biochemical Parasitology **6**: 13-23.

Bossemeyer, D. (1994). The glycine-rich sequence of protein kinases: a multifunctional element. Trends in Biochemical Sciences **19**: 201-205.

Bossemeyer, D., Engh, R. A., Kinzel, V., Ponstingl, H. and Huber, R. (1993). Phosphotransferase and substrate binding mechanism of the cAMP-dependent protein kinase catalytic subunit from porcine heart as deduced from the 2.0 Å structure of the complex with $Mn^{2+}$ adenyl imidodiphosphate and inhibitor peptide PKI(5-24). EMBO Journal **12**: 849-859.

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 Research **20**: 5451-5456.

Brun, R. and Schönenberger, M. (1979). Cultivation and *in vitro* cloning of procyclic culture forms of *T. brucei* in a semi-defined medium. Acta Tropica **36**: 289-292.

Brun, R. and Schönenberger, M. (1981). Stimulating effect of citrate and cis-aconitate on the transformation of *Trypanosoma brucei* bloodstream forms to procyclic forms in vitro. Zeitschrift für Parasitenkunde **66**: 17-24.

Burns, J. M., Jr, Parsons, M., Rosman, D. E. and Reed, S. G. (1993a). Molecular cloning and characterization of a 42-kDa protein phosphatase of *Leishmania chagasi*. Journal of Biological Chemistry **268**: 17155-17161.

Burns, J. M., Jr, Shreffler, W. G., Benson, D. R., Ghalib, H. W., Badaro, R. and Reed, S. G. (1993b). Molecular characterization of a kinesin-related antigen of Leishmania chagasi that detects specific antibody in African and American visceral leishmaniasis. Proceedings of the National Academy of Sciences, USA **90**: 775-779.

Cadena, D. L. and Dahmus, M. E. (1987). Messenger RNA synthesis in mammalian cells is catalysed by the phosphorylated form of RNA polymerase II. Journal of Biological Chemistry **262**: 12468-12474.

Celenza, J. L. and Carlson, M. (1986). A yeast gene that is essential for release from glucose repression encodes a protein kinase. Science **233**: 1175-1180.

Chang, C., Kwok, S. F., Bleecker, A. B. and Meyerowitz, E. M. (1993). Arabidopsis ethylene-response gene ETR1: similarity of product to two-component regulators. Science **262**: 539-544.

Chapman, A. B. and Agabian, N. (1994). *Trypanosoma brucei* RNA polymerase II is phosphorylated in the absence of carboxyl-terminal domain heptapeptide repeats. Journal of Biological Chemistry **269**: 4754-4760.

Chen, J., Rauch, C. A., White, J. H., Englund, P. T. and Cozzarelli, N. R. (1995). The topology of the kinetoplast DNA network. Cell **80**: 61-69.

Chijiwa, T., Hagiwara, M. and Hidaka, H. (1989). A newly synthesized selective Casein Kinase I inhibitor, N-(2-Aminoethyl)-5-chloroisoquinoline-8-sulfonamide, and affinity purification of Casein Kinase I from bovine testis. Journal of Biological Chemistry **264**: 4924-4927.

Clay, F. J., McEwen, S. J., Bertoncello, I., Wilks, A. F. and Dunn, A. R. (1993). Identification and cloning of a protein kinase-encoding mouse gene, *Plk*, related to the *polo* gene of *Drosophila*. Proceedings of the National Academy of Sciences, USA **90**: 4882-4886.

Clayton, C. E. and Fox, J. A. (1989). Phosphorylation of fructose bisphosphate aldolase in *Trypanosoma brucei*. Molecular and Biochemical Parasitology **33**: 73-80.

Cohen, P. (1989). The structure and regulation of protein phosphatases. Annual Review of Biochemistry **58**: 453-508.

Cohen, S., Ushiro, H., Stoscheck, C. and Chinkers, M. (1982). A native 170,000 Epidermal Growth Factor Receptor-kinase complex from shed plasma membrane vesicles. Journal of Biological Chemistry **257**: 1523-1531.

Cooper, J. A., Sefton, B. M. and Hunter, T. (1983). Detection and quantification of phosphotyrosine in proteins. Methods in Enzymology **99**: 387-402.

Coppens, I., Baudhuin, P., Opperdoes, F. R. and Courtoy, P. J. (1988). Receptors for the host low-density lipoproteins on the hemoflagellate *Trypanosoma brucei brucei*. Proceedings of the National Academy of Sciences, USA **85**: 6753-6757.

Coppens, I., Bastin, P., Levade, T. and Courtoy, P. -J. (1995). Activity, pharmacological inhibition and biological regulation of 3-hydroxy-3-methylglutaryl coenzyme A reductase in *Trypanosoma brucei*. Molecular and Biochemical Parasitology **69**: 29-40.

Coppens, I. and Courtoy, P. J. (1995). Exogenous and endogenous sources of sterols in the culture-adapted procyclic trypomastigotes of *Trypanosoma brucei*. Molecular and Biochemical Parasitology **73**: 179-188.

Corbin, J. D. and Lincoln, T. M. (1978). Comparison of cAMP- and cGMP-dependent protein kinases. Advances in cyclic nucleotide research **9**: 159-170.

Cunningham, M. L., Zvelebil, M. J. J. M. and Fairlamb, A. H. (1994). Mechanism of inhibition of trypanothione reductase and glutathione reductase by trivalent organic arsenicals. European Journal of Biochemistry **221**: 285-295.

Dale, S., Wilson, W. A., Edelman, A. M. and Hardie, D. G. (1995). Similar substrate recognition motifs for mammalian AMP-activated protein kinase, higher plant HMG-CoA reductase kinase-A, yeast SNF1, and mammalian calmodulin-dependent protein kinase I. FEBS Letters **361**: 191-195.

Davies, S. P., Hawley, S. A., Woods, A., Carling, D., Haystead, T. A. J. and Hardie, D. G. (1994). Purification of the AMP-activated protein kinase on ATP-γ-Sepharose and analysis of its subunit structure. European Journal of Biochemistry **223**: 351-357.

Davis, R. J. (1994). MAPKs: new JNK expands the group. Trends in Biochemical Sciences **19**: 470-473.

De Bondt, H. L., Rosenblatt, J., Jancarik, J., Jones, H. D., Morgan, D. O. and Kim, S.-H. (1993). Crystal structure of cyclin-dependent kinase 2. Nature **363**: 595-602.

Dekker, L. V. and Parker, P. J. (1994). Protein kinase C-a question of specificity. Trends in Biochemical Sciences **19**: 73-77.

DeLange, T., Michels, P. A. M., Veerman, H. J. G., Cornelissen, A. W. C. A. and Borst, P. (1984). Many trypanosome mRNAs share a common 5′ terminal sequence. Nucleic Acids Research **12**: 3777-3790.

Dickin, S. K. and Gibson, W. C. (1989). Hybridisation with a repetitive DNA probe reveals the presence of small chromosomes in *Trypanosoma vivax*. Molecular and Biochemical Parasitology **33**: 135-142.

Dingwall, C. and Laskey, R. A. (1986). Protein import into the cell nucleus. Annual Review of Cell Biology **2**: 367-390.

Donohue, P. J., Alberts, G. F., Guo, Y. and Winkles, J. A. (1995). Identification by targeted differential display of an immediate early gene encoding a putative serine/threonine kinase. Journal of Biological Chemistry **270**: 10351-10357.

Ducommun, B., Brambilla, P., Félix, M. A., Franza, B. R., Karsenti, E. and Draetta, G. (1991). cdc2 phosphorylation is required for its interaction with cyclin. EMBO Journal **10**: 3311-3319.

Dujon, B. (1996). The yeast genome project: what did we learn? Trends in Genetics **12**: 263-270.

Dunphy, W. G. (1994). The decision to enter mitosis. Trends in Cell Biology **4**: 202-207.

El-Sayed, N. M. A., Alarcon, C. M., Beck, J. C., Sheffield, V. C. and Donelson, J. E. (1995). cDNA expressed sequence tags of *Trypanosoma brucei rhodesiense* provide new insights into the biology of the parasite. Molecular and Biochemical Parasitology **73**: 75-90.

Elion, E. A. (1995). Ste5: a meeting place for MAP kinases and their associates. Trends in Cell Biology **5**: 322-327.

Ellis, D. S. (1986). Sex and the single trypanosome. Parasitology Today **2**: 184-185.

Erondu, N. E. and Donelson, J. E. (1991). Characterization of trypanosome protein phosphatase I and 2A catalytic subunits. Molecular and Biochemical Parasitology **49**: 303-314.

Evangelista, C., Lockshon, D. and Fields, S. (1996). The yeast two-hybrid system: prospects for protein linkage maps. Trends in Cell Biology **6**: 196-199.

Evans, D. A. and Brown, R. C. (1972). The utilization of glucose and proline by culture forms of *Trypanosoma brucei*. Journal of Protozoology **19**: 686-690.

Evers, R. and Cornelissen, A. W. C. A. (1990). The *Trypanosoma brucei* protein phosphatase gene: polycistronic transcription with the RNA polymerase II largest subunit gene. Nucleic Acids Research **18**: 5089-5095.

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

Fairlamb, A. H., Henderson, G. B. and Cerami, A. (1989). Trypanothione is the primary target for arsenical drugs against African trypanosomes. Proceedings of the National Academy of Sciences, USA **86**: 2607-2611.

Fazio, M. J., Marrs, J. A. and Bouck, G. B. (1994). Protein kinases in protists. Protoplasma **181**: 283-290.

Fenton, B. and Glover, D. M. (1993). A conserved mitotic kinase active at late anaphase-telophase in syncytial *Drosophila* embryos. Nature **363**: 637-640.

Ferguson, M. A. J. (1994). What can GPI do for you? Parasitology Today **10**: 48-52.

Ferrari, S. and Thomas, G. (1991). Micro- and macropurification methods for protein kinases. Methods in Enzymology **200**: 159-169.

Ferrell, J.E. and Martin, S.G., Jr. (1989). Thrombin stimulates the activities of multiple previously unidentified protein kinases in platelets. Journal of Biological Chemistry **264**: 20723-20729.

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

Fode, C., Motro, B., Yousefi, S., Heffernan, M. and Dennis, J. W. (1994). Sak, a murine protein-serine/threonine kinase that is related to the *Drosophila* polo kinase and involved in cell proliferation. Proceedings of the National Academy of Sciences, USA **91**: 6388-6392.

Fox, L. A., Sawin, K. E. and Sale, W. S. (1994). Kinesin-related proteins in eukaryotic flagella. Journal of Cell Science **107**: 1545-1550.

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

Gale, M., Jr., Carter, V. and Parsons, M. (1994b). Translational control mediates the developmental regulation of the *Trypanosoma brucei* Nrk protein kinase. Journal of Biological Chemistry **269**: 31659-31665.

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

Galyov, E. E., Hakansson, S., Forsberg, A. and Wolf-Watz, H. (1993). A secreted protein kinase of *Yersinia pseudotuberculosis* is an indispensable virulence determinant. Nature **361**: 730-732.

Garrison, J. C. and Wagner, J. D. (1982). Glucagon and the Ca2+-linked hormones Angiotensin II, Norepinephrine, and Vasopressin stimulate the phosphorylation of distinct substrates in intact hepatocytes. Journal of Biological Chemistry **257**: 13135-13143.

Gibson, W. C. (1986). Will the real *Trypanosoma b. gambiense* please stand up. Parasitology Today **2**: 255-257.

Glover, C. V. C. and Allis, C. D. (1991). Enzyme activity dot blots for assaying protein kinases. Methods in Enzymology **200**: 85-90.

Goldberg, J., Huang, H.-b., Kwon, Y.-g., Greengard, P., Nairn, A. C. and Kuriyan, J. (1995). Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. Nature **376**: 745-753.

Goldfarb, D. S., Gariépy, J., Schoolnik, G. and Kornberg, R. D. (1986). Synthetic peptides as nuclear localization signals. Nature **322**: 641-644.

Golsteyn, R. M., Schultz, S. J., Bartek, J., Ziemiecki, A., Ried, T. and Nigg, E. A. (1994). Cell cycle analysis and chromosomal localization of human Plk1, a putative homologue of the mitotic kinases *Drosophila* polo and *Saccharomyces cerevisiae* Cdc5. Journal of Cell Science **107**: 1509-1517.

Golsteyn, R. M., Mundt, K. E., Fry, A. M. and Nigg, E. A. (1995). Cell cycle regulation of the activity and subcellular localization of PLK1, a human protein kinase implicated in mitotic spindle function. Journal of Cell Biology **129**: 1617-1628.

Gottesdiener, K., Garciá-Ãnoveros, J., Lee, M. G.-S. and Van der Ploeg, L. H. T. (1990). Chromosome organisation of the protozoan *Trypanosoma brucei*. Molecular and Cellular Biology **10**: 6079-6083.

Graham, S. V. (1995). Mechanisms of stage-regulated gene expression in Kinetoplastida. Parasitology Today **11**: 217-223.

Graham, S. V. and Barry, J. D. (1991). Expression site-associated genes transcribed independently of variant surface glycoprotein genes in *Trypanosoma brucei*. Molecular and Biochemical Parasitology **47**: 31-42.

Graham, S. V., Matthews, K. R. and Barry, J. D. (1993). *Trypanosoma brucei*: unusual expression-site-associated gene homologies in a metacyclic VSG gene expression site. Experimental Parasitology **76**: 96-99.

Graham, S. V., Matthews, K. R., Shiels, P. G. and Barry, J. D. (1990). Distinct, developmental stage-specific activation mechanisms of trypanosome VSG genes. Parasitology **101**: 361-367.

Griffith, J. P., Kim, J. L., Kim, E. E., Sintchak, M. D., Thomsom, J. A., Fitzgibbon, M. J., Fleming, M. A., Caron, P. R., Hsiao, K. and Navia, M. A. (1995). X-ray structure of calcineurin inhibited by the immunophilin-immunosuppressant FKBP12-FK506 complex. Cell **82**: 507-522.

Grove, D. S. and Mastro, A. M. (1987). Changes in protein kinase C and cAMP-dependent protein kinase in lymphocytes after treatment with 12-O-tetradecanoylphorbol-13-acetate or concanavalin A: quantitation of activities with an in situ gel assay. Journal of Cellular Physiology **132**: 415-427.

Hagan, I. and Yanagida, M. (1990). Novel potential mitotic motor protein encoded by the fission yeast *cut7*$^{+}$ gene. Nature **347**: 563-566.

Hagan, I. and Yanagida, M. (1992). Kinesin-related cut7 protein associates with mitotic and meiotic spindles in fission yeast. Nature **356**: 74-76.

Hajduk, S. L., Hager, K. M. and Esko, J. D. (1994). Human high density lipoprotein killing of African trypanosomes. Annual Review of Microbiology **48**: 139-162.

Hallberg, R. L., Wilson, P. G. (1981). Regulation of ribosome phosphorylation and antibiotic sensitivity in Tetrahymena thermophila: a correlation. Cell 26: 47-56

Hamanaka, R., Maloid, S., Smith, M. R., O'Connell, C. D., Longo, D. L. and Ferris, D. K. (1994). Cloning and characterization of human and murine homologues of the *Drosophila polo* serine-threonine kinase. Cell Growth and Differention **5**: 249-257.

Hamanaka, R., Smith, M. R., O'Connor, P. M., Maloid, S., Mihalic, K., Spivak, J. L., Longo, D. L. and Ferris, D. K. (1995). Polo-like kinase is a cell cycle-regulated kinase activated during mitosis. Journal of Biological Chemistry **270**: 21086-21091.

Hamm, B., Schindler, A., Mecke, D. and Duszenko, M. (1990). Differentiation of *Trypanosoma brucei* bloodstream trypomastigotes from long slender to short stumpy-like forms in axenic culture. Molecular and Biochemical Parasitology **40**: 13-22.

Hanks, S. K. and Hunter, T. (1995a). The eukaryotic protein kinase superfamily. In *The Protein Kinase Factsbook* (ed. D. G. Hardie and S. K. Hanks), 7-47. London, Academic Press.

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

Hanks, S. K. and Lawton, M. A. (1993). Homology-based approaches for identifying cDNAs that encode eukaryotic protein (serine/threonine) and protein (tyrosine) kinases. In *Protein Phosphorylation* (ed. D. G. Hardie), 173-196. Oxford, IRL Press.

Hanks, S. K. and Quinn, A. M. (1991). Protein kinase catalytic domain sequence database: Identification of conserved features of primary structure and classification of family members. Methods in Enzymology **200**: 38-62.

Hanks, S. K., Quinn, A. M. and Hunter, T. (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science **241**: 42-52.

Hannaert, V. and Michels, P.A.M. (1994). Structure, function, and biogenesis of glycosomes in Kinetoplastida. Journal of Bioenergetics and Biomembranes **26**: 205-212.

Hardie, D. G. (1994). An emerging role for protein kinases: the response to nutritional and environmental stress. Cellular Signalling **6**: 813-821.

Hardie, D. G. (1995). Cellular functions of protein kinases. In *The Protein Kinase Factsbook* (ed. D. G. Hardie and S. K. Hanks), 48-56. London, Academic Press.

Harper, J. F., Sussman, M. R., Schaller, G. E., Putnam-Evans, C., Charbonneau, H. and Harmon, A. C. (1991). A calcium-dependent protein kinase with a regulatory domain similar to calmodulin. Science **252**: 951-954.

Hart, D., Langridge, A., Barlow, D. and Sutton, B. (1989). Antiparasitic drug design. Parasitology Today **5**: 117-120.

Harwood, A. J., Plyte, S. E., Woodgett, J., Strutt, H. and Kay, R. R. (1995). Glycogen synthase kinase 3 regulates cell fate in Dictyostelium. Cell **80**: 139-148.

Haystead, C. M. M., Gregory, P., Sturgill, T. W. and Haystead, T. A. J. (1993). γ-phosphate-linked ATP-Sepharose for the affinity purification of protein kinases. European Journal of Biochemistry **214**: 459-467.

Hemmings, B. A., Yellowlees, D., Kernohan, J. C. and Cohen, P. (1981). Purification of glycogen-synthase kinase 3 from rabbit skeletal muscle. European Journal of Biochemistry **119**: 443-451.

Henderson, G. B. and Fairlamb, A. H. (1987). Trypanothione metabolism: a chemotherapeutic target in trypanosomatids. Parasitology Today **3**: 312-315.

Hesse, F., Selzer, P. M., Mühlstädt, K. and Duszenko, M. (1995). A novel cultivation technique for long-term maintenance of bloodstream form trypanosomes in vitro. Molecular and Biochemical Parasitology **70**: 157-166.

Hide, G., Gray, A., Harrison, C. M. and Tait, A. (1989). Identification of an epidermal growth factor homologue in trypanosomes. Molecular and Biochemical Parasitology **36**: 51-60.

Hide, G., Cattand, P., Le Ray, D., Barry, J. D. and Tait, A. (1990). The identification of *Trypanosoma brucei* subspecies using repetitive DNA sequences. Molecular and Biochemical Parasitology **39**: 213-226.

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.

Hindley, J., Phear, G., Stein, M. and Beach, D. (1987). *Suc1*$^{+}$ encodes a predicted 13-kilodalton protein that is essential for cell viability and is directly involved in the division cycle of *Schizosaccharomyces pombe*. Molecular and Cellular Biology **7**: 504-511.

Hoare, C. A. (1972). The trypanosomes of mammals. Oxford, Blackwell.

Hofstee, B.H.J. (1975). Accessible hydrophobic groups of native proteins. Biochemical and Biophysical Research Communications **63**: 618-624.

Holtrich, U., Wolf, G., Bräuninger, A., Karn, T., Böhme, B., Rübsamen-Waigmann, H. and Strebhardt, K. (1994). Induction and down-regulation of PLK, a human serine/threonine kinase expressed in proliferating cells and tumors. Proceedings of the National Academy of Sciences, USA **91**: 1736-1740.

Hu, S.-H., Parker, M. W., Lei, J. Y., Wilce, M. C. J., Benian, G. M. and Kemp, B. E. (1994). Insights into autoregulation from the crystal structure of twitchin kinase. Nature **369**: 581-584.

Hua, S.-b. and Wang, C. C. (1994). Differential accumulation of a protein kinase homolog in *Trypanosoma brucei*. Journal of Cellular Biochemistry **54**: 20-31.

Hubbard, S. R., Wei, L., Ellis, L. and Hendrickson, W. A. (1994). Crystal structure of the tyrosine kinase domain of the human insulin receptor. Nature **372**: 746-754.

Hug, M., Hotz, H. R., Hartmann, C. and Clayton, C. (1994). Hierarchies of RNA-processing signals in a trypanosome surface antigen mRNA precursor. Molecular and Cellular Biology **14**: 7428-7435.

Hunter, T. (1987). A thousand and one protein kinases. Cell **50**: 823-829.

Hunter, T. (1991). Protein kinase classification. Methods in Enzymology **200**: 3-37.

Hunter, T. (1995). Protein kinases and phosphatases: the yin and yang of protein phosphorylation and signaling. Cell **80**: 225-236.

Hurley, J. H., Dean, A. M., Sohl, J. L., Koshland, D. E. and Stroud, R. M. (1990). Regulation of an enzyme by phosphorylation at the active site. Science **249**: 1012-1016.

Inagaki, M., Watanabe, M. and Hidaka, H. (1985). *N*-(2-Aminoethyl)-5-isoquinolinesulfonamide, a newly synthesized protein kinase inhibitor, functions as a ligand in affinity chromatography. Journal of Biological Chemistry **260**: 2922-2925.

Jenö, P., Jäggi, N., Luther, H., Siegmann, M. and Thomas, G. (1989). Purification and characterization of a 40 S ribosomal protein S6 kinase from vanadate-stimulated Swiss 3T3 cells. Journal of Biological Chemistry **264**: 1293-1297.

Jenö, P. and Thomas, G. (1991). Affinity purification of protein kinases using adenosine 5′-triphosphate, amino acid, and peptide analogs. Methods in Enzymology **200**: 178-187.

Jernigan, J. A. and Pearson, R. D. (1993). Chemotherapy of leishmaniasis, Chagas' disease and African trypanosomiasis. Current Opinion in Infectious Diseases **6**: 794-802.

Johnston, M. (1996). The complete code for a eukaryotic cell. Current Biology **6**: 500-503.

Kardami, E., Pearson, T. W., Beecroft, R. P. and Fandrich, R. R. (1992). Identification of basic fibroblast growth factor-like proteins in African trypanosomes and *Leishmania*. Molecular and Biochemical Parasitology 51: 171-182.

Keith, K., Hide, G. and Tait, A. (1990). Characterisation of protein kinase C like activities in *Trypanosoma brucei*. Molecular and Biochemical Parasitology **43**: 107-116.

Keku, T. O., Seed, J. R., Sechelski, J. B. and Balber, A. (1993). *Trypanosoma brucei rhodesiense*: the inhibition of HL-60 cell growth by the African trypanosomes *in vitro*.

Keyse, S. M. (1995). An emerging family of dual specificity MAP kinase phosphatases. Biochimica et Biophysica Acta **1265**: 152-160.

Kitada, K., Johnson, A. L., Johnston, L. H. and Sugino, A. (1993). A multicopy suppressor gene of the *Saccharomyces cerevisiae* $G_1$ cell cycle mutant gene *dbf4* encodes a protein kinase and is identified as *CDC5*. Molecular and Cellular Biology **13**: 4445-4457.

Knighton, D. R., Zheng, J., Ten Eyck, L. F., Ashford, V. A., Xuong, N.-H., Taylor, S. S. and Sowadski, J. M. (1991a). Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science **253**: 407-414.

Knighton, D. R., Zheng, J., Ten Eyck, L. F., Xuong, N. -H., Taylor, S. S. and Sowadski, J. M. (1991b). Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science **253**: 414-420.

Koide, H., Ogita, K., Kikkawa, U. and Nishizuka, Y. (1992). Isolation and characterisation of the epsilon-subspecies of protein kinase C from rat brain. Proceedings of the National Academy of Sciences, USA **89**: 1149-1153.

Kracht, M., Shiroo, M., Marshall, C. J., Hsuan, J. J. and Saklatvala, J. (1994). Interleukin-1 activates a novel protein kinase that phosphorylates the epidermal-growth-factor receptor peptide T669. Biochemical Journal **302**: 897-905.

Krebs, E. G. and Fischer, E. H. (1956). The phosphorylase *b* to *a* converting enzyme of rabbit skeletal muscle. Biochimica et Biophysica Acta **20**: 150-157.

Kuzoe, F. A. S. (1993). Current situation of African trypanosomiasis. Acta Tropica **54**: 153-162.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**: 680-685.

Lake, R. J. and Jelinek, W. R. (1993). Cell cycle- and terminal differentiation-associated regulation of the mouse mRNA encoding a conserved mitotic protein kinase. Molecular and Cellular Biology **13**: 7793-7801.

Lampson, G. P. and Tytell, A. A. (1965). A simple method for estimating isoelectric points. Analytical Biochemistry **11**: 374-377.

Lane, H. A. and Thomas, G. (1991). Purification and properties of mitogen-activated S6 kinase from rat liver and 3T3 cells. Methods in Enzymology **200**: 268-291.

Lanham, S. M. and Godfrey, D. G. (1970). Isolation of salivarian trypanosomes from man and other animals using DEAE-cellulose. Experimental Parasitology **28**: 521-534.

Lee, K. S., Yuan, Y.-L. O., Kuriyama, R. and Erikson, R. L. (1995). Plk is an M-phase-specific protein kinase and interacts with a kinesin-like protein, CHO1/MKLP-1. Molecular and Cellular Biology **15**: 7143-7151

Letwin, K., Mizzen, L., Motro, B., Ben-David, Y., Bernstein, A. and Pawson, T. (1992). A mammalian dual specificity protein kinase, Nek1, is related to the NIMA cell cycle regulator and highly expressed in meiotic germ cells. EMBO Journal **11**: 3521-3531.

Lincoln, T. M. (1983). cGMP-dependent protein kinase. Methods in Enzymology **99**: 63-71.

Le Guen, L., Thomas, M., Bianchi, M., Halford, N. G. and Kreis, M. (1992). Structure and expression of a gene from *Arabidopsis thaliana* encoding a protein related to SNF1 protein kinase. Gene **120**: 249-254.

Llamazares, S., Moreira, A., Tavares, A., Girdham, C., Spruce, B. A., Gonzalez, C., Karess, R. E., Glover, D. M. and Sunkel, C. E. (1991). *polo* encodes a protein kinase homolog required for mitosis in *Drosophila*. Genes and Development **5**: 2153-2165.

Maeda, T., Wurgler-Murphy, S. M. and Salto, H. (1994). A two-component system that regulates an osmosensing MAP kinase cascade in yeast. Nature **369**: 242-245.

Mancini, P. E. and Patton, C. L. (1981). Cyclic 3′,5′-adenosine monophosphate levels during the developmental cycle of *Trypanosoma brucei* in the rat. Molecular and Biochemical Parasitology **3**: 19-31.

Maslov, D. A. and Simpson, L. (1995). Evolution of parasitism in kinetoplastid protozoa. Parasitology Today **11**: 30-32.

Matthews, K. R. and Gull, K. (1994a). Cycles within cycles: the interplay between differentiation and cell division in *Trypanosoma brucei*. Parasitology 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. Journal of Cell Biology **125**: 1147-1156.

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

Maudlin, I. and Welburn, S. C. (1994). Maturation of trypanosome infections in tsetse. Experimental Parasitology **79**: 202-205.

McConville, M. J. and Ferguson, M. A. J. (1993). The structure, biosynthesis and function of glycosylated phosphatidylinositols in the parasitic protozoa and higher eukaryotes. Biochemical Journal **294**: 305-324.

Michels, P. A. M. and Hannaert, V. (1994). The evolution of kinetoplastid glycosomes. Journal of Bioenergetics and Biomembranes **26**: 213-29.

Milhausen, M., Nelson, R. G., Sather, S., Selkirk, M. and Agabian, N. (1984). Identification of a small RNA containing the trypanosome spliced leader: a donor of the shared 5′ sequences of trypanosomatid mRNAs? Cell **38**: 721-729.

Miller, N. and Lehane, M. J. (1993). Peritrophic membranes, cell surface molecules and parasite tropisms within arthropod vectors. Parasitology Today **9**: 45-51.

Moore, J. D. and Endow, S. A. (1996). Kinesin proteins: a phylum of motors for microtubule-based motility. BioEssays 18: 207-219.

Morgan, D. O. and De Bondt, H. L. (1994). Protein kinase regulation: insights from crystal structure analysis. Current Opinion in Cell Biology **6**: 239-246.

Morgan, G. A., Laufman, H. B., Otieno-Omondi, F. P. and Black, S. J. (1993). Control of $G_1$ to S cell cycle progression of *Trypanosoma brucei* S427c11 organisms under axenic conditions. Molecular and Biochemical Parasitology **57**: 241-252.

Mottram, J. C. (1994). cdc2-related protein kinases and cell cycle control in trypanosomatids. Parasitology Today **10**: 253-257.

Mottram, J. C., Kinnaird, J. H., 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. Journal of Biological Chemistry **268**: 21044-21052.

Mottram, J. C. and Smith, G. (1995). A family of trypanosome cdc2-related protein kinases. Gene **162**: 147-152.

Musacchio, A., Gibson, T., Rice, P., Thompson, J. and Saraste, M. (1993). The PH domain: a common piece in the structural patchwork of signalling proteins. Trends in Biochemical Sciences **18**: 343-348.

Muszynska, G., Andersson, L. and Porath, J. (1986). Selective adsorption of phosphoproteins on gel-immobilized ferric chelate. Biochemistry **25**: 6850-6853.

Nairn, A. C., Bhagat, B. and Palfrey, H. C. (1985). Identification of calmodulin-dependent III and its major $M_r$ 100,000 substrate in mammalian tissues. Proceedings of the National Academy of Sciences, USA **82**: 7939-7943.

Neiman, A. M. (1993). Conservation and reiteration of a kinase cascade. Trends in Genetics **9**: 390-394.

Nishizuka, Y. (1984). The role of protein kinase C in cell surface signal transduction and tumour promotion. Nature **308**: 693-698.

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

Nislow, C., Lombillo, V. A., Kuriyama, R., McIntosh, J. R. (1992). A plus-end-directed motor enzyme that moves antiparallel microtubules *in vitro* localizes to the interzone of mitotic spindles. Nature **359**: 543-547.

Nygård, O. and Nika, H. (1982). Identification by RNA-protein cross-linking of ribosomal proteins located at the interface between the small and large subunits of mammalian ribosomes. EMBO Journal **1**: 357-362.

Nygård, O. and Westermann, P. (1982). Specific interaction of one subunit of eukaryotic initiation factor EIF-3 with 18S ribosomal RNA within the binary complex, EIF-3.small ribosomal subunit as shown by cross-linking experiments. Nucleic Acids Research **10**: 1327-1334.

Ogueta, S. B., Solari, A. and Téllez-Iñón, M. T. (1994). *Trypanosoma cruzi* epimastigote forms possess a $Ca^{2+}$-calmodulin dependent protein kinase. FEBS Letters **337**: 293-297.

Ohkura, H., Hagan, I. M. and Glover, D. M. (1995). The conserved Schizosaccharomyces pombe kinase plo1, required to form a bipolar spindle, the actin ring, and septum, can drive septum formation in $G_1$ and $G_2$ cells. Genes and Development **9**: 1059-1073.

Olsen, S. R. and Uhler, M. D. (1989). Affinity purification of the C-α and C-β isoforms of the catalytic subunit of cAMP-dependent protein kinase. Journal of Biological Chemistry **264**: 18662-18666.

Olsson, T., Bakhiet, M., Edlund, C., Höjeberg, B., Van der Meide, P. H. and Kristensson, K. (1991). Bidirectional activating signals between Trypanosoma brucei and CD8+ T cells: a trypanosome-released factor triggers interferon-γ production that stimulates parasite growth. European Journal of Immunology **21**: 2447-2454.

Olsson, T., Bakhiet, M., Höjeborg, 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.

Olsson, T., Kelic, S., Edlund, C., Bakhiet, M., Höjeborg, B., van der Meide, P. H., Ljungdahl, A. and Kristensson, K. (1994). Neuronal interferon-γ immunoreactive molecule: bioactivities and purification. European Journal of Immunology **24**: 308-314.

Ono, Y., Fujii, T., Igarashi, K., Ogita, K. and Nishizuka, Y. (1988a). Nucleotide sequences of cDNAs for α and γ subspecies of rat brain protein kinase C. Nucleic Acids Research **16**: 5199-5200.

Ono, Y., Fujii, T., Ogita, K., Kikkawa, U., Igarashi, K. and Nishizuka, Y. (1988b). The structure, expression, and properties of additional members of the protein kinase C family. Journal of Biological Chemistry **263**: 6927-6932.

Osmani, S. A., Pu, R. T. and Morris, N. R. (1988). Mitotic induction and maintenance by overexpression of a $G_2$-specific gene that encodes a potential protein kinase. Cell **53**: 237-244.

Overath, P., Chaudri, M., Steverding, D. and Ziegelbauer, K. (1994). Invariant surface proteins in bloodstream forms of *Trypanosoma brucei*. Parasitology Today **10**: 53-58.

Overath, P., Czichos, J., Stock, U. and Nonnengaesser, C. (1983). Repression of glycoprotein synthesis and release of surface coat during transformation of *Trypanosoma brucei*. EMBO Journal **2**: 1721-1728.

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

Palm, L., Andersen, J., Rahbek-Nielsen, H., Hansen, T. S., Kristiansen, K. and Højrup, P. (1995). The phosphorylated ribosomal protein S7 in *Tetrahymena* is homologous with mammalian S4 and the phosphorylated residues are located in the C-terminal region. Journal of Biological Chemistry **270**: 6000-6005.

Parsons, M., Carter, V., Muthiani, A. and Murphy, N. (1995). *Trypanosoma congolense*: Developmental regulation of protein kinases and tyrosine phosphorylation during the life cycle. Experimental Parasitology **80**: 507-514.

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*. Molecular and Biochemical Parasitology **63**: 69-78.

Parsons, M., Nelson, R. G., Watkins, K. P. and Agabian, N. (1984). Trypanosome mRNAs share a common 5′ spliced leader sequence. Cell **38**: 309-316.

Parsons, M., Valentine, M. and Carter, V. (1993). Protein kinases in divergent eukaryotes: identification of protein kinase activities regulated during trypanosome development. Proceedings of the National Academy of Sciences, USA **90**: 2656-2660.

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

Pays, E., Tebabi, P., Pays, A., Coquelet, H., Revelard, P., Salmon, D. and Steinert, M. (1989). The genes and transcripts of an antigen gene expression site from *T. brucei*. Cell **57**: 835-845.

Pays, E., Vanhamme, L. and Berberof, M. (1994). Genetic controls for the expression of surface antigens in African trypanosomes. Annual Review of Microbiology **48**: 25-52.

Pearson, R. B. and Kemp, B. E. (1991). Protein kinase phosphorylation site sequences and consensus specificity motifs: tabulations. Methods in Enzymology **200**: 62-81.

Pépin, J. and Milord, F. (1994). The treatment of human African trypanosomiasis. Advances in Parasitology **33**: 1-47.

Priest, J. W. and Hajduk, S. L. (1994). Developmental regulation of mitochondrial biogenesis in *Trypanosoma brucei*. Journal of Bioenergetics and Biomembranes **20**: 179-191.

Pu, R. T. and Osmani, S. A. (1995). Mitotic destruction of the cell cycle regulated NIMA protein kinase of *Aspergillus nidulans* is required for mitotic exit. EMBO Journal **14**: 995-1003.

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. Journal of Cell Biology **128**: 1163-1172.

Roditi, I. and Pearson, T. W. (1990). The procyclin coat of African trypanosomes. Parasitology Today **6**: 79-82.

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*. Journal of Cell Biology **108**: 737-746.

Rolin, S., Paindavoine, P., Hanocq-Quertier, J., Hanocq, F., Claes, Y., Le Ray, D., Overath, P. and Pays, E. (1993). Transient adenylate cyclase activation accompanies differentiation of *Trypanosoma brucei* from bloodstream to procyclic forms. Molecular and Biochemical Parasitology **61**: 115-126.

Rosenblatt, J., De Bondt, H., Jancarik, J., Morgan, D.O. and Kim, S.-H. (1993). Purification and crystallization of human cyclin-dependent kinase 2. Journal of Molecular Biology **230**: 1317-1319.

Ross, D. T., Raibaud, A., Florent, I. C. Sather, S., Gross, M. K., Storm, D. R. and Eisen, H. (1991). The trypanosome VSG expression site encodes adenylate cyclase and a leucine-rich putative regulatory gene. EMBO Journal **10**: 2047-2053.

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

Sanchez, M. A., Zeoli, D., Klamo, E. M., Kavanaugh, M. P. and Landfear, S. M. (1995). A family of putative receptor adenylate cyclases from *Leishmania donovani*. Journal of Biological Chemistry **270**: 17551-17558.

Sawin, K. E., Mitchison, T. J. and Wordeman, L. G. (1992). Evidence for kinesin-like proteins in the mitotic apparatus using peptide antibodies. Journal of Cell Science **102**: 303-313.

Schaeffer, E., Smith, D., Mardon, G., Quinn, W. and Zuker, C. (1989). Isolation and characterisation of two new Drosophila protein kinase C genes, including one specifically expressed in photoreceptor cells. Cell **57**: 403-412.

Schechter, P. J. and Sjoerdsma, A. (1986). Difluoromethylornithine in the treatment of African trypanosomiasis. Parasitology Today **2**: 223-224.

Seebeck, T. and Gehr, P. (1983). Trypanocidal action of neuroleptic phenothiazines in *Trypanosoma brucei*. Molecular and Biochemical Parasitology **9**: 197-208.

Seed, J. R. and Sechelski, J. B. (1989). Mechanism of Long Slender (LS) to Short Stumpy (SS) transformation in the African trypanosomes. Journal of Protozoology **36**: 572-577.

Seger, R., Ahn, N. G., Posada, J., Munar, E. S., Jensen, A. M., Cooper, J. A., Cobb, M. H. and Krebs, E. G. (1992). Purification and characterization of Mitogen-activated Protein Kinase Activator(s) from Epidermal Growth Factor-stimulated A431 cells. Journal of Biological Chemistry **267**: 14373-14381.

Seiwert, S. D. and Stuart, K. (1994). RNA editing: transfer of genetic information from gRNA to precursor mRNA in vitro. Science **266**: 114-117.

Shabb, J. B. and Corbin, J. D. (1992). Cyclic nucleotide-binding domains in proteins having diverse functions. Journal of Biological Chemistry **267**: 5723-5726.

Sherwin, T. and Gull, K. (1989a). The cell division cycle of *Trypanosoma brucei brucei*: timing of event markers and cytoskeletal modulations. Philosophical Transactions of the Royal Society of London B **323**: 573-588.

Sherwin, T. and Gull, K. (1989b). Visualization of detyrosination along single microtubules reveals novel mechanisms of assembly during cytoskeletal duplication in trypanosomes. Cell **57**: 211-221.

Shoji, S., Titani, K., Demaille, J.G. and Fischer, E.H. (1979). Sequence of two phosphorylated sites in the catalytic subunit of bovine cardiac muscle adenosine 3′:5′-monophosphate-dependent protein kinase. Journal of Biological Chemistry **254**: 6211-6214.

Siegfried, E., Chou, T.-B. and Perrimon, N. (1992). *wingless* signaling acts through *zeste-white* 3, the Drosophila homolog of *glycogen synthase kinase-3*, to regulate *engrailed* and establish cell fate. Cell **71**: 1167-1179.

Simmons, D. L., Neel, B. G., Stevens, R., Evett, G. and Erikson, R. L. (1992). Identification of an early-growth-response gene encoding a novel putative protein kinase. Molecular and Cellular Biology **12**: 4164-4169.

Simpson, L. (1987). The mitochondrial genome of kinetoplastid protozoa: genomic organisation, transcription, replication, and evolution. Annual Review of Microbiology **41**: 363-382.

Smith, A. B., Esko, J. D. and Hajduk, S. L. (1995). Killing of trypanosomes by the human haptoglobin-related protein. Science **268**: 284-286.

Smith, J. L., Levin, J. R., Ingles, C. J. and Agabian, N. (1989). In trypanosomes the homolog of the largest subunit of RNA polymerase II is encoded by two genes and has a highly unusual C-terminal domain structure. Cell **56**: 815-827.

Smith, R. F. and King, K. Y. (1995). Identification of a eukaryotic-like protein kinase in Archaebacteria. Protein Science **4**: 126-129.

Sogin, M. L., Gunderson, J. H., Elwood, H. J., Alonso, R. A. and Peattie, D. A. (1989). Phylogenetic meaning of the kingdom concept: an unusual ribosomal RNA from *Giardia lamblia*. Science **243**: 75-77.

Sommer, J. M. and Wang, C. C. (1994). Targeting proteins to the glycosomes of African trypanosomes. Annual Review of Microbiology **48**: 105-138.

Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M. F., Piwnica-Worms, H. and Cantley, L. C. (1994). Use of an oriented peptide library to determine the optimal substrates of protein kinases. Current Biology **4**: 973-982.

Srinivasan, J., Koselak, M., Mendelow, M, Kwon, Y. G. and Lawrence, D. S. (1995). The design of peptide-based substrates for the CDC2 protein kinase. Biochemical Journal **309**: 927-931.

Stapleton, D., Mitchelhill, K. I., Gao, G., Widmer, J., Michell, B. J., Teh, T., House, C. M., Fernandez, C. S., Cox, T., Witters, L. A. and Kemp, B. E. (1996). Mammalian AMP-activated protein kinase subfamily. Journal of Biological Chemistry **271**: 611-614.

Sternberg, J. and Tait, A. (1990). Genetic exchange in African trypanosomes. Trends in Genetics **6**: 317-322.

Stone, R. L. and Dixon, J. E. (1994). Protein tyrosine phosphatases. Journal of Biological Chemistry **269**: 31323-31326.

Sunkel, C. E. and Glover, D. M. (1988). *polo*, a mitotic mutant of *Drosophila* displaying abnormal spindle poles. Journal of Cell Science **89**: 25-38.

Swanson, R. V., Alex, L. A. and Simon, M. I. (1994). Histidine and aspartate phosphorylation: two-component systems and the limits of homology. Trends in Biochemical Sciences **19**: 485-490.

Sztein, M. B. and Kierszenbaum, F. (1991). A soluble factor from *Trypanosoma brucei rhodesiense* that prevents progression of activated human T lymphocytes through the cell cycle. Immunology **73**: 180-185.

Tait, A. (1980). Evidence for diploidy and mating in trypanosomes. Nature **287**: 536-538.

Tait, A. and Turner, C. M. R. (1990). Genetic exchange in *Trypanosoma brucei*. Parasitology Today **6**: 70-75.

Takio, K., Wade, R. D., Smith, S. B., Krebs, E. G., Walsh, K. A. and Titani, K. (1984). Guanosine cyclic 3′,5′-phosphate dependent protein kinase, a chimeric protein homologous with two separate protein families. Biochemistry **23**: 4207-4218.

Tan, J. L. and Spudich, J. A. (1991). Characterization and bacterial expression of the *Dictyostelium* myosin light chain kinase cDNA: Identification of an autoinhibitory domain. Journal of Biological Chemistry **266**: 16044-16049.

Taylor, S. S., Knighton, D. R., Zheng, J., Ten Eyck, L.F. and Sowadski, J. M. (1992). Structural framework for the protein kinase family. Annual Review of Cell Biology **8**: 429-462.

Taylor, S. S. and Radzio-Andzelm, E. (1994). Three protein kinase structures define a common motif. Structure **2**: 345-355.

ter Kuile, B. H. (1993). Glucose and proline transport in Kinetoplastids. Parasitology Today **9**: 206-210.

Tetley, L. and Vickerman, K. (1985). Differentiation in *Trypanosoma brucei*: host-parasite cell junctions and their persistence during acquisition of the variable antigen coat. Journal of Cell Science **74**: 1-19.

Torri, A. F., Bertrand, K. I. and Hajduk, S. L. (1993). Protein stability regulates the expression of cytochrome *c* during the developmental cycle of *Trypanosoma brucei*. Molecular and Biochemical Parasitology **57**: 305-316.

Turrens, J. F. (1991). Mitochondrial metabolism of African trypanosomes. In *Biochemical Protozoology* (ed. G. H. Coombs and M. J. North), 145-153. London, Taylor and Francis Ltd.

Uchiumi, T. and Ogata, K. (1986). Cross-linking study on localization of the binding site for Elongation factor 1α on rat liver ribosomes. Journal of Biological Chemistry **261**: 9668-9671.

Udo, H., Munoz-Dorado, J., Inouye, M. and Inouye, S. (1995). *Myxococcus xanthus*, a Gram-negative bacterium, contains a transmembrane protein serine/threonine kinase that blocks the secretion of β-lactamase by phosphorylation. Genes and Development **9**: 972-983.

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 mortisans mortisans*. Experimental Parasitology **81**: 618-620.

Van der Ploeg, L. H. T., Cornelissen, A. W. C. A., Michels, P. A. M. and Borst, P. (1984c). Chromosome rearrangements in *Trypanosoma brucei*. Cell **39**: 213-221.

Van der Ploeg, L. H. T., Cornelissen, A. W. C. A., Barry, J. D. and Borst, P. (1984a). Chromosomes of Kinetoplastida. EMBO Journal **3**: 3109-3115.

Van der Ploeg, L. H. T., Schwartz, D. C., Cantor, C. R. and Borst, P. (1984b). Antigenic variation in Trypanosoma brucei analysed by electrophoretic separation of chromosome-sized DNA molecules. Cell **37**: 77-84.

Van der Ploeg, L. H. T., Smith, C. L., Polvere, R. I. and Gottesdiener, K. M. (1989). Improved separation of chromosome-sized DNA from *Trypanosoma brucei*. Nucleic Acids Research **17**: 3217-3227.

Vanhamme, L. and Pays, E. (1995). Control of gene expression in trypanosomes. Microbiological Reviews **59**: 223-240.

Vickerman, K. (1985). Developmental cycles and biology of pathogenic trypanosomes. British Medical Bulletin **41**: 105-114.

Vickerman, K. (1994). The evolutionary expansion of the trypanosomatid flagellates. International Journal for Parasitology **24**: 1317-1331.

Walczak, C. E., Mitchison, T. J. and Desai, A. (1996). XKCM1: A Xenopus kinesin-related protein that regulates microtubule dynamics during mitotic spindle assembly. Cell **84**: 37-47.

Walter, R. D. (1976). Nucleoside-dependent protein kinase from *Trypanosoma gambiense*. Biochimica et Biophysica Acta **429**: 137-146.

Walter, R. D. (1978). Multiple protein kinases from *Trypanosoma gambiense*. Hoppe-Seylers Zeitschrift fur Physiologische Chemie **359**: 601-606.

Walter, R. D. (1980). Effect of suramin on phosphorylation-dephosphorylation reactions in *Trypanosoma gambiense*. Molecular and Biochemical Parasitology **1**: 139-142.

Walter, R. D. and Opperdoes, F. R. (1982). Subcellular distribution of adenylate cyclase, cyclic-AMP phosphodiesterase, protein kinases and phosphoprotein phosphatase in *Trypanosoma brucei*. Molecular and Biochemical Parasitology **6**: 287-295.

Wang, C. C. (1991). A novel suicide inhibitor strategy for antiparasitic drug development. Journal of Cellular Biochemistry **45**: 49-53.

Wang, Q. M., Fiol, C. J., Depaoli-Roach, A. A. and Roach, P. J. (1994). Glycogen-synthase kinase-3-β is a dual-specificity kinase differentially regulated by tyrosine and serine/threonine phosphorylation. Journal of Biological Chemistry **269**: 14566-14574.

Wang, Y. and Roach, P. J. (1993). Purification and assay of mammalian protein (serine/threonine) kinases. In *Protein Phosphorylation* (ed. D. G. Hardie), 121-142. Oxford, IRL Press.

Webster, P. and Russell, D. G. (1993). The flagellar pocket of trypanosomatids. Parasitology Today **9**: 201-206.

Wedegaertner, P. B. and Gill, G. N. (1989). Activation of the purified tyrosine kinase domain of the Epidermal Growth-Factor Receptor. Journal of Biological Chemistry **264**: 11346-11353.

Wei, L., Hubbard, S. R., Smith, R. F. and Ellis, L. (1994). Protein kinase superfamily-comparisons of sequence data with three-dimensional structures. Current Opinion in Structural Biology **4**: 450-455.

Wera, S. and Hemmings, B. A. (1995). Serine/threonine protein phosphatases. Biochemical Journal **311**: 17-29.

Wheeler-Alm, E. and Shapiro, S. Z. (1992). Evidence of tyrosine kinase activity in the protozoan parasite *Trypanosoma brucei*. Journal of Protozoology **39**: 413-416.

Wirtz, E. and Clayton, C. (1995). Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. Science **268**: 1179-1183.

Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M., Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J., Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A., Fraser, A., Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M., Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laisster, N., Latrille, P., Lightning, J., Lloyd, C., Mortimore, B., O'Callaghan, M., Parsons, J., Percy, C., Rifken, L., Roopra, A., Saunders, D., Shownkeen, R., Sims, M., Smaldon, N., Smith, A., Smith, M., Sonnhammer, E., Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M., Vaughan, K., Waterston, R., Watson, A., Weinstock, L., Wilkinson-Sproat, J. and

Wohldman, P. (1994). 2.2 Mb of contiguous nucleotide sequence from chromosome III of *C. elegans*. Nature **368**: 32-38.

Wise, B. C., Raynor, R. L. and Kuo, J. F. (1982). Phospholipid-sensitive $Ca^{2+}$-dependent protein kinase from heart. Journal of Biological Chemistry **257**: 8481-8488.

Woodgett, J. R. (1991). Use of synthetic peptides mimicking phosphorylation sites for affinity purification of protein-serine kinases. Methods in Enzymology **200**: 169-178.

Woods, A., Baines, A. J. and Gull, K. (1992). A high molecular mass phosphoprotein defined by a novel monoclonal antibody is closely associated with the intermicrotubule cross bridges in the *Trypanosoma brucei* cytoskeleton. Journal of Cell Science **103**: 665-675.

Woodward, R. and Gull, K. (1990). Timing of nuclear and kinetoplast DNA replication and early morphological events in the cell cycle of *Trypanosoma brucei*. Journal of Cell Science **95**: 49-57.

Wooten, M. W., Vandenplas, M. and Nel, A. E. (1987). Rapid purification of protein kinase C from rat-brain. European Journal of Biochemistry **164**: 461-467.

Xu, R. -M., Carmel, G., Sweet, R. M., Kuret, J. and Cheng, X. (1995). Crystal structure of casein kinase-1, a phosphate-directed protein kinase. EMBO Journal **14**: 1015-1023.

Ye, X. S., Xu, G., Pu, R. T., Fincher, R. R., McGuire, S. L., Osmani, A. H. and Osmani, S. A. (1995). The NIMA protein kinase is hyperphosphorylated and activated downstream of $p34^{cdc2}$/cyclin B: coordination of two mitosis promoting kinases. EMBO Journal **14**: 986-994.

Zhang, F., Strand, A., Robbins, D., Cobb, M. H. and Goldsmith, E. J. (1994). Atomic structure of the MAP kinase ERK2 at 2.3 Å resolution. Nature **367**: 704-711.

Zhang, W. D., Munoz-Dorado, J., Inouye, M. and Inouye, S. (1992). Identification of a putative eukaryotic-like protein kinase family in the developmental bacterium *Myxococcus xanthus*. Journal of Bacteriology **174**: 5450-5453.

Zheng, J., Knighton, D. R., Ten Eyck, L. F., Karlsson, R., Xuong, N.-H., Taylor, S. S. and Sowadski, J. M. (1993). Crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MgATP and peptide inhibitor. Biochemistry **32**: 2154-2161.

Zheng, P., Fay, D. S., Burton, J., Xiao, H., Pinkham, J. L. and Stern, D. F. (1993). *SPK1* is an essential S-phase specific gene of *Saccharomyces cerevisiae* that encodes a nuclear serine/threonine/tyrosine kinase. Molecular and Cellular Biology **13**: 5829-5842.

Ziegelbauer, K., Quinten, M., Schwarz, H., Pearson, T. W. and Overath, P. (1990). Synchronous differentiation of *Trypanosoma brucei* from bloodstream to procyclic forms *in vitro*. European Journal of Biochemistry **192**: 373-378.

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.

Zoller, M. J., Nelson, N. C. and Taylor, S. S. (1981). Affinity labeling of cAMP-dependent protein kinase with *p*-Fluorosulfonylbenzoyl Adenosine. Journal of Biological Chemistry **256**: 10837-10842.