Identifying substrates of CDK2:cyclin A.

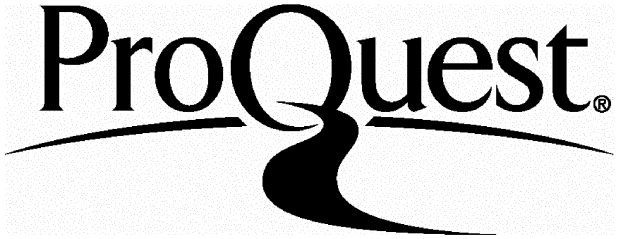
A dissertation submitted to the University of London
for the degree of Doctor of Philosophy

by

Tod Duncan

Imperial Research Cancer Fund,
Clare Hall.

September, 2001.



Abstract.

CDKs and cyclins are the central cell cycle regulators in all organisms studied to date. Full kinase activity of a CDK is only realised upon the binding of a cognate cyclin and phosphorylation on an activating threonine residue. The major cell cycle events, such as entry into S-phase or chromatin condensation, are known to depend on the kinase activity of CDK:cyclin complexes, although the exact changes which these complexes cause to promote particular cell cycle transitions remains unclear. In human cells entry into S-phase is under the control of CDK2 complexed with cyclin E, although soon after entry into S-phase the predominant complex present is CDK2:cyclin A.

Two screening methods were developed based on the *in vitro* expression cloning technique. These screens were employed to identify proteins which could either bind to, or that were substrates of, CDK2:cyclin A. A *Xenopus laevis* and a HeLa cDNA library were constructed, the cDNA inserts of which could be expressed using *in vitro* coupled transcription and translation. Pools of plasmid DNA were prepared from the HeLa cDNA library which were used to generate pools of labelled proteins. These pools were then searched for proteins which could bind to micro-affinity columns of CDK2:cyclin A, or which displayed altered electrophoretic mobility through SDS-PAGE in the presence of active CDK2:cyclin A.

From these screens, 15 *in vitro* substrates of CDK2:cyclin A were identified. The localisation of some of these proteins is described. Using these 15 proteins, the substrate specificity of CDK2:cyclin A and CDK1:cyclin B has been examined. The necessity of an intact hydrophobic patch for efficient phosphorylation of substrates has also been investigated and it appears that there is a difference in the way that CDK2:cyclin E and CDK2:cyclin A recognise their substrates.

Table of Contents

List of figures

Abbreviations

ATP	Adenosine 5'-triphosphate
BAH	Bromo-adjacent homology domain
BSA	Bovine serum albumin
CAK	CDK activating kinase
cAMP	Cyclin adenosine monophosphate
CDC	Cell division cycle
CDK	Cyclin Dependent Kinase
cDNA	complimentary DNA
CFP	Cyan fluorescent protein
Ci	Curie
CKI	Cyclin-dependent kinase inhibitor
CnBr	Cyanogen bromide
CsCl	Caesium chloride
CSF	Cytostatic factor
CTD	Carboxy terminal domain (of RNA polymerase II)
dCTP	Deoxy-cytidine 5'-triphosphate
DIMPIM	Dimethylpimelidate
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
dTTP	Deoxy-thymidine triphosphated
ECL	Enhanced chemiluminescence
EDTA	Ethylene diamine tetraacetic acid
EGTA	Ethylene glyco-bis(b-aminoethyl ether)N,N,N',N' tetraacetic acid
ERK1	Extracellular regulated kinase 1
EtBr	Ethidium Bromide
EtOH	Ethanol
FACS	Fluorescence-activated cell sorter
FITC	Fluorescein isothiocyanate
FPLC	Fast protein liquid chromatography
GEF	GDP-GTP exchange factor
GFP	Green fluorescent protein
GSH	Glutathione
GST	Glutathione-S-transferase
GTP	Guanosine 5'-triphosphate
Gy	Gray
HBS	HEPES buffered saline
HEPES	N-[2-Hydroxyethyl]piperazine-N'-[2-ethanesulfonic acid]
HRP	Horseradish peroxidase
ICRF	Imperial Cancer Research Fund
IPTG	Isopropyl-β-D-thiogalactopyranoside
IVEC	*In vitro* expression cloning
kDa	KiloDaltons
kbp	Kilobase pairs
LiCl	Lithium chloride
MCS	Multiple cloning site

MCM	Minichromosome maintenance
MDa	MegaDaltons
ml	Millilitre
MeOH	Methanol
μM	Micromolar
MPF	Maturation promoting factor
M_r	Molecular mass
mRNA	Messenger RNA
MTA1	Metastasis-associated protein-1
NaOAc	Sodium acetate
NaOH	Sodium hydroxide
NCBI	National Center for Biotechnology Information
NEM	N-ethyl maleimide
NER	Nucleotide excision repair
NLS	Nuclear localisation signal
NTA	Nitrilotriacetic acid
NuRD	Nucleosome remodelling and deacetylase complex
OAc	Acetate
SDS-PAGE	Sodium-Dodecyl Sulphate-Polyacrylamide Gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
Pfu	*Pyrococcus furiosus*
PH	Pleckstrin homology
pI	Isoelectric point
pM	Picomolar
PMSF	Phenylmethylsulphonylfluoride
RFP	Red fluorescent protein
RNA	Ribonucleic acid
RNase	Ribonuclease
RNasin	Ribonuclease inhibitor
rpm	Revolutions per minute
rRNA	Ribosomal RNA
RT-PCR	Reverse-transcription polymerase chain reaction
SANT	Swi3/Ade2/N-CoR/TFIIIB domain
SDS	Sodium dodecyl sulphate
SMART	Simple Modular Architecture Research Tool
SV40	Simian virus 40
Taq	*Thermophilus aquaticus*
TCA	Trichloroacetic acid
TFIIH	Basal transcription factor IIH
TnT™	Coupled transcription-translation system
Tris	Tris[hydroxymethyl]aminomethane
TRITC	Tetramethylrhodamine B isothiocyanate
tRNA	Transfer RNA
Tween-20	Polyoxyethylene-sorbitan monolaurate
UTR	Untranslated region
UV	Ultra-violet

YFP	Yellow fluorescent protein

the cell cycle consisted of a series of interdependent events, where initiation of the next step required the completion of the previous one. For example, yeast mutants with unreplicated DNA failed to enter mitosis. The mechanisms which were believed to be activated in response to failed cell cycle events, and which probably existed to prevent the inappropriate progress of the cell cycle were called checkpoints (Hartwell and Weinert, 1989).

Schizosaccharomyces pombe is an ideal organism to study the interdependence of the cell cycle and growth or cell size. An interesting fission yeast cdc mutant was cdc2, which was required for entry into mitosis. Interestingly, other alleles of cdc2 caused cells to divide at smaller sizes than was normal, which indicated that cdc2 was a cell cycle regulator whose actions could be stimulated as well as perturbed, causing the cell division cycle to be performed before cells had reached the normal size for division. The cloning and sequencing of cdc2 showed that it bore homology to protein kinases (Hindley and Phear, 1984). It was soon discovered that cdc2 and *CDC28* encoded conserved, interchangeable kinases and that a human ortholog existed (Lee and Nurse, 1987). It was already known that there were large changes in protein phosphorylation as cells entered mitosis from work performed on MPF in frogs and starfish, and it was not long before the concerted efforts of yeast geneticists and cell biologists led to the hypothesis that transitions between the cell cycle phases were caused by protein phosphorylation.

Observation of protein synthesis in fertilised sea urchin eggs led to the discovery of proteins which were destroyed at the end of each mitosis, and which accumulated during the next interphase (Evans *et al.*, 1983). These proteins, called cyclins, were believed to cause oscillation in MPF activity which was implicated in driving the alternation of S- and M-phases. MPF was partially purified in 1986 (Nguyen *et al.*, 1986), but purification in 1988 (Gautier *et al.*, 1988b) showed that it contained two subunits. Antibodies against a conserved motif in cdc2 revealed that one of the subunits was cdc2 (Gautier *et al.*, 1988a), and antibodies against frog cyclins revealed that MPF also contained cyclin B.

1.2. The major cell cycle functions of *S. cerevisiae* are controlled by Cdc28 and a number of different cyclins.

Budding yeast possesses five CDKs (Cdc28, Pho85, Kin28, Ssn3 and Ctk1). By far the most well studied is Cdc28 and this kinase controls the major cell cycle events in budding

yeast and is an essential gene. The abundance of the Cdc28 protein is constant throughout the cell cycle and little has been reported on environmental or cell cycle effects on *CDC28* transcription (Mendenhall and Hodge, 1998). The protein is both stable and in excess in the cell and it appears that Cdc28 activity is controlled primarily post-translationally.

The cyclin activators of Cdc28 can be broadly classified into two groups: the Clns (Cln1-3), which act in G1 and are required for Start, and the B-type cyclins (Clb1-6) which control S-phase and mitotic events. The major event between mitosis and S-phase is Start where cells commit to a round of DNA replication and cell division. Start is under the control of Cdc28 complexed with Clns, while events from Start to mitosis are regulated by Cdc28 complexed with Clbs. Each of the Clns and Clbs confer limited functions on Cdc28 but redundancy and overlap in these functions makes a detailed analysis of the specific function of the individual cyclins difficult. The expression level of *CLN1* and *CLN2*, their protein products and their Cdc28-associated kinase activity peaks around Start (Wittenberg *et al.*, 1990) indicating that they play a role in committing the cell to proceeding through the cell cycle. That Cln2 controls cell cycle events at Start and the commitment to cell division is supported by the observation that stable alleles of *CLN2* result in the greatly accelerated passage through Start (Hadwiger *et al.*, 1989). Deletions of either *CLN1* or *CLN2* on the other hand have no significant phenotypes, suggesting that their functions overlap somewhat. Cells with double *CLN1/CLN2* knockouts however show defects in the timing of bud emergence, grow slowly, are aberrantly shaped and are delayed in the initiation of DNA synthesis (Stuart and Wittenberg, 1994). In addition to their functions at or close to Start, Cln1 and Cln2 indirectly induce the initiation of DNA replication by promoting the proteolysis of the Cdc28:Clb inhibitor, Sic1. The bud emergence defect observed in double *CLN1/CLN2* knockouts would also suggest that Cln1 and/or Cln2 play a role in controlling the timing of bud emergence, and it is also likely that they control the events leading to the duplication of the microtubule-organising center, or spindle pole body.

Cln3 has only approximately 20% identity to Cln1 and Cln2 and has greater overall sequence similarity to the B-type cyclins Clb5 and Clb6. The transcription of *CLN3* does not oscillate through the cell cycle like the other cyclins and Cln3 protein levels display only moderate oscillations (Cross and Blake, 1993). Cells lacking *cln3* display more dramatic cell cycle perturbations than cells with *CLN1* or *CLN2* deletions. Cells lacking

cln3 are significantly enlarged and have an extended G1 although the average cell cycle length is similar to that in wild-type cells due to differences in the lengths of other cell cycle stages (Jeoung *et al.*, 1998). Although the effects of deleting *cln3* appear to be more severe than deleting either *cln1* or *cln2*, the protein levels of Cln3 are considerably lower than those of Cln1 and Cln2, and moreover the Cdc28-associated histone H1 kinase activity of Cln3 is lower than that for Cdc28:Cln1/2 complexes (Tyers *et al.*, 1993). It is possible that Cln3 acts as an upstream activator of *CLN1* and *CLN2* transcription in G1 (Amon *et al.*, 1993).

Budding yeast contain six B-type cyclins (Clbs1-6), which are best classified in terms of their expression patterns throughout the cell cycle. *CLBs5* and *6* are expressed in G1 at the same time as *CLNs1* and *2* and cells lacking *CLN1*, *CLN2*, *CLB5* and *CLB6* are inviable (Schwob and Nasmyth, 1993). *CLB5* deleted cells have an extended S-phase, while *CLB6* deleted cells are smaller due to a shortened G1 phase and an earlier transition through Start. Deletion of both *CLB5* and *CLB6* causes a delay in the initiation of S-phase, but once initiated, replication proceeds with normal kinetics. These data support the hypothesis that Clbs5 and 6 are involved in regulating the initiation of S-phase (Schwob *et al.*, 1994), and it has also been shown that they prevent re-replication (Dahmann *et al.*, 1995) and regulate the activity of Cdc28:Cln (Basco *et al.*, 1995). Other genetic analyses using combinations of *CLB* and *CLN* knockouts suggested that Clbs5 and 6 are also involved in the formation of spindles, but that they are not sufficient for the formation of bipolar mitotic spindles. It is unlikely that Clb5 and Clb6 carry out Start-related functions, though, as during G1 phase, their Cdc28-associated kinase activities are kept in check by the action of CDK inhibitor, Sic1.

Transcription of *CLB3* and *CLB4* peaks at the beginning of S-phase shortly after the *CLN1* and *CLN2* peak, and remains high until anaphase (Epstein and Cross, 1992). This periodicity of expression is reflected in the Cdc28:Clb3/4 kinase activity. Deletions of *CLB3* or *CLB4*, or *CLB3* and *CLB4* have no obvious cell cycle defects, but cells lacking *CLB3*, *CLB4* and *CLB5* are unable to form spindles and are inviable, suggesting that Clb3 and Clb4 perform a function that neither Cln1, Cln2 nor Clb6 can compensate for. Although Clb5 and Clb6 are partially able to initiate spindle formation, it is clear that the correct assembly of mitotic spindles requires the activity of Clb3 and Clb4, although in the

absence of Clb3 and Clb4 this function can be completed by Clb1 and Clb2 (Fitch *et al.*, 1992).

Clb1- and Clb2-associated Cdc28 kinase activity peaks just before anaphase and this is reflected in the peak of *CLB1* and *CLB2* transcription (Fitch *et al.*, 1992). In cells arrested in mitosis Cdc28 is predominantly complexed with Clb2. The importance of the Cdc28:Clb2 complex is reflected phenotypically as cells lacking Clb2 are larger than normal and cultures of these cells contain large numbers of budded G2 cells. Cells lacking Clb1 on the other hand have no detectable mitotic cell cycle phenotype. Loss of Clb2 in the absence of *CLB1* or *CLB3* is lethal, although cells lacking Clb2 and Clb4 or Clb2 and Clb5 are viable. Clb2 appears to be important in spindle elongation and interestingly inappropriate expression of Clb2 can induce DNA synthesis.

1.3. Cdc2 controls cell cycle progression in *S. pombe*.

The situation in fission yeast appears to be simpler than that in budding yeast, and the main cell cycle kinase, Cdc2, associates with only four cyclins (Cig1, Cig2, Puc1 and Cdc13), of which only Cdc13 is essential. As in budding yeast, the level of Cdc2 (Cdc28 in budding yeast), remains constant throughout the cell cycle while levels of cyclins oscillate. These oscillations in cyclin levels, and the actions of Cdc25 and Wee1, are largely responsible for the changes in Cdc2 kinase activity. The S-phase cyclin in fission yeast is Cig2, which accumulates just before S-phase and disappears as cells exit S-phase (Mondesert *et al.*, 1996). The fission yeast mitotic cyclin is Cdc13, the levels of which are low in G1, accumulate during G2 and are maintained until the end of mitosis (Moreno *et al.*, 1989).

1.4. Higher eukaryotic cell cycles are controlled by multiple CDKs and multiple cyclins.

The CDK family is much larger in eukaryotes, and four CDKs (1,2,4 and 6) appear to perform the functions of Cdc28 and cdc2. The CDKs 4 and 6 complexed with D-type cyclins integrate extra-cellular signals with the cell cycle machinery via pRb. Inactivation of pRb by CDK4 and 6 phosphorylation releases E2F and derepresses the transcription of E2F responsive genes, including those encoding cyclins E and A, as well as other genes required for the initiation of DNA synthesis (Mittnacht, 1998). Activation of CDK2 by

cyclin E leads to the further phosphorylation of pRb. CDK2 is also associated with cyclin A during S-phase, when the complex plays a poorly understood role in the regulation of DNA synthesis. CDK2:cyclin E controls its own activity by promoting the degradation of cyclin E, and in this way, cyclin E is self-limiting. The expression of E2F responsive genes needs to be restricted however, and this is achieved by phosphorylation of the E2F complex. Only CDK2 complexed with cyclin A can effectively phosphorylate E2F. The phosphorylation by CDK2:cyclin A, on either the DP-1 or E2F-1 components of the transcription factor abolishes DNA binding activity and trans-activation, halting expression of E2F responsive genes (Dynlacht *et al.*, 1994a). The timings of when different cyclins activate CDKs is shown in figure 1-1.

Entry into mitosis is regulated primarily by CDK1 associated with the B-type cyclins, although CDK1 can also form complexes with cyclin A, and there is good evidence that this complex plays a role in preparation for mitosis during late G2 as well as during early phases of mitosis itself. Mitosis requires major biochemical, architectural and physiological changes to the cell while DNA synthesis requires the formation and activation of a complex replication machine. That the CDKs are involved in the regulation of these events is no longer in question and much has been learned about the regulation and control of the cell cycle, and in particular, the regulation and control of CDKs and cyclins (figure 1-2). What is not clear, however, is how CDKs control the "downstream events" which lead to cell cycle events such as initiation of DNA synthesis. Few *bona fide* targets of CDKs have been identified. This is the focus of the work described in this thesis.

Are the CDKs the workhorses of the cell cycle machinery, acting directly on substrates responsible for inducing DNA synthesis, nuclear envelope breakdown, chromatin condensation and spindle formation? To some extent this is true; lamins are phosphorylated and the nuclear envelope disassembles (Peter *et al.*, 1990), condensins are phosphorylated and chromosomes condense (Kimura *et al.*, 1998) and the *cis* golgi matrix protein p130 is phosphorylated and the golgi fragments ((Lowe *et al.*, 1998). These substrates are typical of those identified to date: a protein is modified in a mitotic fashion and the kinases involved are identified. It is much harder to ask the question the other way around: what proteins do the kinases phosphorylate? The CDKs could also be the controllers of poorly defined signal cascades, acting as directors of undefined cell cycle

Figure 1-1. Cyclin associated CDK activity in higher eukaryotes.

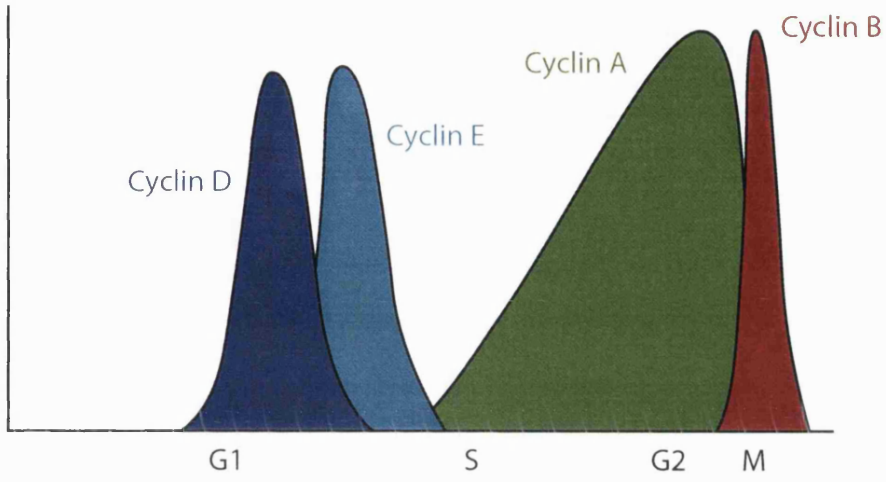


Figure 1-1. Cyclin D-associated CDK activity is induced in mid-late G1 in response to mitogens and is responsible for the transcriptional induction of cyclin E. Cyclin E appears late in G1 and is required for S-phase entry, and disappears shortly after the initiation of DNA replication. Cyclin A does not appear until DNA replication has initiated and gradually accumulates during G2. It is destroyed just before cyclin B. Cyclin B associated kinase activity appears just prior to mitosis and cyclin B is rapidly destroyed just after cyclin A at the metaphase to anaphase transition.

Figure 1-2. Schematic representation of cell cycle regulation in human cells.

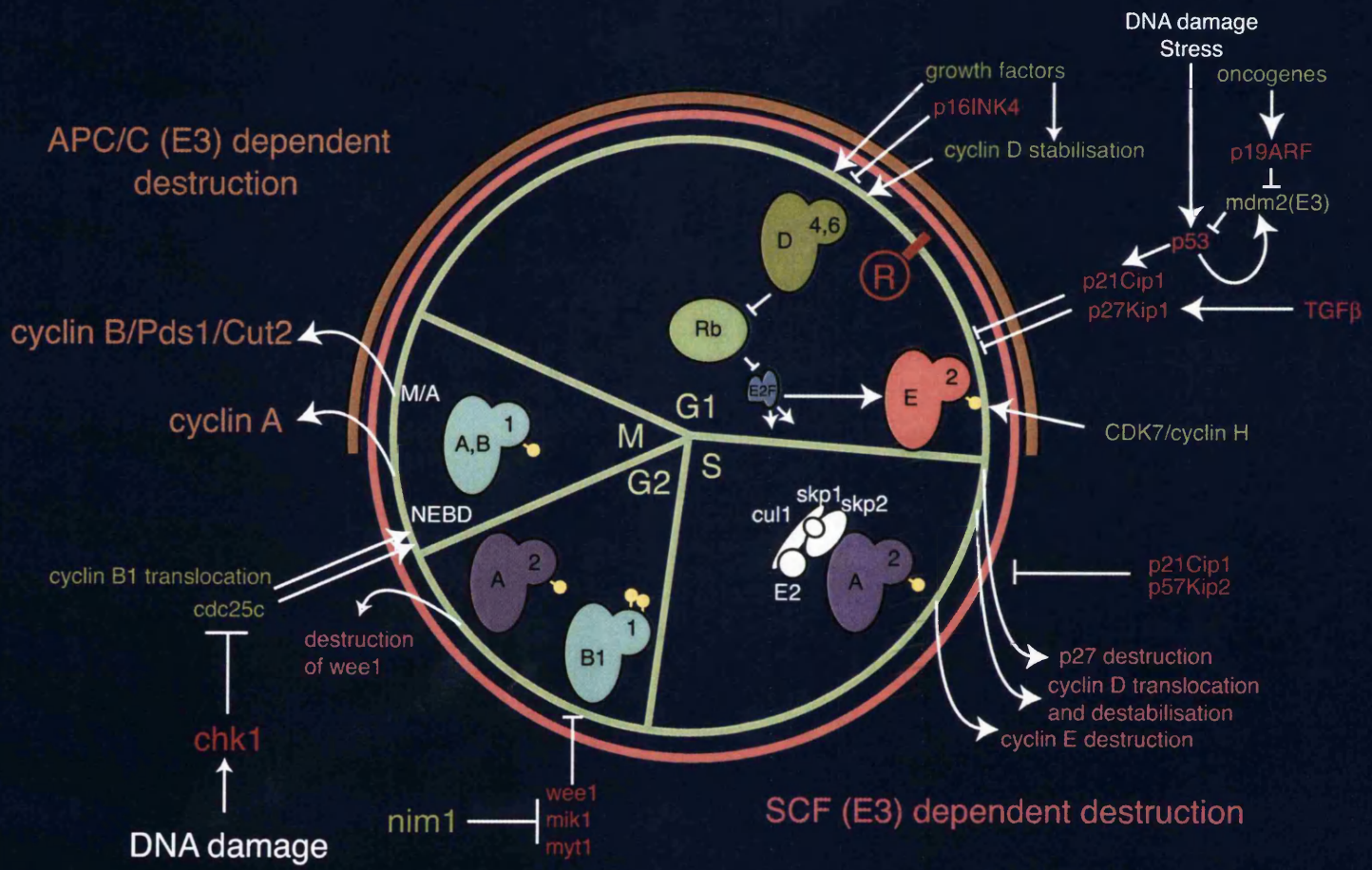


Figure 1-2. Representation of elements of cell cycle control in humans. Green: positive regulators, red: negative regulators, white: inducers of checkpoints, orange semi-circle: time when APC/C (E3) is active, pink circle: time when SCF (E3) is active. Yellow ball and sticks on kinase subunits indicate phosphorylations.

players whose performance results in chromosome duplication and division. CDK activity is ordered temporally, but what of its organisation spatially? How are CDK:cyclin complexes presented in the cell and are they juxtaposed with sites of S- and M-phase specific activity?

1.5. Why has it proven so difficult to identify substrates of CDKs?

The minimal CDK consensus phosphorylation site has been well characterised as S/TPXK/R. The question is, however, is S/TPXK/R all you need, or, are there other primary or secondary sequence determinants that are required for phosphorylation. Do all S/TPXK/R motifs get phosphorylated during mitosis and why cannot CDK2 complexed with cyclin A or E cause entry into mitosis? Identifying substrates of kinases and phosphatases is notoriously difficult and traditional biochemical approaches are time-consuming and laborious. Some success has been achieved by determining the preferred consensus target sequence of a kinase. Potential substrates can be identified by exploring protein sequence databases for a deduced phosphorylation consensus sequence. In cases where enzymes display high primary sequence requirements, and given recent advances in genomics and proteomics, such approaches warrant merit. Hall *et al* (2001) had success in scanning SWISSPROT for sequences which contained the CDK consensus phosphorylation sequence S/TPXK/R and the cyclin interaction motif RXL (Hall *et al.*, 2001). One of the many proteins they identified was HIRA, a homologue of the yeast proteins Hir1 and Hir2 which are involved in histone gene transcription. They showed that HIRA interacted with CDK:cyclin complexes in an RXL dependent manner, could be phosphorylated *in vitro* by CDK2:cyclin A, and localised to the nucleus where CDK2 is active. They also showed that overexpression of HIRA caused S-phase arrest and concluded that HIRA is an *in vivo* substrate of CDK2 containing complexes (Hall *et al.*, 2001). Many kinases such as the mitogen-activated (MAP), and cyclin-dependent (CDK) kinases have redundant or little primary sequence requirements however, and so "sequence-gazing" *per se* could be a misleading endeavour and more empirical assessments are required.

Using a peptide library containing 2.5 billion unique peptides and a fixed serine as the phosphoacceptor, Songyang *et al* (1994) investigated the primary sequence requirements of a panel of kinases including CDK1:cyclin B and CDK2:cyclin A. In agreement with previous studies they identified a sequence similar to the optimal consensus

sequence for CDK1/cyclin B: (K/R)SP(R/P)(R/K/H) (Songyang *et al.*, 1994). Holmes *et al* (1996) extended this type of study by using a series of glutathione-S-transferase (GST) fusions containing the canonical CDK1 phosphorylation site which was systematically altered to determine the fine primary sequence requirements for phosphorylation by CDK1 and CDK2 containing complexes. They found that there was great tolerance (no more than 2-fold differences in efficiency of phosphorylation) of substitutions at the –2 position relative to the phosphoacceptor site of the wild type consensus phosphorylation sequence (Holmes and Solomon, 1996). This is in contrast to the previously reported data, that posited a basic residue in this position. Substitutions at the +2 position were well tolerated however, although there were large differences (20 fold) in the rates of phosphorylation of substrates, with both polar and non-polar residues resulting in efficient phosphorylation. In terms of side-chain characteristics however, no discernible rule could be established for preferences at this position. The recently elucidated structure of CDK2:cyclin A complexed with a peptide substrate explains the preference for a basic residue at the +3 position with which Holmes *et al* (1996) concur. Interestingly, they show that the substitution of a basic residue (K/R) with an alanine, results in a weakening of the interaction between substrate and kinase by approximately 1.8 kcal/mol, which corresponds to the loss of a single ionic interaction, such as that between a basic residue and a negatively charged molecule such as phosphate. The significance of this is discussed below. This type of predictive scale for evaluating potential phosphorylation sites represents an unbiased starting point for establishing the theoretical ability of a protein for being a good CDK substrate.

Theoretically, excellent consensus phosphorylation sites may be inappropriately folded, preventing docking in the active site of a kinase, or may not be exposed to solvent which would also preclude phosphorylation. So although a predictive scale may be a useful starting point to assess the 'quality' of a potential substrate, it should not be considered a substitute for the empirical demonstration that a protein can be phosphorylated by a kinase *in vitro* and *in vivo*. It is unlikely therefore that substrates can be identified based solely on their primary sequences or on the presence of a consensus sequence which is recognised with little specificity. In some cases however, the analysis of primary sequences of known substrates and cyclin-interacting proteins has allowed the identification of motifs involved in substrate specificity or recruitment (Adams *et al.*, 1996). Moreover, some substrates,

including histone H1 which is often used as the prototypical CDK substrate in *in vitro* kinase assays, can be phosphorylated on serine and threonine residues not contained within a consensus phosphorylation motif (Gurley *et al.*, 1995).

Co-crystallisation of phospho-CDK2:cyclin A3 and a synthetic substrate peptide (HHASPRK) revealed the nature of the interactions between substrate and the active site of CDK2 (Brown *et al.*, 1999). Binding of cyclin A to unphosphorylated CDK2 causes significant changes in the conformation of CDK2. Specifically, the activation segment (residues D145 to E172) and N- and C-terminal lobes become rearranged to facilitate binding of protein and ATP substrates in the active cleft. Phosphorylation on T160 of CDK2 causes further changes in the activation segment which realise full activity of the kinase. The phosphorylation of T160 and binding of cyclin A are required to accommodate a proline in the +1 position relative to the phosphoacceptor site (figure 1-3). Arginine in the +2 position relative to the phosphoacceptor makes no contact with CDK2 and has a solvent-exposed side chain. The major site of interaction between the substrate peptide and the CDK2 and cyclin A is via the lysine in the +3 position. The sidechain of this lysine is hydrogen bonded with the phosphate of T160 and also makes contacts with I270 of cyclin A3. These contacts explain the preference for a basic residue in the +3 position. On the N-terminal side of the phosphoaccepting serine, histidine in the –3 position contacts W167 and the histidine in the –2 position contacts the main chain of G205. These observations explain the observed requirements for proline in the +1 position and a basic residue in the +3 position.

In these structures (figure 1-3), AMPPNP is bound in the cleft between the N-terminal and C-terminal lobes of CDK2 (Brown *et al.*, 1999). The phosphoaccepting serine is hydrogen bonded to the catalytic residues D127 and K129, while the γ-phosphate of AMPPNP is oriented toward the phosphoacceptor. A single Mg^{2+} is present which chelates the α and γ phosphates of AMPPNP but is not shown in figure 1-3.

It is possible, and even likely, that CDKs recognise and phosphorylate multiple substrates to regulate cell cycle events. This has meant that genetic analysis in yeast has not shed much light on the important questions. A search for genetic suppressors of CDK mutations is unlikely to be an effective approach to identify downstream targets in complex

onto origins of replication, which is an essential event in priming cells for S-phase. When CDK activity falls after mitosis, it becomes possible to load the origins with the appropriate factors. 'Firing' the origins requires an increase in CDK activity, however, which prevents new origin loading. Once replication has occurred, and origins are cleared of these proteins, there can be no reloading of the origins until Cdc28:Clb activity levels drop again. Cells with a high Cdc28:Clb activity are prevented from loading origins which prevents re-replication without an intervening mitosis. According to this model, Cdc28:Clb2 activity oscillations drive the alternation of DNA replication and segregation and prevent the re-initiation of DNA synthesis until after a mitosis has been successfully completed. The drop in the activity of Cdc28:Clb levels after mitosis then permit the reloading of origins for a new round of DNA replication.

Stern and Nurse (1996) have proposed a similar model for how these alternations are generated in fission yeast (Stern and Nurse, 1996). They suggest that a moderate level of Cdc2:Cdc13 kinase activity is required for cells to initiate S-phase. Higher kinase levels are required for cells to enter mitosis. As in budding yeast, high levels of Cdc2:Cdc13 inhibit origin loading and as long as the kinase levels following S-phase are above a 'threshold' the re-replication of the genome is prevented. These differences in Cdc2:Cdc13 kinase levels and the fact that as in budding yeast, levels of Cdk:cyclin activity must be abolished for the successful completion of mitosis, ensures that cells replicate their DNA only once before committing to cytokinesis. In agreement with this model, deletion of *CDC13* leads to re-replication and failure to perform mitosis, which suggests that high Cdc2:Cdc13 inhibits the loading of proteins at origins required for initiation of DNA synthesis.

These observations have lead to the suggestion that cyclins act only as activators of CDK activity, and that no particular cell cycle function can be attributed to a specific cyclin protein (Stern and Nurse, 1996). Contrary to this model, and from experiments primarily in budding yeast and animal cells, is the argument that the nature of the cyclin may help to dictate the identity of the substrates with which CDK interacts. In this way cyclins are qualitatively different from one another and affect the way in which bound CDK behaves, while in the first model there is no qualitative difference between cyclins and it is the amount, and timing of appearance of kinase activity which is important. It is also possible,

A number of experiments, mainly using yeast cells, have shown that an inappropriately expressed cyclin could rescue cell cycle defects associated with the deletion of other cyclins. Firstly, as described above was the clear demonstration in fission yeast that deletion of all cyclins excepting Cdc13 caused no serious cell cycle defects. In budding yeast, the inviability of cells lacking *CLNs1-3* could be rescued by the overexpression of mammalian cyclins A, B, D and E, or the mitotic cyclin *CLB5* (Epstein and Cross, 1992); also, the lethality of deleting all B-type cyclins was rescued by singly overexpressing *CLB1* (Haase and Reed, 1999). These experiments do not necessarily support the hypothesis that cyclin specificity is not required for proper cell cycle progress if another cyclin is substituted. In all but the fission yeast experiments, the rescue of the phenotypes associated with the deletion of cyclins has required the overexpression of another cyclin. It is possible that the deleted cyclins were providing specialised functions at particular cellular locations, and that overexpression of another cyclin merely floods the cell with Cdc28-associated kinase activity. Indeed, *CLN* deficient budding yeast cells were rescued by the expression of truncated cyclins A, B, D and E in G1, indicating that almost any cyclin able to activate Cdc28 kinase can allow cells to complete G1 (Leopold and O'Farrell, 1991; Lew *et al.*, 1991). These experiments do not disprove a requirement for specific cyclins for particular cell cycle events, but support the idea that a minimum amount of kinase is all that is required for proper cell cycle events. These two hypothesis are difficult to segregate and test experimentally although the observations that cyclins can induce stage specific events independent of timing of expression argues against a substrate availability model as the sole regulator of substrate specificity.

A more persuasive experiment in budding yeast suggesting that cyclin specificity is not an important factor in cell cycle control involved the identification of mutant Cdc28 which was activated in the absence of cyclins (Levine *et al.*, 1999). Cells lacking *CLN1* and *CLN3*, and containing a copy of *CLN2* which was deficient in binding and activating Cdc28, but which contained specific mutations in Cdc28 retained their ability to transit Start efficiently. These data showed that Cdc28 could be activated in a cyclin independent manner, in which case Cln functions were no longer required. The activated Cdc28 was not able to perform essential Clb functions later in the cell cycle (Levine *et al.*, 1999). It seems unlikely that further 'evolution' of Cdc28, allowing it to provide essential Clb functions could occur however as there would be no cyclin-dependent drop in the levels of Cdc28

activity at the end of mitosis, which would lead to arrest in anaphase. Moreover, due to the absence of a drop in Cdc28 activity following mitosis, origins would not be reloaded and there would be no ensuing S-phase. This does not preclude, however, the concomitant 'evolution' of some other kind of periodic Cdc28 control mechanism that could allow a drop in the levels of Cdc28-associated kinase activity in G1 to allow origin loading.

In fact, Cdc28 kinase activity is often regulated in manner that clearly depends on the type of cyclin to which it is bound. For example, Cdc28:Clb kinase activity is inhibited by the CKI Sic1, yet Cdc28:Cln activity is refractory to this inhibitor which ultimately leads to the degradation of Sic1; realising the kinase activity of Cdc28:Clb complexes which allows entry into S-phase (Mendenhall *et al.*, 1995). One interpretation of this phenomenon is that Clb-associated Cdc28 has a different function from Cln-associated Cdc28 kinase. Alternatively, it can be argued that it is not the output of the kinase which is required for a particular function, but that certain CDK:cyclin pairs are not able to perform certain functions due to differences in the ways they are regulated. Similarly, in fission yeast the CKI Rum1 is able to inhibit the kinase activity of Cdc2:Cdc13 and Cdc2:Cig2 complexes but not Cig1 or Puc1 containing complexes. Any apparent differences in abilities to promote cell cycle events may be due to the differences in the ways these complexes are regulated. In animal cells phosphorylation of pRb by complexes containing CDKs 4 and 6 complexed with D-type cyclins is inhibited by the p21/p27 and INK4 classes of CKI (Harper *et al.*, 1995). Inhibition of pRb phosphorylation prevents the release of the E2F transcription factor and thus restricts the transcriptional cascades leading to expression of cyclin E (and eventually cyclin A) which in complex with CDK2 lead to S-phase entry (Dynlacht *et al.*, 1994b). Some viral cyclins however, have evolved a way to usurp this control by being refractory to inhibition by p21/p27 and INK4 when bound to CDK4 and 6. This leads to the untimely inactivation of pRb and promotion of the G1 transcriptional cascade causing amongst other things, the replication of the viral genome (Swanton *et al.*, 1997).

Observations in our own laboratory have indicated that in frog egg extracts an excess of cyclin A3 can lead to mitosis, an event normally promoted by B-type cyclins (Jon Moore, personal communication), and that in this system the nature of the cyclin activating CDK1 is irrelevant as long as a minimum amount of kinase activity is realised. An

interesting speculation is whether low levels of CDK1:cyclin B are able to initiate S-phase in frog egg extracts. This is a technically difficult experiment however as intermediate levels of kinase activity must be generated via modulation of the Wee1/Cdc25 regulatory loop and moreover, the activated kinase must be accumulated in the nuclei of sperm added to frog egg extracts depleted of cyclin E.

In summary, there is ample evidence, mainly in budding and fission yeasts which suggests that indeed there may be no specific requirement for a certain cyclin at a particular point in the cell cycle. Under some circumstances it has been shown that the nature of the cyclin is not important and that as long as sufficient kinase activity can be generated then this is sufficient to ensure the orderly alternation of DNA replication and segregation. A simplified model showing how only quantitative differences in CDK:cyclin activity might drive a minimal cell cycle is depicted in figure 1-4. In this model, minimal amounts of kinase activity are required to cause entry into, and passage through, S-phase, which could be due to the presence of few substrates, while entry into mitosis requires higher levels of kinase activity, which could in turn be due to the presence of more substrates.

1.9. Are cyclins specialised for particular functions?

The experiments showing that inappropriate expression of a cyclin in the absence of other cyclins still resulted in a relatively normal cell cycle could be regarded as flawed, as the inappropriately expressed cyclin was overexpressed and not expressed from its own, or perhaps more appropriately from one of the deleted cyclin promoters. Levine *et al* (1996) reported how the question of cyclin specificity could be addressed by placing cyclin coding sequences under the control of a different cyclin promoter. Thus differences between the abilities of specific cyclins to induce particular cell cycle events could be tested (Levine *et al.*, 1996).

For example, when *CLN2* was placed under the control of the *CLN3* promoter there were differences in cell viability, which presumably reflected the inability of Cln2 to perform some functions of Cln3. Thus when *CLN2* is expressed at a time when Cln3 function is required, Cln2 cannot compensate for the lack of Cln3 protein. The differences in the timing of expression do not completely account for differences in the ability of Cln2 and Cln3 to perform cell cycle functions (Levine *et al.*, 1996). It appeared that there was

Figure 1-4. A quantitative model of a minimal cell cycle.

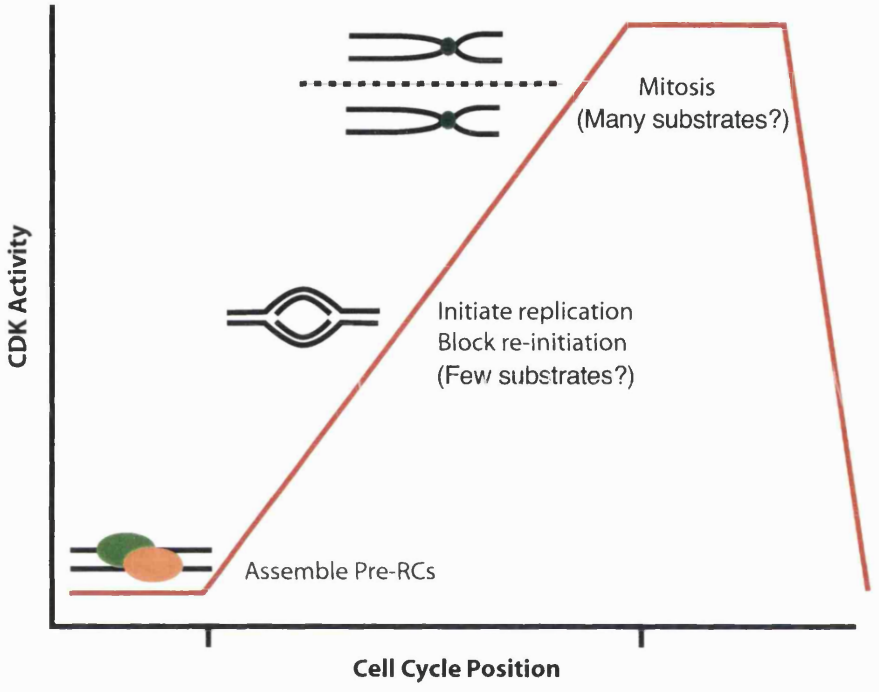


Figure 1-4. Model showing how cell cycles could progress by only quantitative changes in the amount of CDK:cyclin activity. With low kinase activity proteins required for the initiation of DNA replication are loaded onto origins or replication. Initiation requires an elevated level of kinase activity, which must also be sufficiently high to prevent the re-loading, and therefore, reinitation of replication. When kinase activity reaches a sufficiently high level, mitosis is induced. Completion of mitosis requires that CDK:cyclin activity drops, which as also a requisite for the re-loading of origins and an ensuing S-phase. These changes alone are hyothesised to be sufficient to allow the alternation of DNA replication and segregation.

an inherent difference between the two proteins in their abilities to perform different G1 functions. Both cyclins are able to support transition through Start, although they possibly achieve this through distinct mechanisms in that Cln3 acts primarily as a transcriptional activator (including of *CLN1* and *CLN2*), and that the Cln1/2 pair control additional G1 events such as bud emergence and activation of Cdc28:Clb complexes (Dirick and Nasmyth, 1991; Lew and Reed, 1993). So this is not really a proper case of functional redundancy.

A similar phenomenon was observed when cyclin E was placed under the control of the cyclin D1 promoter in cyclin D1 knockout mice (Geng *et al.*, 1999). Knock-in of the cyclin E coding sequence under the control of the cyclin D1 promoter in knockout mice led to rescue of the phenotype associated with loss of cyclin D1. This appears to be a case of functional redundancy between cyclins, and could be used to argue for a case of quantitative cell cycle control, and is strengthened by the fact that expression of cyclin E occurred at the same time and in the same amounts as cyclin D1. It appears however that the rescue of the cyclin $D1^{-/-}$ phenotype by expression of cyclin E is not due to functional redundancy, but is in fact due to the bypass of the requirement for cyclin D1. Normally, cyclin D1 expression leads to the expression of cyclin E and so in this case, cyclin E was not substituting for cyclin D1 functionally.

Cross *et al* (1999) have performed similar experiments to those of Levine *et al* (1996) where one cyclin is placed under the control of another cyclin's promoter. Cross *et al* (1999) performed an analysis of the effects of placing the coding sequence of *CLB2* under control of the *CLB5* promoter which means that cyclin function can be examined independently of timing of expression (Cross *et al.*, 1999). Cells lacking *CLB1* and *CLB2* are inviable, although this was rescued in cells where *CLB2* was under control of the *CLB5* promoter, but interestingly DNA replication was poorly promoted. This suggested that a function of Clb5, namely the promotion of initiation of DNA replication, was only poorly performed by Clb2 and therefore that there is something special about Clb5 which is required for initiation of DNA replication (Cross *et al.*, 1999). A mutant Clb2, which cannot be degraded by the APC blocks cells in mitosis, while a Clb5 mutant also unable to be degraded by the APC and expressed to similar levels to Clb2 does not have a mitotic

arrest phenotype. Clearly, Clb5 is unable to perform a function of Clb2 which prevents exit from mitosis, while both these proteins are able to prevent the loading of origins.

Kim and Kaelin (2001) have reported how cyclin A and cyclin E are differently able to regulate transcription when tethered to DNA. A number of transcriptional regulators, including the pRb homologues p107 and p130, E2F1-3, p300 and NPAT, are known to form stable complexes with either CDK2:cyclin A or CDK2:cyclin E (Ewen *et al.*, 1992; Mudryj *et al.*, 1991; Perkins *et al.*, 1997; Zhao *et al.*, 1998). These interactions mean that CDK2:cyclin A and CDK2:cyclin E can become concentrated at specific genomic loci where they could regulate events such as DNA replication and transcription (Kim and Kaelin, 2001). A plasmid encoding either cyclin A or cyclin E fused to the DNA binding domain of the tetracycline repressor was transfected into cells harbouring a plasmid containing 7 copies of the tetracycline operator sequence upstream of a minimal CMV promoter containing a TATA box, which was in turn upstream of the coding sequence for luciferase. The effect of the cyclins on transcription from the CMV promoter was measured using luciferase.

In this assay cyclin A repressed expression of luciferase nearly as efficiently as when pRb was used, while cycle E stimulated transcription approximately 10 fold, nearly half as efficiently as when E2F was used. The activation of transcription by cyclin E was shown to occur in a dose dependent fashion based on the number of tetracycline operator sequences present on the luciferase expressing plasmid and on the amount of cyclin E expressed in transfected cells. The ability of cyclin E to stimulate transcription was also shown to be dependent on its ability to bind CDK2 and to phosphorylate substrates. Strangely though deletion of the N-terminal 50 residues of cyclin E nearly doubled its ability to activate transcription, without also stimulating its ability to phosphorylate substrates. Cyclin A on the other hand was able to repress transcription in this assay independently of its ability to bind CDK2. Consistent with these data, co-transfection of plasmids encoding dominant-negative CDK2 (dn CDK2) and TETr:cyclin E resulted in a reduction of the activation of transcription relative to when plasmid encoding TETr:cyclin E was transfected alone. Co-transfection of plasmid encoding dn CDK2 and TETr:cyclin A however, had little effect on the ability of cyclin A to repress transcription in this system. (Kim and Kaelin, 2001).

It appears that the concentration of either cyclin A or cyclin E at specific sites in the genome could cause a decrease or increase in the levels of transcription respectively, and that this is a consequence of their inherent abilities to form complexes with certain DNA binding proteins. Cyclin A and cyclin E are functionally distinct in their abilities to repress and activate transcription respectively. A functional model to explain these differences supposes that in late G1-phase cyclin E acts to stimulate transcription of genes needed to enter S-phase, including cyclin A, but that as CDK2:cyclin E complexes disappear and CDK2:cyclin A complexes accumulate these act to repress transcription as cells traverse S-phase. It has also been reported that complexes of CDK1:cyclin B also inhibit transcription during mitosis (Leresche *et al.*, 1996). Although the temporal appearance and disappearance of cyclins during the cell cycle may control transcription, it appears also that cyclin A and cyclin E differ fundamentally in their abilities to regulate transcription (Kim and Kaelin, 2001). How the cell cycle might be regulated in a qualitative fashion is shown schematically in figure 1-5.

1.10. The role of localisation in generating specificity.

The timing of expression of a cyclin is unlikely to be the sole dictator of specificity, even when there is a close correlation between timing of expression and co-ordination of a cell cycle event. The targeting of CDK activity could also be dictated by cyclin-dependent protein-protein interactions, or by the cyclin-dependent targeting of CDK to specific structures or compartments within the cell. For example, the human CAK enzyme CDK7:cyclin H:MAT1 (Tassan *et al.*, 1994), is always, but not exclusively found complexed with the class II transcription factor TFIIH which recruits RNA polymerase II to promoters (Tassan *et al.*, 1995). TFIIH contains a number of subunits including XPB and XPD which play roles in NER (Cappelli *et al.*, 1999), and is capable of phosphorylating the CTD of RNA polymerase II. Mice lacking MAT1 are inviable, although mitotic cell lineages derived from these animals showed cell cycle arrest characterised by an inability to enter S-phase. Analysis of the CTD of RNA polymerase II showed that serine 2 and serine 5 were not phosphorylated, although this had no affect on transcriptional (or translational) integrity, suggesting that the CDK7:cyclin H:MAT1 does not play a role in regulating RNA polymerase II dependent transcription in normally cycling cells

Figure 1-5. A qualitative model of cell cycle regulation.

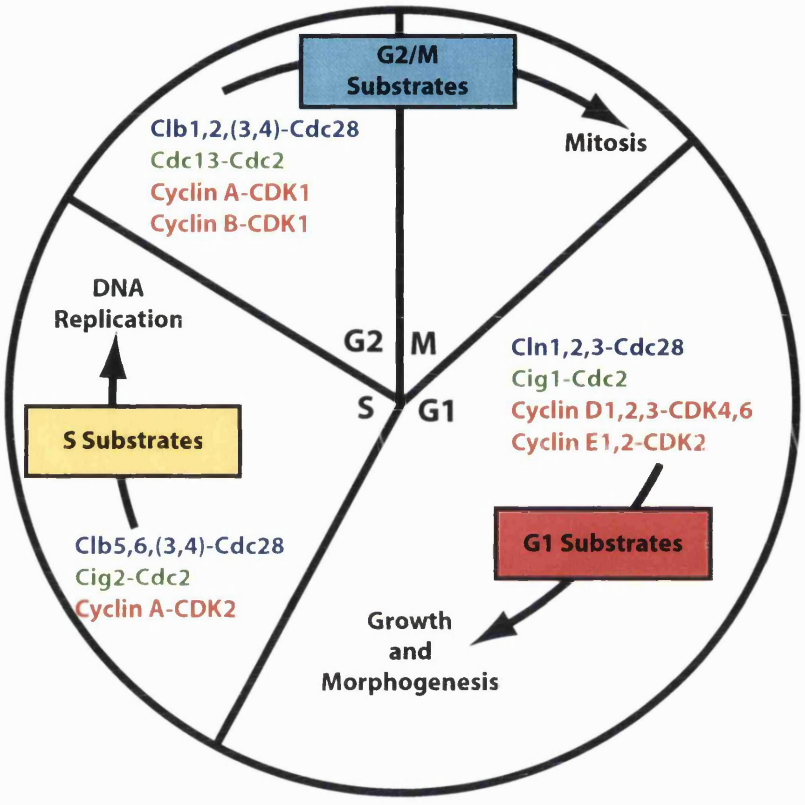


Figure 1-5. Cell cycle regulation by multiple CDK:cyclin complexes. Different CDK:cyclin complexes regulate transition through certain cell cycle stages. In human cells cyclins D1, D2 and D3 complexed with CDKs 4 and 6 regulate transition from G_0 into and through G1, while in *S. cerevisiae* transition through G1 and committment to the cell division cycle is under control of Clns1, 2 and 3 complexed with Cdc28 and in *S. pombe* passage through G1 is controlled by Cdc2:Cig1 complexes. Entry into, and passage through S-phase in human cells is controlled by CDK2:cyclin E and CDK2:cyclin A respectively. In *S. cerevisiae* these events are under the control of Clb5 and 6 complexed with Cdc28, although some functions may be performed by Clbs3 and 4 complexed with Cdc28. S-phase in *S. pombe* is under the control of Cdc2:Cig2 complexes. Mitotic entry is under the control of CDK1 complexed with cyclins A and B in human cells, while in *S. cerevisiae* these functions are under the control of primarily Cdc2 complexed with Clbs1 and 2. In *S. pombe* mitotic events are under the control of Cdc28:Cdc13. Red text indicates human complexes, green text represents *S. pombe* complexes and blue text represents complexes of *S. cerevisiae*.

(Rossi *et al.*, 2001). When cells were irradiated with UV however, there was a considerable reduction in

the amount of CDK7:cyclin H:MAT1 present in TFIIH and it has been suggested that this could represent a link between DNA damage and transcriptional NER (Adamczewski *et al.*, 1996). In this example, the localisation of CDK7:cyclin H in TFIIH is dependent on the presence of MAT1 and so although not cyclin dependent, the localisation of CDK7 appears to be under the control of a CDK:cyclin associated protein, and the role of CDK7:cyclin H in properly co-ordinating cell cycle progression appears to be dependent on the presence of MAT1.

If cyclins are involved in the targeting of CDK activity to particular locations in the cell, then this may in some ways act like temporal regulation of cyclin expression, in that the exposure of certain substrates to active kinase is limited. Potentially a significant contribution of localisation to cyclin specificity could be due also to the sequestration of certain complexes away from substrates whose inappropriate phosphorylation could have adverse effects on the cell. As was mentioned before the apparent target specificity of some cyclins could be a reflection of the regulatory controls which are exerted over certain CDK:cyclin complexes but not others. This could be achieved in part by localisation, where exposure to certain regulators of activity is dependent on the localisation of a CDK:cyclin complex to an appropriate subcellular location. It begins to appear that the specific output of a CDK:cyclin complex is the sum of a number of factors acting through decidedly non-linear pathways.

A dependence on cyclin localisation for some cell cycle events has been demonstrated and as might be predicted cyclins localise at the structures and in the compartments where they need to act (Hood *et al.*, 2001; Pines and Hunter, 1991). Possibly the best characterised correlation between cyclin localisation and specific cell cycle events is the localisation of cyclin B1 and cyclin B2 to microtubules and the golgi respectively. CDK1 complexed with cyclin B2 can phosphorylate the cis-golgi matrix protein GM130 and this phosphorylation prevents interaction of GM130 with the vesicle docking protein p115. Disruption of this interaction leads to the mitotic fragmentation of the golgi which is required for the partitioning of the organelle into the two new daughter cells (Lowe *et al.*, 1998). Cyclin B1 localises to microtubules, however, (Jackman *et al.*,

1995) and in starfish oocytes cyclin B1 is reported to be associated with MAP4 (Ookata *et al.*, 1995). The functional significance of the microtubule specific localisation of cyclin B1 is unclear but the different mitotic localisation of cyclin B1 and cyclin B2 suggests distinct mitotic functions.

Draviam *et al* (2001) have described the elements in cyclin B1 and B2 which regulate the distinct subcellular localisations which these proteins display. By switching the N-termini of the two cyclins to create hybrid cyclins in which cyclin B1 bore the N-terminus of cyclin B2 and cyclin B2 bore the N-terminus of cyclin B1 they showed that the localisation of these proteins could be switched (Draviam *et al.*, 2001). Replacement of the N-terminus of cyclin B1 with that of cyclin B2 caused a dramatic re-localisation of cyclin B1 from microtubules to the golgi. Consequently, cyclin B1 was able to promote the fragmentation of the golgi. Cyclin B2 whose N-terminus was replaced by that of cyclin B1 showed localisation on microtubules and resulted in changes in cytoskeletal architecture (Draviam *et al.*, 2001). These striking results show that specificity of these cyclins for inducing certain cell cycle events depends on the particular subcellular localisation of a cyclin. These data could be used to argue that it is not specifically Cdc2:cyclin B1 that is responsible for reorganising the cytoskeleton, or not specifically Cdc2:cyclin B2 which fragments the golgi because simply swapping their localisations swaps their supposedly specific functions. It is an intrinsic quality, namely the N-termini, of cyclin B1 and cyclin B2 however which is essential for this localisation. It is this intrinsic quality which dictates where these cyclins localise and therefore where Cdc2 is activated to cause specific cell cycle events.

In *S. cerevisiae* the G1 cyclins Cln2 and Cln3 localise to different compartments. Cln2 is predominantly a cytoplasmic protein and Cln3 is predominantly nuclear. Cln3 is proposed to be involved in causing the transcription of genes, including *CLN1* and *CLN2*, which are needed for cells to traverse start, and so the localisation of Cln3 in the nucleus is expected (Miller and Cross, 2000). Cln3 is transported to the nucleus in a NLS-dependent fashion, and removal of the NLS causes Cln3 to accumulate in the cytoplasm. A mutant of Cln3 lacking an NLS is equally good at rescuing certain strains as Cln2 and this rescue can be abrogated by the addition of an exogenous NLS to Cln3 lacking it's own NLS (Miller and Cross, 2001). These data suggest that it is the particular localisation of Cln3 which is

important for its specific function, and moreover that altering the subcellular localisation of Cln3 can alter its abilities to mimic those of Cln2. This indicates that, besides the temporal differences in expression of Cln2 and Cln3, they are exposed to different substrates and it is this difference in exposure which accounts for their ability to promote distinct cell cycle events.

In simple eukaryotes it has been clearly shown that a minimal cell cycle can occur relatively normally in the absence of mulitple cyclins (Haase and Reed, 1999; Nasmyth, 1996). Modern higher eukaryotes have significantly more complex genomes however and it seems likely that multiple cyclins have evolved to facilitate the replication and segregation of complex genomes. In higher eukaryotes a qualitative mode of cell cycle regulation, where cyclins regulate CDK activity by being variably expressed, localised, responsive to regulatory inputs and are capable of making different protein-protein interactions, is a much simpler model to rationalise than is quantitative mode of regulation. Even in *S. cerevisiae* a qualitative mode of cell cycle regulation can be evoked and it is important to bear in mind that *S. cerevisiae* and *S. pombe* are as evolutionarily diverged as the yeasts and humans are. It is possible that a prototypical B-type cyclin controlled both DNA replication and segregation in a primordial eukaryotic cell. The six *CLB*s of budding yeast possibly evolved as a consequence of gene duplication events and that functional specialisation followed as genomes became more complex. The function of a Clb parallels its timing of expression, and although the difference in functions of different Clbs could be attributed to different timing of expression of transposable cyclins, it is also suggestive that divergent cyclins could have evolved to fulfil specific cell cycle demands (Nasmyth, 1995).

1.11. Substrate recruitment to cyclin.

At the centre of the qualitative mode of cell cycle regulation is the question: How do CDK:cyclin complexes specifically recognise and recruit substrates? Krek *et al* (1994) described a 42 residue peptide derived from E2F1 (residues 67-108) which could specifically block the interaction of *in vitro* translated cyclin A with GST-E2F1 (Krek *et al.*, 1994). Based on this work Adams *et al* (1996) found a minimal 7mer peptide (PVKRRLD) which also efficiently blocked the interaction of cyclin A and E2F1 (Adams *et al.*, 1996). Sequence analysis of multiple cell cycle regulatory and cyclin-binding proteins found that they all contained sequences similar to the one able to block interaction

between cyclin A and E2F1. Alignments of these sequences showed a conserved core motif defined as ZRXL, where Z tends to be basic or cysteine and X tends to be basic. Peptides spanning the ZRXL motifs in p107, p130 and p21 were shown to be able to block the binding of E2F1 to cyclin A and peptides spanning the ZRXL motifs in E2F1, p21, p27, p57, p107 and p130 blocked the binding of cyclin A and cyclin E to GST-p21. Unlike cyclin A, cyclin E has not been found in complexes with E2F1 and therefore it is possible that the RXL motif may play a role in mediating cyclin-substrate interactions. Specificity may be context dependent however and sequences outside the minimal ZRXL motif could play a role in modulating specificity. None of the peptides capable of blocking the interaction between cyclin A or cyclin E and E2F1 or GST-p21 respectively were able to block the interaction of cyclin B and GST-p21 however. There may be RXL-independent sequence elements in p21 which are responsible for mediating cyclin B-p21 interactions, although these remain to be identified.

The E2F1-derived minimal sequence capable of blocking interaction between cyclin A and E2F1 is often abbreviated to the 'RXL motif' as these are the residues which are best conserved in alignments of cyclin binding proteins. Ironically however, alanine scanning mutagenesis of the peptide showed that substitution of the R and L of the RXL motif are well tolerated (Adams *et al.*, 1996). Substitutions at the –1, -2 and +4 positions relative to the R of RXL on the other hand are not well tolerated and show greatly reduced abilities to block the binding of cyclin A to GST-E2F1. That RXL-containing sequence elements are important in mediating cyclin-substrate interactions was shown by deleting them from known substrates. Substrates which lacked their RXL-containing motifs were poorly phosphorylated by CDK:cyclin complexes. These data show directly that the ZRXL motifs specifically target substrates to CDK:cyclin complexes. Further to this, it has been shown that binding of p21 and p107 to cyclin A and cyclin E is mutually exclusive (Dimri *et al.*, 1996) and that p21 can disrupt the binding of p130 to CDK:cyclin complexes (Shiyanov *et al.*, 1996).

Possession of a consensus CDK phosphorylation motif is not necessary to ensure efficient phosphorylation if substrates can bind to CDK:cyclin complexes via an RXL motif. The presence of independent phosphoacceptor and cyclin interaction sites means that these elements could in theory be on different but physically associated proteins. In

this way, one protein containing a RXL motif could act as an adapter molecule for another protein containing a consensus CDK phosphorylation motif. It was also shown that RXL-containing protein substrates were able to stabilise CDK:cyclin complexes. It is a theoretical possibility that the presence of different substrates at different times during the cell cycle could lead to the assembly and/or stabilisation of different CDK:cyclin complexes. This could add yet another level of complexity to a qualitative model of cell cycle regulation.

1.12. The hydrophobic patch of cyclin A.

The identification of the RXL motif in multiple cell cycle regulatory and cyclin binding proteins required only the analysis of the primary sequences of these proteins. Schulman *et al* (1998) extended this type of analysis to the tertiary structure of cyclin A in the hope of identifying elements which were involved in protein-protein interaction. Detailed analysis of the crystal structure of CDK2 complexed with cyclin A identified a hydrophobic patch on the surface of cyclin A (Schulman *et al.*, 1998). The residues which compose the hydrophobic patch are well conserved between cyclins and some of these residues fall within the characteristic MRAIL sequence. It was suggested that, because the majority of intermolecular binding energies are contributed by hydrophobic interactions (Clackson and Wells, 1995), this hydrophobic patch could be involved in the recruitment of proteins, possibly substrates, to cyclin. Supporting this hypothesis was the observation that, in the crystal structure of CDK2 complexed with cyclin A and p27, the RNFL motif of p27 contacts the hydrophobic patch of cyclin A (Jeffrey *et al.*, 1995b).

The crystal structure of phospho-CDK2:cyclin A3 complexed with a peptide derived from p107 which contained the RXL recruitment motif (1-RRLFGEDPPKE-11) showed that the recruitment peptide bound in the hydrophobic groove of cyclin A (comprising residues M210, I213, L214, W217, E220, L253 and Q254) (figure 1-3). The side chain of 1R formed ion pairs with E220 of cyclin A3 while 2R made no contacts with any hydrophobic patch residue. Residue 3L makes a hydrogen bond and extensive van der Waals interactions with residue Q254 and van der Waal interactions with W217. Residue 4F makes extensive van der Waals contacts with cyclin A3 residues M210, I213, and L253, while the sequence C-terminal to residue 4F in the peptide was not apparent in the crystal

structure suggesting that this sequence is not important in contributing to specificity (Brown *et al.*, 1999).

The crystal structure of cyclin A showed that the residues M210, I213, L214, W217, L253 and Q254 contribute sidechains to the hydrophobic patch of cyclin A (figure 1-3). In human cyclin E and human cyclin B these residues are completely conserved while in human cyclin D1 L214 of cyclin A is replaced by a valine. Cyclin A which had been mutated in the hydrophobic patch (M210A, L214A and W217A) was called hpm cyclin A. Immunoprecipitates of cyclin A from extracts prepared from cells transfected with plasmid encoding HA-tagged cyclin A contained p107 while immunoprecipates from cells expressing HA-tagged hpm cyclin A did not. When lysates of cells expressing HA-tagged cyclin A were mixed with recombinant, bacterially produced E2F1, cyclin A could be recovered in immunoprecipitations using anti-E2F1 antibodies and immunoprecipates using antibodies against the HA-tag contained GST-E2F1 (Schulman *et al.*, 1998). HA-tagged hpm cyclin A was able to bind E2F1 in this assay with greatly reduced affinity. Similarly, p21 was unable to interact with hpm cyclin A alone, but was able to interact with cyclin A alone, or cyclin A complexed with CDK2. These data suggest that the hydrophobic patch of cyclin A is an important structural element in the recruitment of substrates to cyclin A. They also indicate that the hydrophobic patch of cyclin A is the only component of cyclin A which is required for its interaction with p21.

Immunoprecipitates of HA-tagged cyclin A or HA-tagged hpm cyclin A from cells were tested for their ability to phosphorylate a panel of substrates. Immunoprecipitates containing HA-tagged cyclin A were able to phosphorylate histone H1, p107, E2F1 and the C-terminus of pRb. Immunoprecipitates containing HA-tagged hpm cyclin A however were deficient in their ability to phosphorylate p107, E2F1 and the C-terminus of pRb but were still able to phosphorylate histone H1. Histone H1 appears to not possess an RXL motif and is unlikely to be recruited to cyclin in a hydrophobic patch-dependent manner (Schulman *et al.*, 1998). The hydrophobic patch mutations do not disrupt binding of cyclin A to CDK2 or CDK1 and therefore the inability of immunoprecipitates containing hpm cyclin A to phosphorylate RXL containing substrates suggests that the hydrophobic patch plays a role in substrate targeting to CDK.

Under the conditions used for these kinase assays, significantly higher concentrations of histone H1 (relative to p107, E2F1 and pRb) were required to achieve efficient phosphorylation. Approximately 25 times more histone H1 was required to achieve comparable levels of phosphorylation as those seen when p107, E2F1 and pRb were used, although the concentrations of substrates used were not normalised in terms of molarity or number of phosphoacceptor sites per substrate molecule, and rates were not measured, making direct comparisons between the different substrates difficult. Given the high number of SP and TP sites present in histone H1 however it is likely that on a stoichiometric basis, histone H1 is a significantly poorer substrate for cyclin A containing CDK complexes than are p107, E2F1 and pRb. The basis for the preference of RXL containing substrates could be attributed to the hydrophobic patch of cyclin A. It is possible that the hydrophobic patch of cyclin may serve to increase the local concentration of substrate at the bound CDK with the net effect of lowering the K_M of the substrates for the kinase. An alternative model could be that the hydrophobic patch of cyclin serves to correctly orient the phosphoacceptor site in the catalytic cleft of the CDK. If this model were correct then the addition of an endogenous, non-hydrophobic patch substrate binding sequence to hpm cyclin A would result in changes in the residues of substrates which were phosphorylated by bound CDK. This assumes however, that the non-hydrophobic patch substrate targeting motif added to hpm cyclin A would not also orient phosphoacceptor sites in the active site of CDK as does the RXL motif.

A number of viral oncoproteins bind to the pocket domain of p107 and a peptide derived from the papilloma virus E7 protein containing the sequence LYCYEQL is able to bind the pocket domain of p107 (Arroyo *et al.*, 1993). To test if the hydrophobic patch of cyclin A acted to properly orient the phosphoacceptor residues of substrates in the active site of bound CDK, the p107 interacting motif containing the consensus sequence LXCXE was introduced at two points in cyclin A which were both solvent exposed and dispensable for CDK binding. Immunoprecipitates containing LXCXE tagged hpm cyclin A or cyclin A were used to phosphorylate p107 in the presence of $\gamma[^{32}P]$-ATP *in vitro*. The tryptic digests of the phosphorylated p107 was analysed by two-dimensional SDS-PAGE. Most of the phosphopeptides generated when p107 was phosphorylated using immunoprecipitates containing cyclin A were also generated when immunoprecipitates containing LXCXE hpm cyclin A were used to phosphorylate p107.

The phosphopeptide map of p107 phosphorylated by wild type cyclin A shows five distinct peptides which are strongly phosphorylated. The map derived from one LXCXE-targeted cyclin A shows four well phosphorylated peptides, while the other insertion mutant shows three, with the fourth being phosphorylated to a significantly lesser extent. Schulman *et al* (1998) argue that this is a demonstration that the LXCXE motif is able to qualitatively substitute for the hydrophobic patch's role of substrate recruitment to cyclin A in a position-independent fashion. This supports a model where the hydrophobic patch of cyclin A interacts with substrates to merely raise the local concentration of substrate. Clearly however, LXCXE-directed phosphorylation of p107 by LXCXE hpm cyclin A containing complexes is different to wild type phosphorylation of p107 by cyclin A containing complexes. This suggests that the recruitment of substrate to CDK by cyclin is not only to increase the local concentration of substrate, but also to correctly position the phosphoacceptor(s) in the active site. Further to these conclusions, the observation that hpm cyclin A could direct the phosphorylation of histone H1 also suggests that there is an additional RXL and hydrophobic patch independent mechanism for recruiting substrates to CDK.

The hydrophobic patch is also conserved in the mitotic cyclins of *S. cerevisiae*. Clb5 interacts with p27 in yeast-2-hybrid assays and this can be disrupted by mutation of the hydrophobic patch, or by mutating Q241 (Q254 in cyclin A) which disrupts hydrogen bonding with the R and L of the RXL motif of p27 (Cross and Jacobson, 2000). These mutations only partly reduced the binding of the Sic1 inhibitor to Clb5, however, suggesting that there are additional contacts between Sic1 and Clb5. Mutation of the hydrophobic patch residues greatly reduced biological activity of Clb5, although did not affect its kinase activity measured towards histone H1. Moreover, similar mutations in Clb2 also reduced its ability to perform its mitotic functions, although they also appeared to increase the ability of Clb2 to perform functions normally restricted to Clb5. These data suggest that there is a conserved substrate targeting motif on the surface of Clb5 and Clb2 (Cross and Jacobson, 2000). Strangely however a peptide derived from E2F1 and containing the RXL motif was unable to inhibit the kinase activity of human cyclin B containing CDK complexes (directed towards a RXL containing substrate), and was also unable to block the binding of cyclin B to p21 (Adams *et al.*, 1996).

1.13. Can the hydrophobic patch of cyclin mediate substrate specificity?

This raises a difficult point about the role of the 'hydrophobic pocket', for the residues contributing to the hydrophobic patch of cyclin A are well conserved between cyclin B, cyclin E and cyclin D1. Figure 1-6 shows an alignment of the C-termini of B- and A-type cyclins from a number of different species. The conservation of residues between A- and B-type cyclins which contribute to the hydrophobic patch make it difficult to understand how this patch could play a role in generating substrate specificity. Moreover, although p21 interacts with cyclin B, it is a very poor inhibitor of CDK1:cyclin B compared to CDK complexes containing cyclin A and cyclin E. These observations suggest strongly that there is another motif or structural element in cyclin B which is responsible for its interaction with p21. This could also imply that there is another element responsible for substrate recruitment and/or selection in cyclin B. There is overlap in the expression pattern of cyclin A and cyclin B in human cells. Both cyclin A and cyclin B associate with and activate CDK1 during mitosis, and presumably cyclin A is complexed with CDK2 during G2 and mitosis also. There is little data however showing what fraction of cyclin A is complexed with CDK1 or CDK2 during G2 and mitosis, or addressing the question of whether there is redistribution of cyclin A from CDK2 to CDK1 prior to entry to mitosis.

If cyclin A and cyclin B are inherently different in their abilities to promote particular cell cycle events, and if this difference is not solely due to variation in localisation patterns, then given the conservation of the hydrophobic patch between cyclin A and cyclin B, it is probable that there are additional sequence elements which are required for dictating substrate specificity. These elements are likely to be solvent exposed, and if accountable for differences in substrate specificity, variable between cyclin A and cyclin B. The alignment of cyclin As and cyclin Bs in figure 1-6 is coloured to highlight residues which are conserved between cyclin A of different species (red) and between cyclin B (cyan) of different species. Residues which are conserved between cyclin A and cyclin B are shown in yellow, while residues which are buried at the CDK interface are marked with arrows. The residues contributing to the hydrophobic patch are not clustered in the primary sequence of cyclin A but are spatially proximal in the crystal structure of cyclin A. It is unlikely that analysis of the sequence of cyclins would expose motifs such as the hydrophobic patch of cyclin A. Mapping residues which are conserved between

Figure 1-6. Alignment of the C-termini of cyclins A and B from a variety of species.

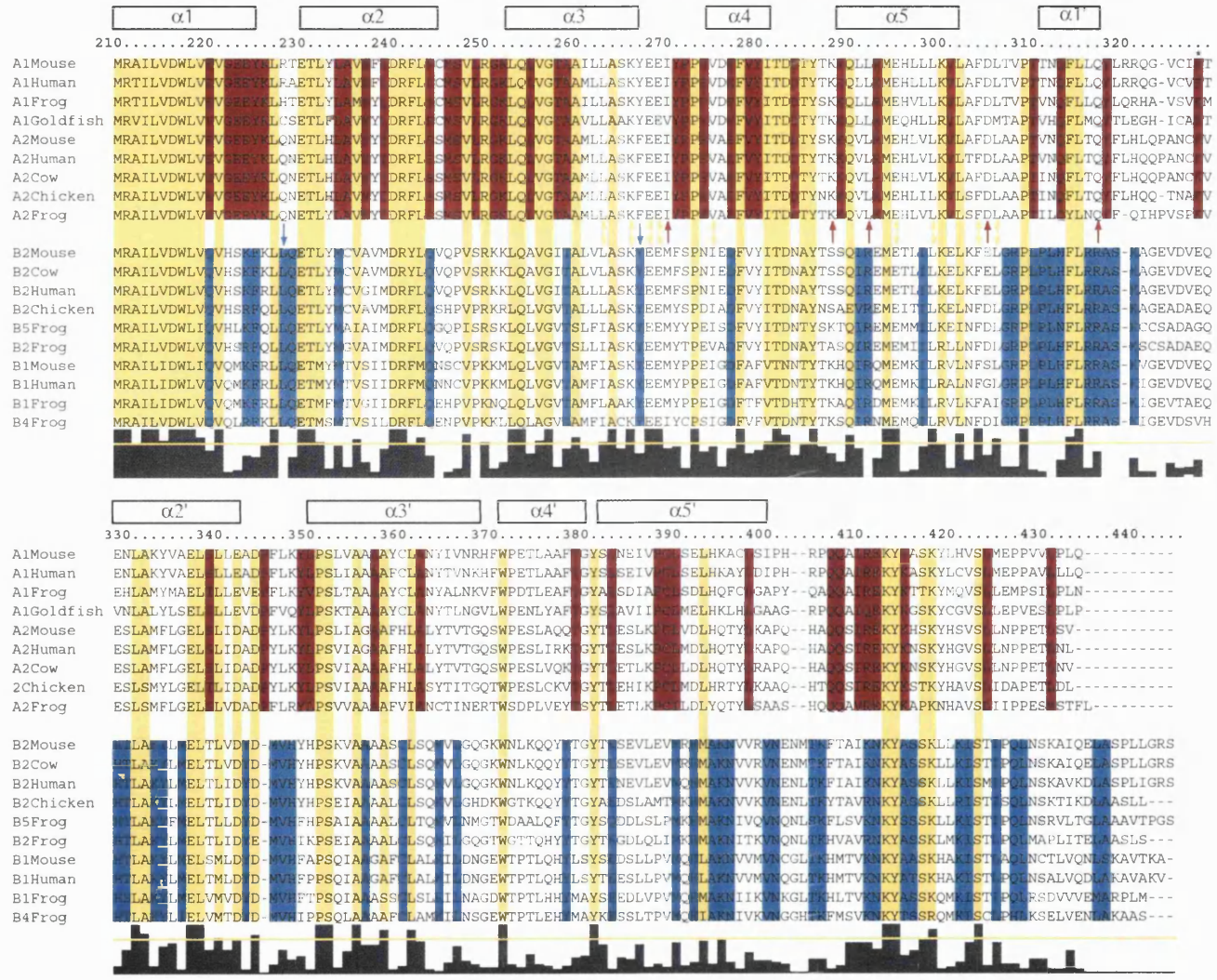


Figure 1-6. Sequence differences between cyclins A and B. Yellow:common between A and B, cyan: conserved in B, red:conserved in A, arrows:masked at the CDK interface.

different cyclins (such as cyclins A and B), and which are unique to specifically cyclin A or cyclin B for example, may reveal sequence elements responsible for mediating protein-protein interactions.

Figure 1-7 shows stereo images of the carbon backbone of cyclin A3 (ΔN170) which have had residues which are conserved between cyclins A and B (yellow), specific to cyclin A (red), specific to cyclin B or conserved within but not between cyclins A and B (magenta), spacefilled. The hydrophobic patch residues are coloured for orientation, (M210, orange; I213, purple; L214, green; W217, pink; L253, grey and Q254, brown) although they are highly conserved between all the cyclins aligned in figure 1-5. When residues were found to be conserved or specific to either cyclin A or B, but which were not solvent exposed due to masking by CDK interaction, they were not highlighted in the crystal structures (figure 1-7), and are marked by an appropriately coloured arrow in the alignments. Alignment of the C-termini of cyclins A and B shows that 60 of the residues are well conserved and it is unlikely that these residues would be involved in substrate discrimination between cyclins A and B. Fifty-two residues are well conserved in, and unique to cyclin A (red) and these are well distributed throughout the C-terminus of cyclin A. There are 63 residues which are well conserved and unique to the C-termini of B-type cyclins (cyan). Mapping the residues unique to B-type cyclins shows that the majority of them cluster in the lower part of the molecule when it is viewed so that the hydrophobic patch is in the top right quadrant of the molecule and bound CDK would bind to the left hand side of cyclin. Buried residues have not been excluded from these structures, although the backbone of the molecule shows that many of the residues are in fact solvent exposed.

The N-terminus of cyclin A does not form crystals, and the N-termini of cyclin B appears to be highly disordered (Hiro Yamano, personal communication) suggesting that the N-termini of these cyclins have little tertiary structure and so a similar analysis using the N-termini of cyclins A and B is not possible. Based on the data of Draviam *et al* (2001) it is likely that there would be significant differences between the N-termini of cyclins A, B1 and B2 however (Draviam *et al.*, 2001).

Figure 1-7. Conserved and unique residues in cyclins A and B.

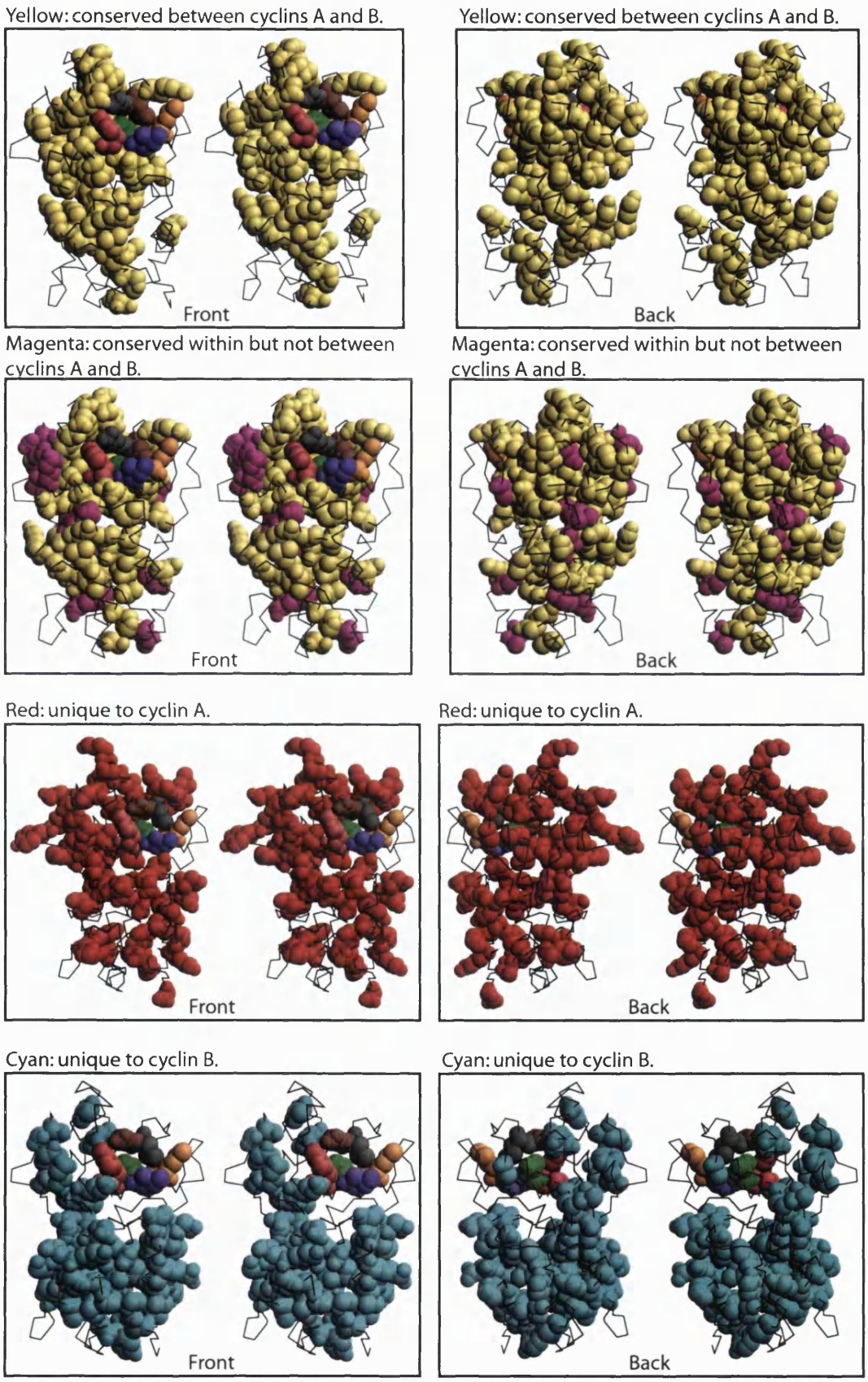


Figure 1-7. Residues shown in figure 1-5 were mapped onto the crystal structure of cyclin A3 using RasMac2.7. In all images the residues contributing to the hydrophobic patch of cyclin A are coloured thus: M210, orange; I213, purple; L214, green; W217, pink; L253, grey; Q254, brown. For details see text.

1.14. Screens and selections for identifying substrates of kinases.

Many systematic screening techniques currently in use, such as the yeast-2-hybrid and far-western screening, are based on interaction trap methodologies, which rely on the establishment of relatively stable interactions between two or more molecules. Unfortunately this type of interaction is not necessarily that which occurs between a protein kinase and its substrates. It became desirable therefore, to develop novel *systematic* screening methods based not on interaction traps, but instead utilising the enzymatic properties of the kinase of interest.

Fukunaga and Hunter (1997) described a solid-phase phosphorylation screen using the Extracellular-Regulated MAP kinase (ERK1). In their study they described the use of recombinant ERK1 to phosphorylate proteins expressed from a HeLa cDNA phage λ expression library in the presence of γ[^{32}P]-ATP. This approach successfully identified the previously characterised substrates of ERK1 ($p90^{RSK2}$ and c-myc) and a number of novel *in vitro* substrates, including MNK1 which they showed to be a genuine physiological substrate of ERK1 (Fukunaga and Hunter, 1997a). Solid phase phosphorylation screening does not rely on the establishment of stable protein-protein interactions as the output of the screen (detection of incorporation of [^{32}P] into target proteins) does not require that the interaction between kinase and substrate persist. This is a high-throughput screen and the isolation of cDNAs encoding proteins of interest is relatively simple. Unfortunately though, the expression of target proteins from a HeLa cDNA library is performed in *E. coli* and this has limitations including low levels of expression, insoluble or incorrectly folded proteins and degradation of expressed proteins. False positives may also be problematic as proteins expressed in *E. coli* and exposed to kinase *in vitro* have been removed from their normal physiological cellular context, where they may never encounter the kinase being used. This type of screen has been used successfully however to identify NPAT, a substrate of CDK2:cyclin E (Ma *et al.*, 2000) which has been characterised extensively as regulating the expression of histones which are required to package newly replicated DNA into nucleosomes during S-phase (Zhao *et al.*, 1998; Zhao *et al.*, 2000).

As early as 1995 biochemical techniques were used to identify proteins which specifically interacted with cyclins (Kellogg *et al.*, 1995a). Cyclins from yeast and *Xenopus laevis* were expressed as GST-fusions in *E. coli* and used to prepared affinity

chromatography columns. Affinity columns of GST-Clb2 or GST-Clb3 and GST-Xl cyclin A, GST-Xl cyclin B1 or GST-Xl cyclin B2 were loaded with proteins made from either yeast cells or *Xenopus* eggs respectively. After extensive washing bound proteins were separated using SDS-PAGE and identified using protein sequencing techniques (Kellogg *et al.*, 1995a). This approach lead to the identification of a 60 kDa protein which specifically interacted with B-type cyclins called NAP1. NAP1 in *Xenopus* extracts interacted with cyclins B1 and B2 but not cyclin A, while in yeast extracts, a NAP1 homologue interacted with Clb2 but not Clb3. Complexes of CDK1:cyclin B were able to phosphorylate NAP1 *in vitro* while complexes of CDK1:cyclin A were not; these data suggest that NAP1 is a substrate of CDK1:cyclin B. Moreover, these observations also suggest that an affinity chromatography approach could be used to assess the specificity of binding of interacting proteins to specific cyclins.

Given that screening for substrates needs to be done *in vitro*, it is beneficial to screen eukaryotic proteins using a eukaryotic expression system and any *in vitro* system which mimics the physiological state would be preferable. The advent of small-pool *in vitro* expression cloning (IVEC), used in conjunction with frog egg extracts that are able to recapitulate cell cycle events, allowed screens to be conducted which had the potential to identify substrates of mitotic kinases which were more likely to be physiologically relevant (Lustig *et al.*, 1997; Stukenberg *et al.*, 1997). *In vitro* transcription and translation was used to generate pools of [^{35}S]-labelled proteins which were incubated in the presence of either an interphase or mitotic extract prepared from frog eggs. Interphase extracts are able to form nuclei around naked DNA and replicate chromosomes once only, while the mitotic extracts are able to disassemble nuclear envelopes, condense chromosomes and form metaphase spindles. These extracts are expected to perform the physiological phosphorylations that are responsible for orchestrating the cell cycle oscillations during the rapid cell divisions of early embryogenesis. Following incubation of a labelled pool in the different extracts, substrates can be identified by differences in electrophoretic mobility through SDS-PAGE, or by generation of MPM2 reactivity. This approach was successfully used to identify a number of novel substrates of CDK1:cyclin B (Stukenberg *et al.*, 1997).

The recent completion of the first draft of the human genome, and the lack of success using traditional genetic approaches to identify kinase substrates has made

proteomic approaches an increasingly attractive possibility in identifying substrates of kinases. Shah *et al* (1997) used a knowledge of the structures of a number of kinases including cAMP-dependent kinase, CDK2 and v-Src to engineer the ATP-binding site of v-Src to accept an analogue of ATP (N^6-(cyclopentyl)) which was not recognised and could not be used by the wild-type enzyme or other kinases. It was shown that the catalytic efficiency of the engineered kinase using the ATP-analogue as substrate was similar to that of wild-type kinase using ATP substrate while the ATP-analogue could not be used by the wild-type enzyme (Liu *et al.*, 1998b). Because the substrate specificity of the engineered kinase was identical to the wild-type enzyme, the engineered v-Src could be used to identify its substrates even in the context of a complex mixture of cellular kinases. This type of approach is becoming increasingly successful and has been used to identify, amongst others, a novel substrate of JNK (heterogeneous nuclear ribonucleoprotein K) (Habelhah *et al.*, 2001). Extensive structural data is available for CDK2 and so it is an attractive, amenable target for this approach. If the active site of CDK2 could be engineered to utilise a $\gamma[^{32}P]$-labelled ATP analogue, then this kinase could be used to phosphorylate cell extracts which could be separated using 2D gels. Technical advances in protein sequencing and the completion of the sequencing of the human genome makes the identification of proteins phosphorylated exclusively by modified CDK2 significantly simpler. A similar approach has been used, albeit with limited success, by engineering the active site of Cdc28 (K. Shokat, personal communication).

1.15. The scope of this thesis.

This aim of this thesis was to develop a screen or selection which would permit the identification of substrates of CDKs, or binding partners for cyclins. By incorporating aspects of different types of screens it was hoped that greater success would be achieved. Initially the screen was to be based on affinity chromatography which had previously proved successful in identifying novel substrates of cyclins (Honey *et al.*, 2001; Kellogg *et al.*, 1995a). An affinity chromatography approach would also allow the identification of proteins which are not necessarily substrates of cyclins, but which might be involved in regulating cyclin or CDK activity, or could act as adapter or targeting molecules for CDK:cyclin complexes. Screening cell extracts for proteins which interact with cyclins or CDKs can cause difficulties in the isolation and identification of interacting proteins.

Screening pools of *in vitro* transcribed and translated proteins allows the easy identification of proteins of interest and circumvents the need for proteomics. Using a eukaryotic *in vitro* transcription and translation system to generate pools of proteins means that proteins are more likely to be folded correctly than they would if they were expressed in an *E. coli* based expression system (such as that used in Solid-phase phosphorylation screening). The ligand to be used; a cyclin A construct missing its N-terminal 170 residues called cyclin A3 can be purified in active form in large amounts from *E. coli* and its structure has been determined both in monomer form as well as in complex with CDK2 and CDK2:p27. Kinase activity can be generated by co-expression of GST-CDK2 and Civ1 in bacteria followed by mixing lysates from these cells and from cells expressing cyclin A3. The complexes can be purified to virtual homogeneity in a two-step purification scheme.

IMAC25/Tx100	20 mM Tris-Cl pH 8.0, 1 M NaCl, 25 mM imidazole, 0.5 % (v/v) Triton X-100, 0.5 % (v/v) Tween-20
pGEX lysis buffer	50 mM Tris-Cl pH 7.5, 2 mM EDTA, 1 mM DTT, 0.25 mM PMSF, 2 mg/ml lysozyme
2xTY	1 % (w/v) bactotryptone, 1 % (w/v) yeast extract, 0.5 % (w/v) NaCl
LB	1 % (w/v) bactotryptone, 0.5 % (w/v) yeast extract, 1 % (w/v) NaCl

2.6. Non-directional cDNA library construction in λYES.

2.6.1. Total RNA preparation from *Xenopus laevis* oocytes.

Stage six *Xenopus laevis* oocytes were frozen in liquid nitrogen and ground to a fine powder using a pestle and mortar. Pulverised material was stored at -80°C. An equal volume of extraction buffer containing 0.2 M LiCl, 50 mM Tris-Cl pH 8, 10 mM EDTA pH 7.5 and 0.5% SDS was mixed with Tris-Cl pH 8 equilibrated phenol containing 0.1% 8-hydroxyquinoline and heated to 90°C. A single phase was formed by mixing and to a 20 ml volume of extraction buffer/phenol emulsion one gram of powdered oocytes was added. The resulting emulsion was mixed vigorously and spun in a bench-top swing-out Sorvall centrifuge at 2800 g for 10 minutes at room temperature. The colourless, aqueous phase was re-extracted with an equal volume of phenol and re-spun. Contaminating phenol in the aqueous phase was removed by three ether extractions. LiCl was added to 2 M and incubated on ice for one hour. Precipitated RNA was collected by centrifugation at 10000 rpm for 10 minutes at 4°C in siliconised, baked Corex tubes. The RNA pellet was washed with 70% EtOH and re-suspended in 100 µl of 10 mM Tris-Cl pH 8, 1 mM EDTA pH 8 per gram of ground oocytes. Typical yields were approximately 2.5 mg RNA per gram of ground oocytes. Alternatively, RNA was extracted using Trizol™ Reagent (Life Technologies, Scotland) according to the manufacturers instructions.

2.6.2. Poly A^+ RNA selection.

One gram of oligo dT cellulose (Collaborative Research Inc.) was washed once with 25 volumes of loading buffer containing 500 mM NaCl, 10 mM Tris-Cl pH 7.5 and 1 mM EDTA, once with 25 volumes of storage buffer containing 100 mM NaOH and 5 mM EDTA, then with water until the pH returned to 6 and finally was equilibrated three times in loading buffer. RNA at a concentration of 2 mg/ml was heated to 65°C for five minutes and added to 0.5 volumes of equilibrated, packed oligo dT cellulose. An equal volume of 2x loading buffer was added to make the slurry 500 mM for NaCl. The resulting slurry was cooled to room temperature and incubated for 10 minutes before being poured into a siliconised, baked Pasteur pipette plugged with glass wool. The flow through was heated to 65°C for five minutes, cooled rapidly on ice and loaded onto the column. This was

repeated so that the RNA passed over the column a total of three times. The column was washed with five column volumes of loading buffer and five columns of loading buffer containing 250 mM NaCl. The column was eluted with 10 successive column volumes of elution buffer which had been heated to 65°C. The A_{260} of each fraction was determined and RNA containing fractions were pooled, heated to 65°C for five minutes and added to an equal volume of 2x loading buffer. This material was passed through the column three times as before. Poly A^+ RNA was eluted and detected as before. To the pooled RNA containing fractions 0.1 volumes of 4 M NH_4OAc pH 5.2 and 2.5 volumes of EtOH were added. This was chilled in a baked, siliconised Corex tube at -80°C until solid. RNA was collected by centrifugation at 11 951g for 30 minutes at 4°C. The RNA pellet was washed with 70% EtOH, air dried and resuspended in 10 µl of 10 mM Tris-Cl pH 8, 1 mM EDTA pH 8 per gram of starting material. Purity was determined by the ratio of A_{260}:A_{280} and translation in rabbit reticulocyte lysate. Typical yields were 75 µg poly A^+ RNA per gram of ground oocytes.

2.6.3. cDNA synthesis from poly A^+ RNA and adapter ligation.

First and second strand cDNA synthesis were performed using a cDNA synthesis module from Amersham (Catalogue number RPN1256, Amersham International plc, England) in accordance with the manufacturers instructions. Briefly, 5 µg poly A^+ RNA were reverse transcribed in a volume of 100 µl using 100 U of AMV reverse transcriptase from an anchored oligonucleotide dT_{25} primer. Second strand synthesis was performed in a volume of 500 µl using *E. coli* DNA polymerase I and primed using the *E. coli* RNAse H nicked RNA template from the first strand synthesis. Nicks were filled using phage T4 DNA polymerase. A small amount of $\alpha[^{32}P]$dCTP was included in each synthesis reaction to allow determination of yields and efficiencies of each reaction. First and second strand synthesis was monitored by scintillation counting and analysis by denaturing and native agarose gel electrophoresis for first and second strand reactions respectively. The following adapters were formed by annealing 5' phosphorylated oligonucleotides at a concentration of 100 mM each to 65°C for five minutes and slowly cooling to room temperature in a buffer containing 10 mM Tris-Cl pH 7.5, 100 mM NaCl, and 1 mM EDTA.

C G A G G C G G C C G C G T C G A C
C G C C G G C G C A G C T Gp5'

Double stranded cDNA was ligated to adapters in a 15 µl reaction containing 0.13 nmol of the adapters, 50 mM Tris-Cl pH 7.8, 200 µg/ml BSA, 10 mM $MgCl_2$, 1 mM spermidine, 0.33 mM ATP, 1 mM hexammine cobalt chloride and 88 U/µl T4 DNA ligase for 16 hours at 4°C. Precipitation in the presence of 5 mM spermine and 100 mM KCl in a final volume of 100 µl largely removed adapters. The cDNAs were analysed by 1% agarose gel electrophoresis and those longer than 0.5 kbp were excised from the gel and purified using a Qiagen gel extraction kit (Catalogue number 28704, Qiagen, Germany), according to the manufacturers instructions. The purified cDNA was EtOH precipitated, washed and resuspended in a small volume of TE buffer.

2.6.4. λ YES vector preparation.

λYES (a gift of Dr. Steve Elledge, Baylor College of Medicine, Houston, USA) was amplified as a plasmid by transformation into *E. coli* strain DH5α. Extra-chromosomal lysogens were selected by growth at 30°C on LBA agar. Growth must be at or below this temperature as lambda repressor is temperature-sensitive and is required to maintain lysogeny. A single colony was inoculated into 10 ml 2x TYA media and following growth overnight at 30°C was used to inoculate 400 ml 2x TY media. The culture was grown at 30°C until the culture reached saturation, usually about 24 hours. Lambda DNA was prepared according to standard alkaline lysis methods and the final pellet was dissolved in 4 ml of TE buffer to which 4.4 g CsCl and 400 µl of 10 mg/ml EtBr were added. This was spun at 10 000 g for 10 minutes to pellet insoluble material, and the supernatant was loaded into a polypropylene quickseal tube which was spun in a vTi65.2 rotor at 50000 rpm for 16 hours at 22°C. Supercoiled lambda DNA was removed using a 1 ml syringe and 20 gauge needle under visible light. EtBr was removed by extraction against 20 volumes of water saturated n-butanol. CsCl was removed by spin dialysis using a Centricon-30 microconcentrator. Following extraction once with an equal volume of Tris-Cl pH 8 equilibrated phenol and three times with equal volumes of ether, lambda DNA was precipitated with 2.5 volumes of EtOH. Typical yields were about 300 µg lambda DNA per 400 ml of culture. 40 µg of vector was cut with 500 U of Xho I for 10 hours at 37°C. DNA was precipitated by the addition of EDTA to 20 mM, 0.1 volumes 4 M NH_4OAc and

2.5 volumes EtOH. Precipitated DNA was washed with 70% EtOH and resuspended in 190 µl TE buffer, and re-precipitated by the addition of 10 µl 100 mM spermine. DNA was washed with spermine wash buffer containing 70% EtOH, 10 mM $Mg(OAc)_2$ and 300 mM NaOAc pH 7.2, followed by 70% EtOH. DNA was resuspended in 90 µl TE. Yields were typically 20% of the starting material. The Xho I site was filled with a single thymine residue to make the sticky-ends incompatible. 10 µl of 10x Taq buffer containing 500 mM KCl, 100 mM Tris-Cl pH 8.3, 20 mM $MgCl_2$ and 0.1% gelatin was added to the lambda DNA along with 1 µl of 2.5 mM dTTP. This was heated to 72°C for five minutes and 1 U Taq polymerase or Sequenase™ was added. The reaction proceeded for two minutes at 72°C and was then stopped by precipitation with spermine as before. The T-filled lambda was resuspended in 20 µl TE.

2.6.5. Ligation of adapted cDNA to T-filled λYES.

2 µl of ligation buffer containing 50 mM Tris-Cl pH 7.8, 10 mM $MgCl_2$, 200 µg/ml BSA, 1 mM spermidine, 1 mM ATP pH 7.8, 7 mM DTT and 120 U T4 DNA ligase was mixed with 1 µl of 100 ng/µl adapted cDNA and 1 µl of 2 µg/µl T-filled λYES. The ligation reaction proceeded for 16 hours at 4°C.

2.6.6. Preparation of packaging extracts.

Preparation of packaging extracts was based on the method of Gunther *et al* (1993) (Gunther *et al.*, 1993). Temperature sensitivity of strains BHB 2688 and NM 759 (a gift of Dr. Peter Glazer, Boyer Center for Molecular Medicine, New Haven, CT, USA) was checked by streaking on LB agar plates and incubation at 30°C and 42°C. 80 ml of NZY media was inoculated with a single colony of BHB 2688 and grown overnight at 30°C. Three flasks of 500 ml NZY were inoculated with 25 ml of the overnight culture of BHB 2688 and grown with shaking at 32°C until the OD_{600} reached 0.6. The cultures were transferred to a 65°C shaking waterbath and the internal temperature was monitored until it reached 45°C when they were transferred to a 45°C waterbath for 15 minutes. The cultures were then placed in a shaking incubator at 38°C for 3 hours. Cells were harvested by centrifugation at 3 422 g for 20 minutes at 4°C. Pellets were resuspended in a total of 3 ml sucrose solution containing 10% sucrose w/v and 50 mM Tris-Cl pH 8. 500 µl aliquots were pipetted into tubes containing 25 µl lysozyme solution containing 2mg/ml lysozyme and 10 mM Tris-Cl pH 8 and dropped into liquid nitrogen. The contents of the tubes were

thawed on ice and combined into two tubes which were spun at 45000 rpm in a Ti 90 rotor at 4°C. The supernatant was stored in 80 µl aliquots at -80°C and was called "freeze-thaw" extract. 30 ml of NZY was inoculated with a single colony of NM 759 and grown overnight at 30°C. 500 ml of NZY media was inoculated with 25 ml of the overnight culture of NM 759 and grown until the OD_{600} reached 0.3. The culture was treated as before but the cell pellets were resuspended in 3.6 ml of sonication buffer containing 20 mM Tris-Cl pH 8, 10 mM EDTA pH 8 and 5.8 mM β-mercaptoethanol. Cells were sonicated 10 times in two second bursts and then spun at 10 000 g in an SS-34 rotor at 4°C for 10 minutes. To the supernatants were added 0.5 volumes sonication buffer and 0.17 volumes packaging buffer containing 6 mM Tris-Cl pH 8, 12.8 mg/ml spermidine, 8 mg/ml putrescine, 10 mM $MgCl_2$, 5 mM β-mercaptoethanol and 30 mM ATP pH 7. 120 µl aliquots were frozen in lN_2 and stored at -80°C and was called "sonication" extract.

2.6.7. Packaging of ligated cDNA/λYES into infective phage particles.

80 µl of "freeze-thaw" extract from strain BHB 2688 was mixed with 120 µl of "sonication" extract from strain NM 759 to which was added 1 µl of concatenated λYES-cDNA ligation reaction. Control reactions of self-ligated and unligated λYES were also performed to determine the background of non-recombinant phage. Packaging reactions proceeded for 2 hours at room temperature and were terminated by the addition of 300 µl SM buffer containing 130 mM NaCl, 8 mM $MgSO_4$, 50 mM Tris-Cl pH 7.5 and 0.01% w/v gelatin. To each reaction a drop of chloroform was added, mixed gently and spun briefly to sediment the chloroform. Alternatively, concatenated λYES-cDNA was packaged using either the Packagene® Lambda DNA Packaging System (Catalogue number K3152, Promega Corporation, USA) or Gigapack® III Gold Packaging Extracts (Catalogue number 200202, Stratagene Cloning Systems, USA) according to the manufacturers instructions. To determine the primary titre of infective phage, serial dilutions of the packaging reactions were made and infected into *E. coli* strain LE392 suspended in 10 mM $MgSO_4$. Infected cells were plated in top agar onto LB plates and incubated at 37°C until a lawn of bacteria appeared.

2.6.8. Determining percentage recombinant and mean insert size

Plaques were picked at random into PCR buffer and amplified for 30 cycles using primers flanking the Xho I site of λYES. Phage isolated from these plaques were also infected into

$\alpha[^{32}P]$dCTP was included to monitor reaction efficiency. The following adapters containing a Sal I half-site were ligated to the double stranded cDNA in a reaction containing 5 U T4 DNA ligase and 200 ng/µl adapters in a volume of 50 µl for 16 hours at 16°C.

5' pT C G A C C C A C G C G T C C G
G G G T G C G C A G G Cp 5'

Adapted cDNA was precipitated with 0.5 volumes of 7.5 M NH_4OAc and 3 volumes of EtOH. cDNA was then digested with 60 U of Not I for two hours at 37°C. Sal I adapted, Not I digested double stranded cDNA was fractionated through Sephacryl S-1000 and fractions were analysed by agarose gel electrophoresis and scintillation counting. Fractions containing cDNA greater in length than 0.5 kbp were quantified and used in ligation reactions.

2.7.3. pTex-1 vector preparation.

5 µg pTex-1 (a gift of Dr. Stephan Geley, ICRF Clare Hall, South Mimms, UK) was digested with 20 U Sal I and 10 U Not I and purified on a 1% agarose gel. Linearised plasmid was purified using a Qiagen Gel Purification kit according the manufacturers instructions. This material was re-digested with Sal I and Not I and re-purified. Plasmid DNA was self-ligated using a Rapid ligation kit (Boehringer Mannheim GmbH, Germany) according to the manufacturers instructions and transformed into chemically competent *E. coli* strain DH5α to confirm completion of digestion.

2.7.4. Ligation of cDNA to pTex-1, and introduction into *E. coli*.

30 ng of double stranded, adapted cDNA was ligated to 50 ng of Sal I/Not I digested pTex-1 in a reaction volume of 20 µl containing 50 mM Tris-Cl pH 7.6, 10 mM $MgCl_2$,1 mM ATP, 5% w/v PEG 8000, 1 mM DTT, 1 U T4 DNA ligase for 16 hours at 4°C. Ligation products were precipitated in the presence of 5 µg yeast tRNA by addition of 0.5 volumes 7.5 M NH_4OAc and 2.5 volumes of ice cold EtOH. Precipitated material was resuspended in 5 µl ddH_2O and 1 µl was added to 25 µl of electrocompetent DH10B *E. coli* (Gibco BRL, Scotland). Following incubation at 4°C for five minutes cells were electroporated in 0.2 cm cuvettes using a Gene Pulser Electroporation System (BioRad, England). The Gene Pulser was set to 25 µF and 2500 V, with the Pulse Controller Unit set at 4000 Ω,

generating 12500 V/cm and a pulse length of 8-10 ms. Cells were recovered for one hour at 37°C with vigorous shaking in 1 ml SOC media. Transformation efficiency was determined by plating serial dilutions of the cells on LBA agar plates. Primary libraries were stored at -80°C in 10% glycerol until required.

2.7.5. Determining percentage recombinant and mean insert size.

Random colonies were picked into PCR reactions and inserts were amplified using primers flanking the Sal I and Not I sites in pTex-1. Colonies were also inoculated into LBA media and grown overnight. DNA prepared from these cultures using Qiagen miniprep kits (Qiagen, Germany) according to the manufacturers instructions was digested with Sal I and Not I, or translated *in vitro* in the presence of [^{35}S]-methionine. PCR and restriction digestion products were analysed on 1% agarose gels and translated proteins were analysed by SDS-PAGE and autoradiography.

2.8. Creating a GST-fusion library.

The coding sequence for GST lacking a stop codon was cloned into pTEX1, placing it 3' to the T7 promoter, the human influenza virus leader sequence and a single c-myc epitope, A linker was introduced containing a *lox* site 3' to the GST sequence. This construct called pHOST *iv* myc GST was sent to Dr. Steve Elledge (Baylor College of Medicine, Houston, TX, USA) who fused a HeLa cDNA to pHOST *iv* myc GST so that the cDNAs could be expressed as fusions to GST using *in vitro* transcription and translation systems.

2.9. PCR techniques.

2.9.1. Amplification of DNA by PCR.

PCR reactions were performed in a volume of 100 µl using either *Taq*, *Pfu* or *Pfu* turbo polymerase. The following table summarises the reaction conditions used for each of these polymerases:

Reagent	*Taq*	*Pfu*/*Pfu* turbo
Template (ng)	5	100
[Primer] (pmol)	20	30
[dATP, dCTP, dGTP, dTTP] (μM)	200 each	200 each
Enzyme (U)	2	5
[Tris-Cl pH 8.8] (mM)	10	20
[$MgSO_4$] (mM)	-	2
[$MgCl_2$] (mM)	1.5	-
[KCl] (mM)	50	10
[$(NH_4)_2SO_4$)] (mM)	-	10
[Triton X-100] (% v/v)	-	1
[Gelatin] (% w/v)	0.001	-
[BSA] (μg/ml)	-	100
Polymerisation time (min/kb)	1	2/1

Unless otherwise specified 30 cycles were performed with an annealing temperature of T_m for the specific oligonucleotide minus 4°C for 15 seconds, polymerisation at 72°C and denaturation at 94°C for 15 seconds. Reactions were performed using a Peltier Thermal Cycler 2000 (MJ research, USA). When the melting temperature of an oligonucleotide in a pair differed significantly, the lower one was used. Enzyme was always added to the reactions after they had been incubate at 94°C briefly.

2.9.2. Screening bacterial colonies for the presence of plasmid using PCR.

Following transformation, *E. coli* were screened for the presence of recombinant plasmid constructs using PCR. A small part of single colonies from a plate of transformants were picked into PCR reactions. Prior to the addition of polymerase the reactions were heated in the thermal cycle for 30 seconds to lyse the bacteria. The remainder of the colony was spotted onto a duplicate LB plate with the appropriate antibiotic selection.

2.9.3. Using PCR to identify phage containing cDNA inserts.

Packaged phage were infected into *E. coli* strain LE 392. Bacteria were plated in top agar onto LB plates and incubated at 37°C overnight to allow the formation of plaques. Plaques were allowed to develop until they were approximately 3 mm in diameter before being

picked from the plate with the narrow end of a Pasteur pipette directly into a PCR reaction containing oligonucleotides flanking the MCS of the vector. Thirty cycles of amplification were performed and PCR products were analysed by agarose gel electrophoresis and visualised using EtBr.

2.9.4. Site-directed mutagenesis using PCR.

Site-directed mutagenesis was performed based on the protocol of Horton *et al.* (1991). Two complementary, overlapping mutagenic oligonucleotides were synthesised and used in two separate PCR reactions in combination with either a 5' or 3' flanking non-mutagenic oligonucleotide to generate 5' and 3' PCR products with a short common region where the mutagenic oligonucleotides were complimentary. The 5' and 3' PCR products were mixed and diluted 500 fold in PCR buffer and used in a PCR reaction containing the 5' and 3' flanking, non-mutagenic oligonucleotides to amplify the whole sequence of interest.

2.9.5. RT-PCR.

Single stranded cDNA was synthesised from total RNA using SuperScript II™ reverse transcriptase (Life Technologies). 1 μg total RNA was treated with 10 U DNase I in a buffer containing 50 mM Tris-Cl pH8.3, 75 mM KCl, 3 mM $MgCl_2$ in a volume of 25 μl for 15 minutes at 37°C. DNase I was inactivated by the addition of 2.5 μl of 25 mM EDTA and incubation at 65°C for 10 minutes. The DNase digested RNA was immediately made to 50 μl in a buffer containing 250 ng oligo dT oligonucleotide, 50 mM Tris-Cl pH8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT, 0.2 mM dNTPs, 20 U Superscript II™ reverse transcriptase and incubated at 42°C for 2 hours. Reverse transcriptase was inactivated by incubation at 95°C for 5 minutes. 5 μl of single stranded cDNA was used as template in 100 μl PCR reactions.

2.10. Preparation and analysis of DNA.

Different methods of DNA preparation have been used to meet the requirements of the particular application for which the DNA was required. The techniques used are summarised in the table below.

clearing spin at 3000 rpm for 5 minutes was performed and the supernatant was loaded into a Beckman Quickseal ultracentrifuge tube and spun overnight in a Vti 65.2 rotor 200 000 g. Plasmid DNA visible as a dark red band under visible light was collected by sidepuncture and extracted five times with 10 volumes of water saturated butanol to remove ethidium bromide. The DNA was diluted in an equal volume of TE and precipitated by the addition of 2.5 volumes of room temperature absolute ethanol.

2.10.3. Qiagen minipreps.

As a convenient alternative to large scale DNA purification, Qiagen miniprep kits were often used in accordance with the manufacturers recommendations.

2.11. Extraction of DNA from agarose gels using Qiagen columns.

As an alternative to purifying DNA from agarose gels using hot phenol the Qiagen Gel Extraction Kit was used according the directions of the manufacturer.

2.12. Restriction enzyme analysis.

Restriction enzymes were purchased from New England Biolabs (NEB) and digests were performed using their supplied buffers. Typically, 1-2 U of enzyme was used per μg of DNA. For double digests the suggested NEB buffer was used, and when suitable conditions could not be provided by the addition of a single buffer, the reactions were performed sequentially.

2.13. Vector preparations for cloning.

Appropriately digested plasmid DNA was routinely dephosphorylated to prevent self ligation. The 5' phosphate was removed by the addition of 0.5 U calf intestinal alkaline phosphatase per μg of DNA (CIAP, Boehringer BCL, molecular biology grade) to restriction digests for the final 30 minutes of the reaction. Digested dephosphorylated plasmid DNA was purified from agarose gels using a Qiagen Gel Extraction kit.

2.14. Ligation of DNA fragments.

For ligations when single stranded overhangs were present (sticky-ends), 50-100 ng dephosphorylated vector and a 4-5 fold molar excess of DNA fragments were used. For

Protein purification techniques.

2.16. Over-expression and purification of hexahistidine tagged cyclin A3.

Initial purifications used bovine cyclin A3, while later ones used human A3. Cyclin A3 is a deletion of the N-terminal 170 amino acids of cyclin A2, which allows efficient over-expression in and purification from *E. coli*. Plasmid pET21d (Novagen, USA) containing an ORF coding for cyclin A3 followed by six histidines was transformed into *E. coli* strain BL21(DE3)pLysS. A culture was grown at 37°C with vigorous shaking in 400 ml LBA media supplemented with 33 μg/ml chloramphenicol to an OD_{600} of 0.6, when expression was induced by the addition of 0.1 mM IPTG. Growth continued at 37°C for 30 minutes and then at 30°C for a further 3 hours. Harvested cells were resuspended in 20 ml buffer N containing 50 mM Tris-Cl pH 7.5, 300 mM NaCl, 0.01% monothioglycerol, 0.01% NaN_3, 0.25 mM PMSF, 0.7 μg/ml pepstatin, 0.5 μg/ml leupeptin, 1 mM benzamidine, 20 μg/ml DNAseI, 20 μg/ml RNAseI, 0.1% Triton X-100, 5mM imidazole pH 8 and 10 mM $MgCl_2$. Following lysis by freeze-thaw, lysates were incubated on ice for 30 minutes then centrifuged at 10000 g at 4°C for 30 minutes in an SS-34 rotor. Cleared lysate was passed through a 0.2 μm filter syringe and loaded onto a 0.5 ml Ni^{2+}-NTA agarose column (Qiagen, Germany). The column was washed with 25 column volumes of N buffer containing 10 mM imidazole pH 8 and 25 column volumes of buffer N containing 50 mM imidazole pH 8. Bound material was eluted with successive column volumes of buffer N containing 150 mM imidazole pH 8.

2.16.1. Size exclusion chromatography of cyclin A3.

Fractions containing cyclin A3 were pooled and made to 0.1 M $MgCl_2$. Pooled cyclin A3 was concentrated to 10 mg/ml and loaded onto a AcA-34 gel filtration column equilibrated in 50 mM Tris-Cl pH8.5, 0.1 M $MgCl_2$, 0.01% monothioglycerol and 0.01% NaN_3. Two peaks of cyclin A3 elute, one corresponding to the molecular weight of monomeric cyclin A3, the other is presumably aggregated cyclin A3. Alternatively, 5 mg of cyclin A3 from the Ni^{2+}-NTA column was loaded onto a Sephadex-200 size exclusion column and resolved using an ÄKTA FPLC according to the manufacturers instructions.

In vitro small-pool expression cloning techniques

2.18. Preparing plasmid pools for *in vitro* transcription and translation.

Small aliquots of primary library were plated onto LB agar plates containing 50 μg/ml carbenicillin at a colony forming density calculated to be 100 per plate. Density varied from 60 to 120 colony forming units per plate. Plates were incubated at 37°C until colonies were on average 1 mm in diameter. Ten ml of LB media was added to each plate and they were rotated gently at room temperature until all colonies were resuspended. Cells were transferred to 14 ml snap-top tubes and incubated with vigorous shaking for one hour at 37°C. Cells were then pelleted by centrifugation at room temperature in a bench top Sorvall centrifuge at 3 000 g. Spent media was discarded and plasmid DNA was extracted from the pelleted cells using an automated Qiagen Biorobot 9600 workstation at the ICRF Equipment park. Yields were typically 50-100 ng/μl.

2.19. *In vitro* coupled transcription and translation.

Pools of plasmid library were transcribed and translated using a TnT® Quick Coupled Transcription/Translation System (Catalogue number L2081, Promega Corp., USA). 10 μl reactions containing 1 μl of plasmid pool at approximately 100ng/μl and 9 μl TnT® Quick Master Mix supplemented with 0.7 μCi/μl [^{35}S]-methionine were incubated at 30°C for 90 minutes. 1 μl of the reaction products were separated by 12.5% SDS-PAGE and visualised by autoradiography. Individual clones were transcribed and translated using rabbit reticulocyte lysate purchased from Ambion Inc, USA and treated with 500 U of micrococcal nuclease and supplemented with 20 μM hemin, 50 μg/ml creatine kinase, 1 mM $CaCl_2$ and 60 μg/ml calf liver tRNA. 10 μl reactions typically contained 10-50 ng plasmid template, 0.65 μCi/μl [^{35}S]-methionine, 0.74 mM each deoxy-rNTP, 3 mM $MgCl_2$, 37 mM KCl and 0.05 μl T7 RNA polymerase. Reactions proceeded for two hours at 37°C and 0.5-1 μl of the translated products were separated by SDS-PAGE and visualised by autoradiography. The incorporation of biotinyl-lysine into translated proteins was accomplished by the omission of [^{35}S]-methionine from the reactions, and their supplementation with biotinyl-lysyl tRNA and unlabelled methionine. Reactions were incubated as before.

2.20. Small-pool expression cloning.

Rounds of screening typically involved *in vitro* transcription/translation of 48 or 96 pools of plasmid library. 1 μl of plasmid pool was used per coupled transcription/translation and reactions proceeded for 90 minutes at 30°C.

2.20.1. Screening by affinity chromatography.

8 μl from each *in vitro* transcription-translation reaction was diluted with 192 μl of buffer containing 50 mM Tris-Cl pH 7.5, 150 mM NaCl, 0.1% NP-40 and 5 μg of GST-CDK2:cyclin A3. Following an incubation of one hour on ice the GST-CDK2:cyclin A3 and any bound [^{35}S]-labelled translation products were harvested using GSH-Sepharose using a microcolumn format. The diluted translation reaction containing GST-CDK2:cyclin A3 was passed through a siliconised P200 tip plugged with glass wool containing 10 μl of GSH-Sepharose. The tip is supported in a P200 tip box which is used to catch the flow through and subsequent washes. The columns were washed three times and the GSH-Sepharose was harvested by cutting one cm from the tapered end of the tips and inverting each one in a 0.5 ml Eppendorf tube and a brief spin in a microcentrifuge. The proteins captured by the GSH-Sepharose were eluted by heating in SDS-sample buffer. The captured proteins were separated by SDS-PAGE and visualised using autoradiography. In the example shown in figure 5-2, there were four proteins in the pool interacting with GST-CDK2:cyclin A3 or GSH-Sepharose.

2.20.2. Screening using the kinase activity of CDK2:cyclin A3.

1 μl of translation products were retained as input into the kinase reactions to act as negative controls. 1 μl of translation reaction was incubated in the presence of 227 nM GST-CDK2:cyclin A3 in a reaction volume of 5 μl containing 50 mM β-glycerophosphate pH 8, 5 mM NaF, 1 mM DTT and 15 mM $MgCl_2$ for 15 minutes at 30°C. Reactions were terminated by the addition of an equal volume of 2x SDS-sample buffer and heating to 100°C for 2 minutes. The electrophoretic mobility of [^{35}S]-labelled proteins were compared with those not incubated with GST-CDK2:cyclin A3.

2.21. Isolating single clones encoding proteins of interest.

cDNAs encoding proteins whose mobility changed in the presence of CDK2 were isolated by transforming the pool of interest into *E. coli* strain DH5α and plating at a colony forming density of approximately 200-500 per 10 cm plate. Individual colonies were picked into LBA agar filled wells of a 96-well microtitre dish. Colonies from wells in columns were then picked into 12 individual cultures (1-12), and colonies from rows were picked into 8 different cultures (A-H). Thus each of the colonies in each of the wells is represented in two cultures; one from a row and one from a column (1A-12H). DNA was extracted from the cultures and re-screened for the clone(s) of interest. Such a clone should be present in both a reaction prepared from a "column" culture and from a "row" culture. Cross-referencing the columns and rows allows the identification of the well containing the cDNA of interest.

2.22. Screening *in vitro* transcribed and translated biotinylated proteins.

Translation reactions containing individual plasmids containing cDNAs of interest and biotinyl lysyl tRNA were diluted in 200 µl buffer containing 50 mM Tris-Cl pH 7.5, 150 mM NaCl, 0.1% NP-40. To the diluted reactions was added 15 µl of monomeric streptavidin agarose (Promega Corporation). Following a brief incubation on ice, the streptavidin agarose was pelleted by gentle centrifugation and washed extensively. The resin was resuspended in 15 µl of histone H1 kinase buffer containing 100 ng of GST-CDK2:cyclin A3 and 0.1 µCi $\gamma[^{32}P]$-ATP. Reactions were incubated for 30 minutes at 37°C. Bound proteins were eluted by boiling in SDS-sample buffer and separated by SDS-PAGE. Phosphorylated proteins were detected using autoradiography.

Separation and detection of proteins

2.23. SDS-polyacrylamide gel electrophoresis (SDS-PAGE).

The method of SDS-PAGE was based on Andersen *et al.* (Andersen *et al*, 1973) with the main difference being that neither the stacking nor resolving gel mixes contained SDS. All polyacrylamide gels described here were 15 % resolving gels. 100 ml of the resolving gel mix contained 50 ml 30 % (w/v) acrylamide, 8.6 ml 1 % (w/v) bisacrylamide, 25 ml 1.5 M Tris-Cl pH 8.8 and 16.4 ml H_2O. 150 ml of the stacking gel mix contained 25 ml 30 % (w/v) acrylamide, 20 ml 1 % (w/v) bisacrylamide, 18.8 ml 1 M Tris-Cl pH 6.8 and 86.2 ml H_2O. Stocks of 30 % (w/v) acrylamide and 1 % (w/v) bisacrylamide were deionized with MB5113 mixed bed ion exchange resin (BDH) and filtered through a 0.22 μm filter (Millipore) and kept at 4°C. Resolving gel mixes were polymerized with a final concentration of 0.05 % (v/v) TEMED (N,N,N',N'-tetramethyl-ethylenediamine) and 0.05 % (w/v) ammonium persulphate; stacking gel mixes with a final concentration of 0.1 % (v/v) TEMED and 0.1 % (w/v) ammonium persulphate. Both mini (plates 12 x 8 cm) and standard size (plates 20 x 13 cm) gels were cast and run using apparatus purchased from Cambridge Electrophoresis. Samples to be analysed by SDS-PAGE were resuspended in SDS-sample buffer and incubated in a boiling water bath for 2 minutes. Bead-bound proteins were eluted in SDS-sample buffer by boiling for 5 minutes. SDS-PAGE was carried out in SDS-PAGE running buffer (25 mM Tris base, 192 mM glycine, 0.1 % (w/v) SDS). Mini gels were typically run with a limiting voltage of 230 V at the beginning of the run and with current limiting at 20 mA during the second half of the run; standard gels were run at 250 V and at 30 mA, respectively. Electrophoresis was continued until the bromophenol blue dye front had just run off the bottom of the gel.

2.24. Detection of proteins by coomassie blue staining.

Gels for autoradiography were stained for 5 minutes with Coomassie blue (5 g/l Coomassie blue R250, 45 % (v/v) methanol, 45 % (v/v) glacial acetic acid); gels for analysis by scanning densitometry were stained for 20 minutes. Gels were destained in hot destain (25 % (v/v) methanol and 7 % (v/v) glacial acetic acid.

2.25. Immunoblotting

Protein samples were separated by SDS-PAGE and transferred to 0.2 μm nitrocellulose (Hybond C-super) membrane in transfer buffer (20 mM Tris base, 150 mM glycine, 0.1 % (w/v) SDS, 20 % (v/v) methanol). After transfer, proteins on the membrane were visualised by staining with 0.2 % (w/v) Ponceau-S in 3 % (w/v) trichloroacetic acid and destained. The positions of the molecular weight markers and the positions of the lanes were marked on the membrane with a soft pencil and the protein staining was washed off again using 1 x TBST. The membrane was blocked with TBST containing 10 % (w/v) skimmed milk powder for 1 hour on a rotating platform. The membrane was then incubated with the primary antibody in TBST, 10 % (w/v) skimmed milk powder for one hour at room temperature. The membrane was then washed twice for 10 minutes with TBST at room temperature. The primary antibody was detected by probing the membrane with a secondary anti-mouse, anti-rabbit or anti-goat antibody conjugated to horseradish peroxidase (Dako) diluted 4000 fold in TBST containing 2 % (w/v) skimmed milk powder for 1 hour at room temperature. The membrane was washed twice with TBST containing 2 % (w/v) skimmed milk powder and then twice with TBST alone at room temperature to remove unspecifically bound antibodies. Binding was detected using the enhanced chemiluminescence system (ECL, Amersham) according to the manufacturer's instructions.

2.25. Cell cycle arrests.

HeLa OHIO cells (obtained from the Cell Culture facility at ICRF, Clare Hall, South Mimms, UK) were arrested at metaphase by the addition of nocodazole to the growth media. Briefly, cells seeded into quadruplicate 6 well plates at 50% confluence were treated with 1 μM nocodazole for 16 hours. The drug was removed by gently washing the cells twice in E4 media containing 10% FCS. Cells were harvested at various time points after the release either by treatment with trypsin (for FACS analysis) or were lysed directly in 1x SDS-sample buffer. HeLa OHIO cells were arrested at the beginning of S-phase using a double thymidine block. Cells seeded into quadruplicate 6 well plates at 30% confluence were treated with 2 mM thymidine for 16 hours. The thymidine was removed by two washes with E4 containing 10% FCS and returned to the incubator for 8 hours. The cells were then treated with 2 mM thymidine for a further 12 hours and released into fresh E4 containing 10% FCS. Cells were harvested at various time points after release as before. Swiss 3T3k cells were arrested at G_0 by serum starvation. Cells seeded into quadruplicate 6 well plates at 40% confluence were arrested by incubation in E4 media containing 0.2% FCS for 48 hours. Cells were released into the cell cycle by the addition of fresh media containing 10% FCS and harvested at various time points as described before. HeLa OHIO cells were arrested in S-phase using hydroxyurea. Cells seeded into quadruplicate 6 well plates at 30% confluence were treated with 2 mM hydroxyurea for 16 hours and released into the cell cycle by washing the drug away with two changes of E4 containing 10% FCS. Cells were harvested at various time points as before.

2.26. Preparation of cell lysates.

Cells were either lysed directly in the plates in which they were grown, or they were scraped from the plates and pelleted before being lysed. For plate lysis: 10 cm plates of 70-80% confluent cells were washed three times with PBS and 1 ml of lysis buffer containing 100 mM Tris-Cl pH7.5, 150 mM NaCl, 1% NP-40, 1 μg/ml pepstatin, 2 μg/ml aprotinin, 1 μg/ml leupeptin, 10 μg/ml benzamidine was added to the plate and incubated at 4°C for 30 minutes. Cells were resuspended in the lysis buffer and cellular debris was pelleted centrifugation at 20 800 g at 4°C. Alternatively, 10 cm plates were washed three times with PBS and the cells were scraped into 1 ml of PBS. The cells were pelleted by centrifugation

washed extensively in ddH_2O, mounted onto slides with a drop of glycerol and sealed with nail varnish.

2.29. Transient transfection of plasmid DNA into tissue culture cells.

DNA for transfection was prepared either using the Qiagen midiprep kit (Catalogue number 12145, Qiagen, Germany) or $CsCl_2$ density gradient centrifugation. For calcium phosphate mediated transfections, 459 µl of ddH_2O was added to 11 µl of 1mg/ml DNA which was made to 250 mM $CaCl_2$ by addition of 69 µl of tissue culture grade 2 M $CaCl_2$. 11 µl of 35mM $(Na_2H)PO_4$/35 mM $(NaH_2)PO_4$ was added to 550 µl of 2x HBS containing 42 mM HEPES and 364 mM NaCl pH 7.05 (with NaOH). While air was bubbled through the HBS/PO_4 mixture, the $DNA/CaCl_2$ mixture was added dropwise. This was incubated at room temperature for 30 minutes, and vortexed for 10 seconds before 300 µl was added dropwise to the surface of a 35 mm dish seeded with cells 24 hours previously at 30% confluence in the standard E4 medium containing 10% FCS. The transfection complexes were mixed with the media by gently swirling the plates which were then returned to the incubator. 16 hours later media and transfection complexes were removed and the cells were washed once with versene (0.02% w/v EDTA in PBSA) to remove free Ca^{2+}, and fresh media was added. Cells were analysed between 24 and 48 hours post-transfection. Alternatively, transfections with Superfect™ Transfection Reagent (Qiagen, Germany), FuGENE 6 Transfection Reagent (Roche Molecular Biochemicals, USA) or *Trans*IT® Transfection Reagent (Mirus/PanVera, USA) were performed according to the manufacturers instructions.

2.30. Microinjection of plasmid into living HeLa cells and microscopy of injected cells.

Cells were grown to 40% confluency before being subjected to a double thymidine block. The cells were microinjected using an Eppendorf automatic microinjection system mounted on an Zeiss Axiovert 35. Images were obtained on an Axiovert 135TV equipped with a Princeton Instruments MicroMax1300 camera driven by AQM200 image acquisition software. The microscope was adapted for live cell imaging (37°C and CO_2 supply). Phase and fluorescence shutters were controlled by Lambda 10-2 (from Sutter Instruments). Exposure times were 50 ms for phase and 400 ms for fluorescence pictures. The concentration of plasmids used was 50 ng/µl. The volume injected is difficult to estimate

because the system is only semiquantitative; although it was in the 1-10 fl-range. The injection marker was TRITC-labelled high molecular weight dextran at 0.5 mg/ml. Images were converted to 8-bit images and assembled using Adobe QuickTime and Adobe Photoshop.

2.31. Peptide synthesis and antibody production.

2.31.1. Preparation of peptides for use as immunogen.

Peptides were synthesised by the Peptide Synthesis Laboratory at 44 Lincolns Inn Fields, London and were supplied as lyophilised powders. Peptides were dissolved at a concentration of 10 mg/ml in dd H_2O. For use as immunogens peptides were coupled to KLH by mixing 2 ml of 10 mg/ml KLH and 1 ml of peptide at 10 mg/ml in dd H_2O . While stirring 1 ml of 0.3% gluteraldehyde in dd H_20 was added over 30 minutes to the KLH/peptide solution. KLH-peptide conjugate was buffered to pH 9 by dilution in PBS and this was supplied to the antibody production facility at ICRF Clare Hall, South Mimms who via the external contractor Murex, followed standard immunisation protocols.

2.31.2. ELISA.

50 µl of antigen at a concentration of 2 µg/ml was applied overnight to the wells of a Falcon 3912 microtitre plate in 100 mM sodium carbonate/bicarbonate buffer pH 9.6. The wells were emptied and blocked with 1 %NP-40 in PBS for one hour and then rinsed with water. Serum from hyperimmune animals was diluted serially 1/100-1/256000 in 1% NP-40 in PBS and applied to the wells of the dish for one hour. Wells were rinsed three times with 0.1% NP-40 in PBS and three times with PBS. 50 µl of HRP-conjugated to swine anti-rabbit secondary antibody was applied to the wells at a dilution of 1/2500 for one hour. Wells were rinsed three times with 0.1% NP-40 in PBS and three times in PBS. HRP-conjugated secondary antibody was detected by the addition of 50 µl of tetramethylbenzimidine at a concentration of 0.1 mg/ml in citrate phosphate buffer containing 0.02% hydrogen peroxide.

2.31.3. Preparation of peptide affinity chromatography columns.

CnBr-activated Sepharose (AP Biotech, Upsalla, Sweden, 17082001) was swollen in pH 2 HCl. The swollen resin was washed with 1 L of dd H_20 and then 100 ml of 100 mM

sodium carbonate buffer. Peptide dissolved in dd H_20 at a concentration of 1 mg/ml was added to the resin at a ratio of 1 mg peptide/ml of gel immediately. The resin-peptide mix was incubated overnight on a rotating wheel. The peptide coupled resin was washed with of 100 ml Tris-HCl pH 8, 0.5 M NaCl, and with 100 ml of 100 mM citrate buffer pH 4, 0.5 M NaCl and again with 100 ml Tris-HCl pH 8, 0.5 M NaCl and finally with 100 ml of 100 mM citrate buffer pH 4, 0.5 M NaCl. Unreacted sites on the resin were blocked with 100 mM Tris-Cl pH 8, 0.5 M NaCl.

2.31.4. Affinity purification of antibodies.

5 ml columns were pre-eluted with 10 ml of 100 mM NaOAc pH 2 and equilibrated in PBS. 10 ml crude serum was applied to the resin and both ends of the column were sealed and it was placed on a rotating wheel overnight at 4°C. The peptide-resin/serum was allowed to settle before the flow through was collected by gravity flow. The column was washed with 50 ml of 20 mM Tris-Cl pH 8, 5 mM EDTA, 0.5 M NaCl, 1 % NP-40 followed by 50 ml of 20 mM Tris-Cl pH 8, 5 mM EDTA, 150 mM NaCl. Bound antibodies were eluted in 15 ml 100 mM sodium citrate pH 3. Any remaining antibodies were eluted with 15 ml glycine-HCl pH 2.2. Eluates were immediately neutralised by addition of 5 ml 1 M Tris-Cl pH 8.8. Both eluates were dialysed against PBS overnight. Eluates were tested for antibodies by ELISA.

Chapter Three

Construction of cDNA libraries.

The identification of substrates of protein kinases and phosphatases has been notoriously difficult. The lack of success of systematic screens probably relates, at least in part, to the nature of the interaction between a kinase (or phosphatase) and a substrate. Most screening techniques, such as yeast-2-hybrid and far-western, rely on the establishment of relatively stable interactions between two or more molecules (Cowell, 1997; Fukunaga and Hunter, 1997b). Unfortunately the interactions between a protein kinase and its substrates are often transient and so techniques using interaction as a selection criterion may not necessarily be suitable. Certainly, some protein kinase-substrate interactions are strong enough (K_d in the μM range) to be detected by interaction based screens. It is also possible that other proteins which are not necessarily substrates but which have important roles in regulating either the activity, or localisation of a kinase could be identified by interaction-dependent screens.

3.1. Which type of screen and which type of library?

Yeast-2-Hybrid screens have met with limited success in identifying novel binding partners or substrates for CDKs and cyclins. The reasons for this are two-fold. Firstly, overexpression of many cyclins in yeast is toxic due to their inappropriate expression during the cell cycle and consequent association with endogenous CDC28. Secondly, if a CDK were to be used instead of a cyclin as the bait protein, then any potentially interesting clones would be masked by cyclin clones that would predominantly be identified, or the important enzyme-substrate interactions may be dependent on the formation of CDK:cyclin complexes. Clearly, yeast-2-hybrid screens are severely limited in their scope for the identification of substrates and interacting partners for CDKs and cyclins. *In vitro* screens, such as 'far-western' screening and solid-phase phosphorylation screening face a different set of problems when using CDKs or cyclins as bait. Both screens require the purification of substantial amounts of active, recombinant protein, and this can be problematic (Fukunaga and Hunter, 1997b). Another potential shortcoming of these screens is the dependence on *E. coli* as the expression organism for target proteins. Many eukaryotic proteins expressed in bacteria are either insoluble, inappropriately folded or both, making it likely that both genuine interactors and substrates will be missed; raising the possibility that

Figure 3.1. Genomic map of λYES and its derivatives and the scheme for preparing the vector for library construction.

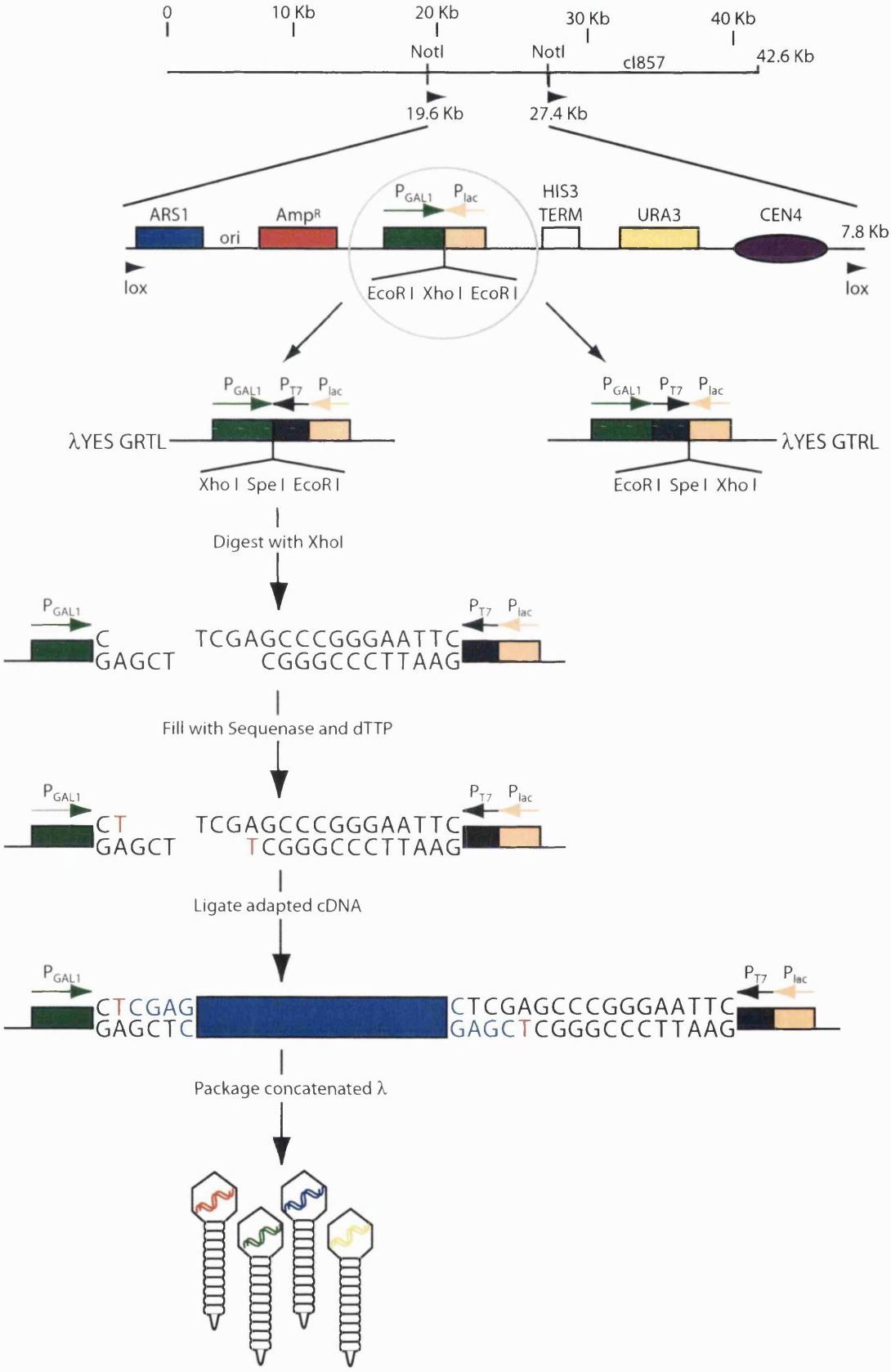


Figure 3.1. GRTL and GTRL refers to the orientation of the polylinker (R) relative to the GAL (G), lac (L) and T7 (T) promoters. Circular Lambda is digested with Xho I and the half-sites are filled with a single T-residue making the half-sites self-incompatible. Adapted cDNA is ligated into the modified half-sites and the recombinant molecules are packaged into infective phage particles.

containing ampicillin, and the *ori* origin of replication which permits low copy number propagation in *E. coli*. For selection in *S. cerevisiae* strains that are auxotrophic for uracil biosynthesis, the vector contains the URA3 selectable marker. The ARS1 origin of replication causes the plasmid to be replicated once per cell cycle, and the CEN4 sequence ensures that the plasmids are partitioned appropriately into the two daughter cells during mitosis. To facilitate gene analysis, λYES contains a plasmid (bearing the genetic elements described above), which can be excised using *Cre*-dependent recombination. The plasmid sequence is flanked by lox sites which in the presence of the *Cre* recombinase from phage P1, undergo recombination to release a 7.8 kb plasmid (figure 3-1).

λYES is derived from λGT6 which is identical to λGT7 except that the deletion in the c*I* gene is replaced with the full length c*I*857 gene. Wild type c*I* encodes repressor which inhibits the expression of genes essential for the lytic lambda life cycle. c*I*857 encodes a temperature-sensitive mutant of the repressor that at high temperatures (37°C) is unable to repress expression, causing transition to the lytic growth cycle. Vector DNA was prepared in *E. coli* by transforming the vector into *E. coli* DH5α and selecting for growth on ampicillin containing plates at 30°C. Due to the presence of the temperature-sensitive repressor the lytic pathway is suppressed at low-temperatures and the vector is propagated as a low copy number extra-chromosomal lysogen. Consequently, colonies were formed when transformed bacteria were plated on ampicillin containing plates incubated at 30°C, but not when duplicate plates were incubated at 37°C (Elledge *et al.*, 1991b).

I found that large amounts of circularised Lambda DNA could be obtained using conventional alkaline-lysis methods, and purified using CsCl density centrifugation. Typically, 400-500 μg of vector could be prepared from a one litre culture of *E. coli* harbouring the modified λYES. I found that this approach was significantly simpler than traditional methods requiring the growth of large amounts of phage, from which DNA is purified.

3.2.1. Vector preparations.

There are many control steps in library construction that are essential for making a good library containing a high fraction of recombinant clones with full length-inserts. The first of these is to ensure that vector is completely digested; as any undigested material would be

Figure 3-2. Scheme showing the steps in *Xenopus laevis* cDNA library construction

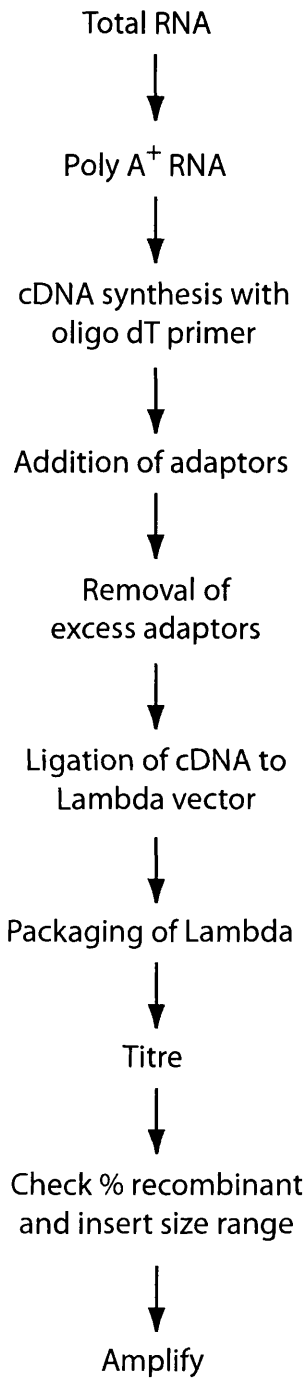


Figure 3-2. Total RNA is purified from *Xenopus laevis* oocytes, which is then used as a source for the purification of poly A^+ RNA. cDNA is synthesised using AMV reverse transcriptase and an oligo dT primer. Adapters are constructed and ligated to the cDNA modifying the blunt ends of the cDNA, and once the excess adapters have been removed and the cDNA selected for products greater than 500 bp, it is ligated into the λYES phagemid which has been digested with Xho I. The phage genomes are packaged using *in vitro* packaging reactions. The library is titred by infecting small aliquots of the packaging reactions into LE 392 *E. coli*. The number of phage containing cDNA inserts is checked using PCR which also allows the determination of the mean insert size. The primary phage library is stored at -80°C or amplified.

obtained from one gram of oocytes, and the quality of the RNA was assessed by agarose gel electrophoresis. Sharp staining of rRNA bands with a ratio of 2:1 is indicative of undegraded RNA. If these bands are smeared or degraded, the mRNA population would likely be completely degraded as it represents only 2% of the total RNA population. Figure 3-3 shows a non-denaturing agarose gel of total *Xenopus laevis* RNA, and although the RNA was heated and cooled quickly prior to loading, it is likely that the additional bands (bracketed) are due to secondary structures in the RNAs. Figure 3-3 suggests that the RNA obtained from the oocytes was of high quality, and should be suitable for subsequent preparation of mRNA.

The total RNA was passed over an oligo-dT cellulose column to isolate poly A^+ mRNA. Figure 3-4 shows a chromatogram of the fractions from an oligo dT column. It is important that poly A^+ RNA be purified by at least two passages through an oligo dT affinity resin to eliminate contaminating rRNAs and tRNAs. The RNA was heated to disrupt secondary structures and loaded onto an oligo dT cellulose column in 0.5 M NaCl to promote binding of the poly A tract to the oligo dT. Only a small percentage of total RNA is polyadenylated mRNA; and the rest of the RNA was discarded by washing the column until the A_{260} of the flow through was close to zero. The interaction between the oligo dT agarose and polyadenylated RNA was disrupted by eluting in low salt buffer then finally in water. Eluted material was adjusted to 0.5 M NaCl, heated to 65°C for ten minutes, snap cooled then reloaded onto the column, which was washed with high salt buffer until no material absorbing at A_{260} was eluted. Fractions containing polyadenylated mRNA are seen to elute in low salt buffer. These two-fold purified poly A^+ fractions were largely free of contaminating rRNAs and tRNAs as assessed by agarose gel electrophoresis of the purified fractions (not shown), providing a template which could be used for first strand cDNA synthesis.

To check the purity and integrity of the poly A^+ RNA I added it to a cell free translation system derived from lysates of rabbit reticulocytes. Figure 3-5 shows the results of translation reactions using either total RNA or oligo dT purified RNA as template in the presence of [^{35}S]-methionine, which have been analysed using SDS-PAGE and autoradiography. The presence of many sharp bands, some larger than 100 kDa, in the lanes using mRNA as template indicate that there are many long mRNAs but few

Figure 3-4. Purification of poly A^+ RNA

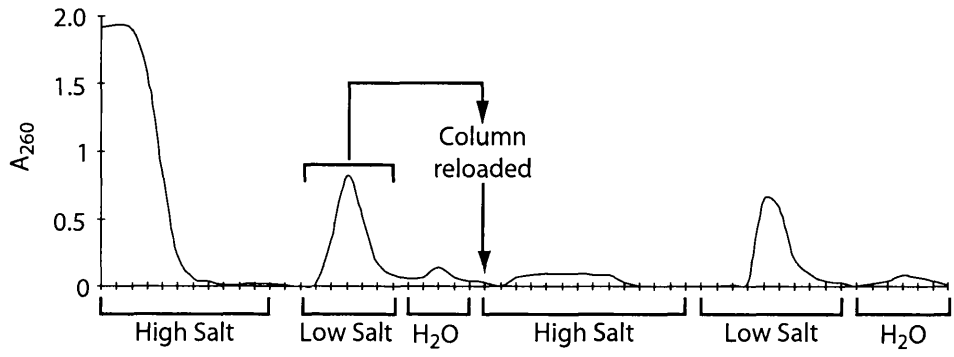


Figure 3-4. 10 mg of total RNA from stage VI *Xenopus laevis* oocytes was heated to 65°C for 10 minutes and placed on ice for 5 minutes before being loaded onto an oligo dT agarose column. The column was washed with buffer containing 0.5 M NaCl until the A_{260} of the flow through reached 0. Fractions were eluted with buffer containing 50 mM NaCl. The column was washed with water and the eluted fractions were adjusted to 0.5 M NaCl before being re-heated, cooled, pooled and reloaded onto the column. The column was washed with buffer containing 0.5 M NaCl and fractions containing poly A^+ RNA were eluted with buffer containing 50 mM NaCl.

Figure 3-5. Quality check of the RNA used for library construction

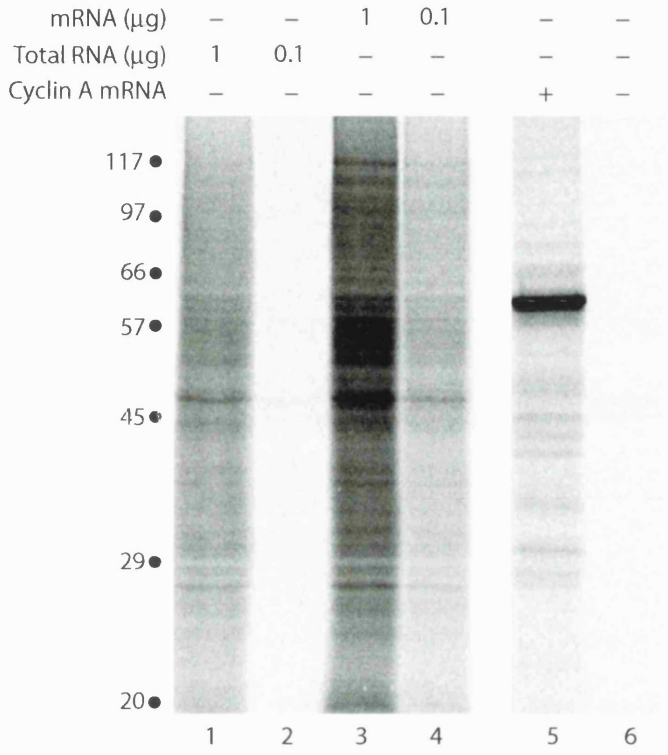


Figure 3-5. Total *Xenopus laevis* oocyte RNA (lanes 1 and 2), poly A^+ RNA (lanes 3 and 4) or cyclin A mRNA transcribed from a cDNA clone (ESP11) were added to a rabbit reticulocyte lysate translation mix in the presence of [^{35}S]-methionine. The reactions were diluted into SDS-gel loading buffer, and analysed by SDS-PAGE. Molecular weight standards were run in an adjacent lane. The dried gel was exposed to a fluorographic screen.

impurities present, a conclusion also supported by an A_{260}:A_{280} ratio close to 2. Should there be few high molecular weight species, an abundance of low molecular weight products or if the bands appear smeared, then this would be indicative of degraded RNA. If this were the case then the RNA should not be used for cDNA synthesis.

First strand cDNA synthesis was primed using oligo dT which tends to generate lower yields, but increases the number of full-length clones obtained compared to syntheses using random hexamer oligonucleotides as primers. Reverse transcription was performed using AML reverse transcriptase. By incorporation of $\alpha[^{32}P]$-dCTP during synthesis it was possible to determine the efficiency of first strand synthesis which is the yield limiting step in synthesis of double-stranded cDNA. In the first strand synthesis I used 1μg of mRNA as template and typically generated approximately 50-100 ng of single-stranded cDNA; a characteristic yield of only 5-10% of the starting material. Figure 3-6 shows an autoradiogram of a 1.4 % denaturing agarose gel of the products of first strand cDNA synthesis reactions using either an RNA ladder or purified poly A^+ RNA as template. Reverse transcription of the RNA ladder was performed using random hexamers as primer as these RNAs do not possess polyA^+ tracts. In the reactions using RNA ladder as template several synthesis products can seen which correspond to single-strand cDNA copies of the RNA ladder. When poly A^+ RNA was used, a smear extended from approximately 100 bp (relative to the double-stranded DNA markers) to nearly 8 Kbp, with most of the products present in the 700-4000 bp range.

Second-strand synthesis is an efficient reaction which can easily be monitored by the incorporation of $\alpha[^{32}P]$-dCTP (not shown). The reaction is catalysed by DNA polymerase I using the first cDNA strand as template, presumably primed from nicks introduced into the original RNA template by limiting digestion with RNAse H, which is an inherent activity of AMV reverse transcriptase. Second-strand synthesis fragments were ligated using *E. coli* DNA ligase, an enzyme unable to ligate blunt ends and therefore unlikely to generate chimeric cDNAs. Blunt-ended double-stranded cDNA was then ligated to specially prepared adapters with T4 DNA ligase to make it compatible with the T-filled half-sites of λYES GRTL. Because the ends of the vector were not conventional half-sites, I prepared adapters that were suitable for cloning into the vector. Figure 3-7 shows a 20% native polyacrylamide gel of the single-stranded oligonucleotides which were

Figure 3-6. First strand cDNA synthesis using *Xenopus laevis* poly A$^+$ RNA as template.

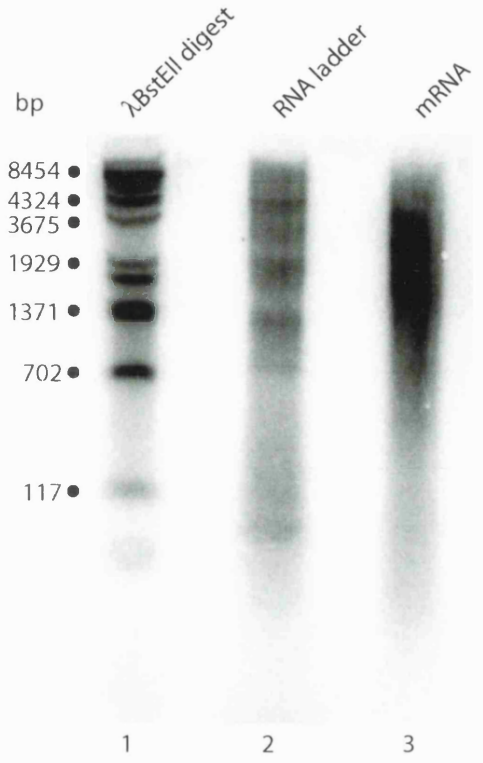


Figure 3-6. One µg of Poly A$^+$ RNA from *Xenopus laevis* oocytes (lane 3) or an RNA ladder (lane 2) was used as template in reverse transcription reactions containing $\alpha[^{32}P]$-dCTP. Reverse transcription using poly$^+$ RNA was primed using an oligo dT primer. Reverse transcription using the RNA ladder was primed using random hexamers. 10% of the reaction products were separated on a 1.4% denaturing agarose gel. Lambda digested with BstE II and end labelled using $\gamma[^{32}P]$-ATP were run as molecular weight markers (lane 1). The dried gel was exposed to a fluorographic plate.

Figure 3-7. Adapter construction

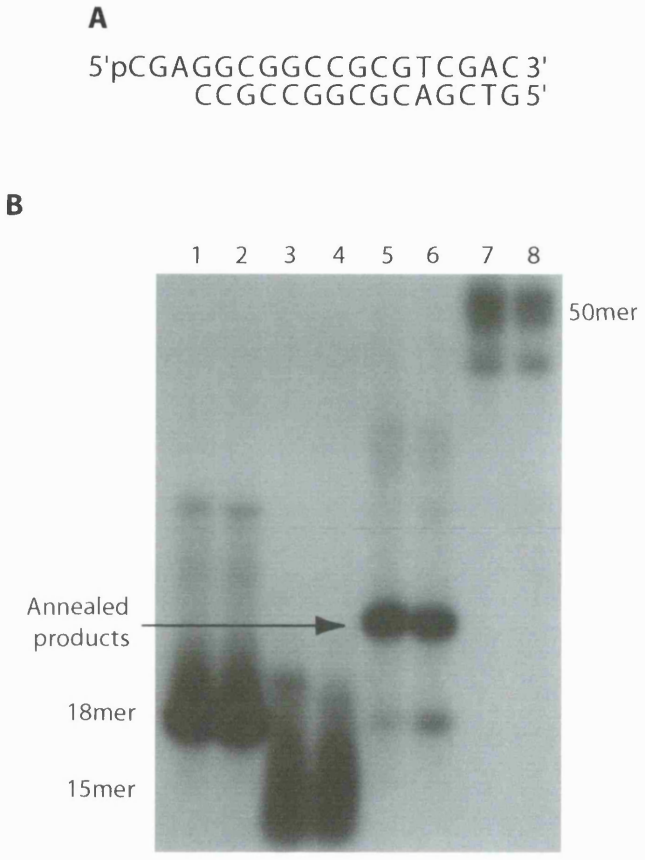


Figure 3-7. (**A**) Sequences of the oligonucleotides used in adapter preparation. (**B**) 20% Native polyacrylamide gel of [^{32}P]-phosphorylated oligonucleotides and annealed adapters. Lanes 1 and 2 contain phosphorylated 18mers, lanes 3 and 4 contain phosphorylated 15 mers, and lanes 6 and 7 contain the annealed 18 and 15mers. Lanes 7 and 8 contain a phosphorylated random 50mer as a marker.

end-labelled using $\gamma[^{32}P]$-ATP and T4 polynucleotide kinase and the products of an annealing reaction using these oligonucleotides. Stoichiometric amounts of each oligonucleotide were used to ensure that the reaction could proceed to completion. These adapters are ligated to the cDNA after which, excess adapters were removed by precipitating the ligation reaction and gel purifying the ligation reaction on a 1% agarose gel. A slice of gel containing cDNAs longer than 500bp was cut from the gel and melted using hot phenol. The cDNA was purified by ethanol precipitation from the aqueous phase. Yields from this step were frequently low, and recovery efficiencies were poor. To improve recoveries I used a commercially available gel purification kit from Qiagen. Size selecting for cDNAs greater than 500 bp ensured that the resulting libraries did not contain a predominance of small inserts. This is an important step as on a molar basis there would be many more unligated adapters and small cDNA fragments than large cDNAs. This would result in libraries containing only adapters and very short cDNA inserts.

3.2.3. Library characteristics and features.

Once I had optimised each step of this process, I ligated 50 ng of adapted cDNA to 1 µg of T-filled vector in a 4 µl reaction to achieve the highest possible concentrations of cDNA and vector ends. One-µl aliquots of the reaction were packaged *in vitro* into infective phage particles, which were titred on LE392 *E. coli*. To check that the resulting phage contained cDNA inserts I designed primers complimentary to the GAL1 and lac promoter regions and used these to amplify inserts directly from phage plaques using the polymerase chain reaction. This allowed me to estimate both the percentage of the phage which were recombinant and the mean insert size of the library. These data are summarised in Figure 3-8.

As I intended to use this library to express inserts from the T7 promoter using T7 RNA polymerase I needed to convert the clones to a form from which I could easily prepare pools of plasmid DNA. These pools would then be used in coupled transcription-translation reactions to generate $[^{35}S]$-methionine labelled proteins. One feature of the λYES phagemids is that they have been modified from their original λGT6 backbones to contain a plasmid flanked by two lox sites. The *Cre* recombinase can catalyse a recombination event between the two lox sites (see figure 3-1) (Abremski and Hoess, 1984; Hamilton and Abremski, 1984), and this results in the excision of a 7.8 kb plasmid

Figure 3-8. Characteristics of the final *Xenopus laevis* oocyte cDNA library.

1° Titre (pfu/μg)
1.6×10^6

2°Titre (pfu/ml)
3.5×10^{13}

% Recombinant
90

Mean Insert Size
1.7 kb

Insert Size Range
0.5 - 6.5 kb

Figure 3-8. *Xenopus laevis* oocyte cDNA library characterstics. Primary titre was determined by infection of packaged virus into LE392 and plating serial dilutions into top agar overlaid LB plates. Secondary titre was determined as for primary titre following a single round of phage amplification as before. The number of clone containing phage was determined by picking 10 whole plaques into PCR reactions containing primers flanking the polylinker and performing 30 cycles of amplification. Amplification products were analysed by agarose gel electrophoresis. Insert size range was estimated from the PCR reactions containing phage plaques, and the mean insert size was calculated from this.

(pYES, see figure 3-1) which can be propagated in *E. coli*. Although this reaction can proceed *in vitro*, by far the most efficient way to induce the recombination event is to infect known amounts of virus into an *E. coli* strain expressing the *Cre* recombinase from a Lambda lysogen (BNN132). Elledge *et al.* (1991) reported that plaque forming units were converted to ampicillin resistant colonies with an efficiency of nearly 50% when using BNN132 (Elledge *et al.*, 1991b). Disappointingly however, I was never able to reproduce this efficiency using the published protocol, and could never achieve a conversion of pfu to cfu of greater than 1%. Moreover, when bacteria infected with phage were grown in culture, they would often become lytic, yielding low yields of poor quality DNA. It is possible that this was due to the phage entering their lytic life-cycle, and not remaining lysogenic as they should. Also, strain BNN132 is a JM107 derivative which contains the endA nuclease and bacteria containing this nuclease have a tendency to yield low quality plasmid DNA.

3.3. Assessing the suitability of the *Xenopus* library for screening purposes.

Having invested significant time and effort in constructing this library, I decided to investigate whether pools of plasmid DNA could be produced from the phage, and whether these pools could direct *in vitro* coupled transcription-translation. Firstly however, to check that indeed [^{35}S]-labelled proteins could be synthesised from the cloned inserts, I randomly picked 6 clones which had been converted from phage to plasmid using *in vivo Cre*-based recombination to release the plasmid pYES from the phage, and used a commercially available kit from Qiagen to maximise purity and yield, prepared miniprep DNA from these clones. 1 µl of plasmid, corresponding to approximately 20 ng DNA, from each clone was used as template in a 10 µl coupled transcription-translation reaction. To test how many clones could be pooled, aliquots of each plasmid were mixed and 1 µl of this mixture was used in a 10 µl translation reaction. The products of the reactions were analysed by SDS-PAGE and autoradiography (see figure 3-9). Labelled translation products were synthesised from 5 out of 6 of the single clones. Clone 5 did not direct synthesis of a labelled protein, which may be due to either the clone being incorrectly oriented relative to the T7 promoter or it lacking a functional start codon (or internal methionine codon), or indeed a *Xenopus* derived cDNA insert. Though five out of this small sample of six clones

Figure 3-9. Coupled transcription and translation of random clones from the *Xenopus laevis* library.

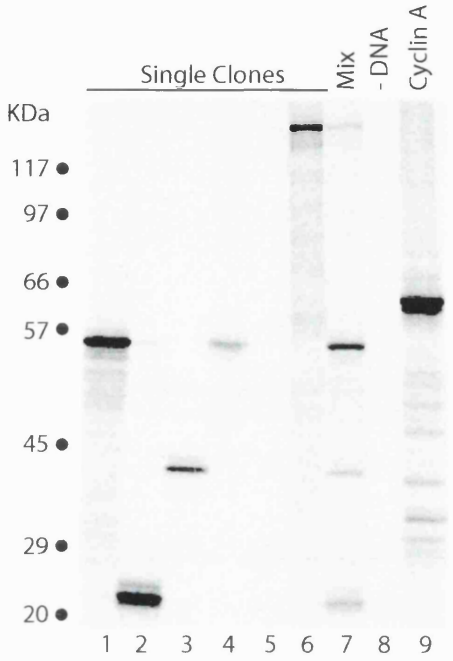


Figure 3-9. Fluorogram of 12.5% SDS-polyacrylamide gel of pYES library clones transcribed and translated in the presence of [^{35}S]-methionine. Individual clones (lanes 1-6) were individually transcribed and translated in the presence of [^{35}S]-methionine. A mixture of all six clones (lane 7) was transcribed and translated to show the decrease in the amounts of product produced by increasing the number of different templates in a reaction. No DNA (lane 8) and a plasmid containing cDNA for cyclin A (lane 9) were used as negative and positive controls respectively.

tested expressed proteins, only about half of the cloned inserts are expected to be appropriately oriented relative to the T7 promoter because this library was not constructed directionally.

3.3.1. The *Xenopus laevis* oocyte library was not suitable for screening purposes.

The next challenge was to scale up the *in vivo* conversion of phage to plasmid so that pools containing approximately 100 clones could be produced and to make sure that the pools of plasmid DNA would support efficient transcription and translation. I titred the number of phage required to generate approximately 100 colony forming units when infected into BNN132 and plated on ampicillin containing plates. Typically this required 10000 phage. I scraped the cells from plates, grew them in culture and prepared maxiprep DNA from them. When the DNA from these 'pooled' clones was used as template in coupled *in vitro* transcription-translation reactions a number of products were visualised by SDS-PAGE and fluorography. On average however only approximately 10-12 bands were synthesised per reaction. For these experiments I used 10 µl translation reactions programmed with either 1 µg or 0.5 µg of plasmid DNA (2 µl or 1 µl respectively). The first four lanes of figure 3-10 show the reaction products of translations using 2 µl of template, while lanes 5 to 9 show the reaction products of reactions using only 1 µl of template. The intensity of the bands in the first four lanes is greater than of those in lanes 5-9, suggesting that a higher amount of template is required to enhance the efficiency of the reactions. Increasing the volume of template per reaction resulted in dilution of the reticulocyte lysate however, and caused a drop in the level of incorporation of [^{35}S] into translation products. When a similar experiment was performed using DNA derived from minipreps of identical cultures, there was insufficient plasmid template per clone to direct efficient transcription and translation (not shown).

It seemed that this library would be unsuitable to be used in small-pool expression cloning. I had intended to prepare 400 or so plasmid pools each containing approximately 100 clones by preparing cultures from plated cells, and extracting the DNA using an automated miniprep robot. The yields of plasmid from the miniprep robot were half those obtained when performed manually, and were equally poor at supporting efficient translation. Had I wished to continue with this library I would have needed to prepare

Figure 3-10. Testing the efficiency of transcription and translation using pools of plasmid derived from the *Xenopus laevis* library.

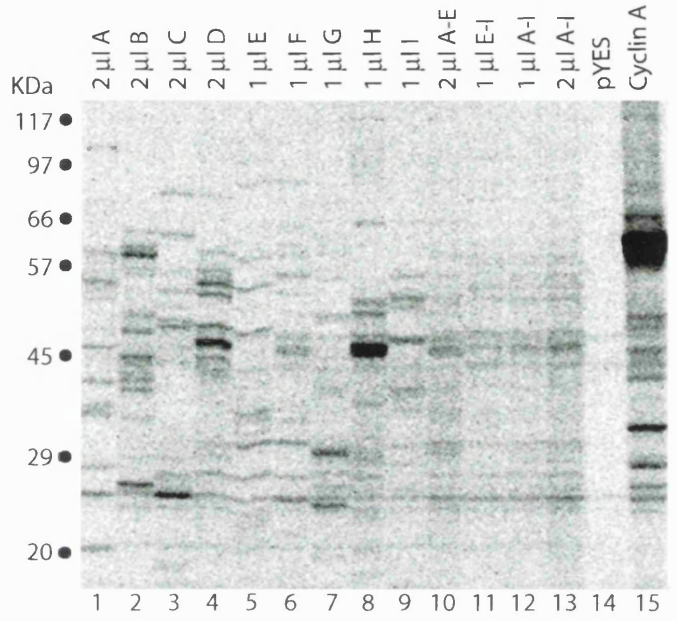


Figure 3-10. Fluorogram of 12.5% SDS-polyacrylamide gel of pooled pYES library clones transcribed and translated in the presence of [^{35}S]-methionine. Pools of plasmid containing 100 independent clones were prepared from colonies harvested off ampicillin containing plates, which were the products of infection of λYES library into *E. coli* strain BNN132. Two µl of plasmid from 4 independent pools (A-D) was used in coupled transcription-translation reactions in the presence of [^{35}S]-methionine (lanes 1-4). For comparison of the efficiency of translation, 1 µl of template from 5 independent pools (E-I) was also transcribed and translated (lanes 5-9). Small aliquots of pools A-I were mixed and 2 µl of this was used as template in a coupled transcription-translation reaction (lane 10). For comparing the concentration dependence on efficiency of translation 1 or 2 µl of template from a mixture of pools E-I, or A-I were transcribed and translated (lanes 11, 12 and 13 respectively). A single weakly translating 47 kDa protein can be detected when empty pYES is used as template (lane 14). A cyclin A containing template was used as a positive control for transcription and translation (lane 15).

maxiprep DNA from each pool to obtain sufficient amounts of plasmid. Given the amount of time to prepare maxiprep DNA, this seemed an unviable option and I began to construct a new library in a different, more suitable vector.

3.3.2. Finding an alternative strategy for preparing suitable libraries.

Given the problems I had encountered with converting phage to plasmid using the λYES based system, and the potential to transform *E. coli* efficiently using electroporation, I decided to construct a plasmid-based cDNA library. The manipulations of plasmid were much simpler than similar manipulations with phagemids. I used the pTEX-1 vector constructed in our lab by Dr. Stephan Geley which contains the T7 promoter proximal to the human influenza virus leader sequence (which enhances translation efficiency) and two copies of the c-*myc* epitope (figure 3-11). The presence of an extended polylinker, containing a Not I site, meant that I was also able to construct the library directionally. This meant that I would be able to effectively screen twice as many clones because all clones would be oriented correctly relative to the T7 promoter. I also decided to use mRNA prepared from human cells as the starting material for cDNA synthesis.

3.4. Constructing a HeLa cDNA plasmid library.

I prepared poly A^+ RNA from HeLa cells, and used this to prepare double stranded cDNA. First strand synthesis was performed using an oligo dT primer which contained a 5' Not I site to facilitate directional cDNA cloning and RNAse H deficient Moloney-Murine Leukaemia Virus (M-MLV) reverse transcriptase. Though the AMV reverse transcriptase has traditionally been the enzyme used to synthesise cDNA *in vitro*, the cloning of the M-MLV gene has provided a valuable alternative. The modification of the M-MLV reverse transcriptase to eliminate the endogenous RNAse H activity is significant because this activity is detrimental to the first strand cDNA synthesis reaction. Initiation of first strand synthesis requires the annealing of a complimentary primer (here oligo dT with the poly A^+ tract) and this hybrid RNA:DNA molecule is a substrate both for the polymerase activity of the reverse transcriptase and also for the RNAse H activity of the enzyme. The extent to which the RNAse H activity is able to hydrolyse the hybrid molecule prior to the initiation of transcription dictates the number of initiation events that can occur. Hydrolysis of the RNA in the hybrid molecule reduces the efficiency of first strand synthesis by effectively

Figure 3-12. Resolution of adapted cDNA from adapters and digestion fragments.

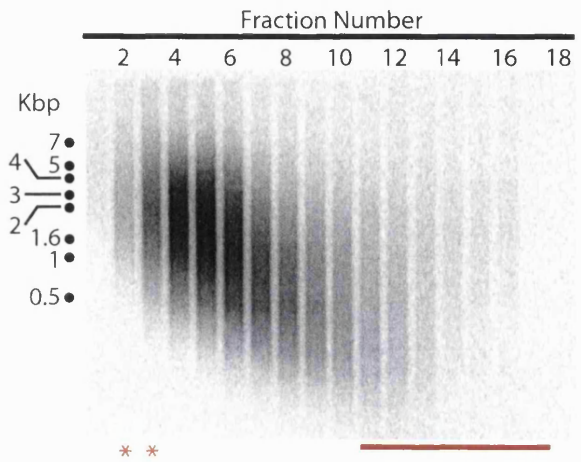


Figure 3-12. Fluorogram of adapted double-stranded [^{32}P]dCTP-labelled cDNA. Double-stranded cDNA was ligated to adapters. The adapted cDNA was digested with Not I enzyme to release 3' adapters and expose the 3' Not I site to allow directional cloning. The adapted, digested cDNA was fractionated using a Sephacryl-S1000 size exclusion column. Column fractions were collected and resolved on a 0.8% agarose gel. The dried gel was exposed to a fluorographic plate. A 1 Kb ladder was end labelled using γ[^{32}P]-ATP and run in an adjacent lane as molecular weight markers. The red stars indicate the fractions which were used in library construction, and the red bar shows where unligated adapters and the Not I restriction fragments elute.

Figure 3-13. Characteristics of the HeLa cDNA library.

1° Titre (cfu/ml)
4.3 x 10^6

% Recombinant
85

Mean Insert Size
1.3 kb

Insert Size Range
0.4 - 5 kb

Figure 3-13. HeLa cDNA library characterstics. Primary titre was determined by plating serial dilutions of *E. coli* electroporated with ligation reactions containing adapted cDNA and Not I/Sal I digested vector onto LB plates containing 50 μg/ml ampicillin. To determine the insert size range 20 individual colonies were picked directly into PCR reactions, and subjected to 30 rounds of amplification using primers flanking the polylinker. Amplification products were analysed by agarose gel electrophoresis. The mean insert size was also calculated from this data.

high titre libraries and because of the apparent ease of removing plasmid from the phage (Elledge *et al.*, 1991a). This library is far from optimal for small-pool expression cloning due to the poor quality and quantity of DNA obtained from BNN132. Yoshimi Tanaka has successfully isolated however a number of clones, including Wee1 and INCENP-E, from the λYES *Xenopus laevis* oocyte library in a small-pool expression screen for substrates of CDK1:cyclin B. Reassuringly many of these clones are full length which alleviates the need to re-screen libraries for missing 5' cDNA ends and confirms one aspect of the quality of this library.

The plasmid based HeLa library is much better suited to the demands of small-pool expression cloning. The cloning vector is small (2.8 Kb compared to 7.8 Kb of pYES), has a high copy number origin of replication and can be grown easily in most common laboratory strains of *E. coli* to yield relatively large amounts of plasmid DNA. The presence of the extended polylinker means that libraries can also be constructed directionally allowing the screening of more primary clones. One disadvantage of plasmid based libraries however, is their tendency to become biased during amplification, a problem not generally associated with phage libraries. This bias is caused by the growth advantage that *E. coli* harbouring smaller plasmids have compared to cells harbouring larger plasmids. Competition during growth in culture can cause larger clones to be lost while smaller ones become over-represented. This effect can be overcome by amplifying plasmid libraries on plates, or in semi-solid agar where there is little competition for resources. This was not a consideration here though, as when it came to preparing plasmid pools I did this directly from the primary library, not from an amplified stock.

Chapter Four

Developing and optimising a screening method to identify binding partners of cyclin A.

Having prepared the cDNA library to be used in the small-pool expression screen the next challenge was to develop a screening method which would identify protein partners for cyclins or CDKs. I hoped to use an affinity chromatography based screen, where the pools of transcribed and translated proteins would be searched for proteins which bound to cyclin A. The first step in this was to identify a suitable cyclin A expression system which would allow the production of large amounts of pure recombinant protein. Once this had been achieved I needed to find an appropriate method for immobilising the cyclin A and checking that the cyclin A-resin could efficiently capture known interacting proteins before embarking on a screen itself.

Many cyclins are notoriously difficult to express and purify using the common heterologous expression systems currently available such as *E. coli* and *Sf*9 cells (Kellogg *et al.*, 1995b). Cyclin B is especially susceptible to degradation and insolubility in both baculovirus infected *Sf*9 cells and *E. coli*, although limited success in purifying active complexes of CDK1:cyclin B has been met by co-infecting baculovirus encoding CDK1 and cyclin B into *Sf*9 cells, or by mixing lysates of these infected cells (Desai *et al.*, 1992). Our laboratory has experimented extensively with various *E. coli* strains and expression systems using different tags which are supposed to help with the purification of poorly expressed and insoluble proteins but has met with limited success. It is unclear what causes the problems with expression of cyclins in *E. coli* although one particular strategy has been successful for the expression of cyclin A in *E. coli*.

The purification of active CDKs in *E. coli* is also problematic. Although CDK2 expresses to high levels and can easily be purified as a GST-fusion protein from soluble lysates of *E. coli*, the generation of active CDK2 requires phosphorylation on T160. To achieve this an active CDK-activating kinase (CAK) must also be purified using a heterologous expression system (Darbon *et al.*, 1994; Poon *et al.*, 1993; Wu *et al.*, 1994). Human CAK comprises CDK7 complexed with cyclin H, and these proteins can to date only be efficiently expressed and purified from baculovirus-infected insect cells.

Alternatively, CDK2 can be also be purified from insect cells infected with a baculovirus encoding CDK2. An endogenous CAK in the insect cells is capable of phosphorylating T160 and activating the enzyme. There are also activities in the insect cells however capable of phosphorylating T14 and Y15 which causes some inhibition of the enzyme.

4.1. Bovine and human cyclin A2.

The predicted primary amino acid sequence of full length human cyclin A2 is shown in figure 4-1. Human cyclin A2 is a protein of 432 amino acids with a predicted molecular weight of 48.5 kDa but which has an apparent molecular weight of approximately 60 kDa through SDS-PAGE. Cyclin A2 is destroyed by the ubiquitin-dependent degradation machinery and the sequences which regulate the stability of cyclin A2 reside between amino acids 47-83 (Geley *et al.*, 2001). Cyclin A2 possesses two cyclin folds between residues 208-306 and 310-397. For comparison, the deduced amino acid sequence of the partial bovine homologue is also shown. The bovine cyclin A2 cDNA which was isolated by Adamcewski lacked 24 residues at its N-termini (GenBank accession number P30274), although given the high degree of homology between the human and bovine forms, it is likely that these residues are also highly conserved between the two species. Figure 4-1 also shows an alignment of human and bovine cyclin A2 with the conserved residues boxed in grey. If gaps corresponding to the missing 24 amino acids are including in the alignment then the human and bovine forms are 85.6% identical. If the missing 24 residues are neglected in the alignment, then the identity between the two homologues increases to 90.7%. Many of the residues which are different between the two proteins are conserved substitutions and do not cause differences between the structures of the human and bovine forms (Brown *et al.*, 1995; Brown *et al.*, 1999).

4.1.1. Cyclin A3 (ΔN170).

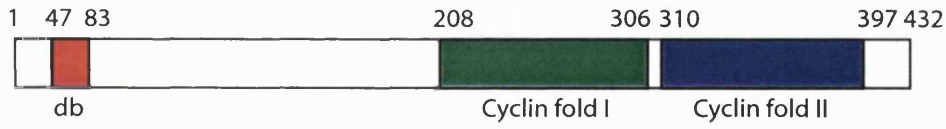
Deletion of the first 170 residues of cyclin A2 yields a protein of apparent molecular mass 30 kDa by SDS-PAGE, which is called cyclin A3 (GenBank accession number CAA48398). Cyclin A3 lacks the N-terminal sequences responsible for ubiquitin-dependent degradation, but retains both cyclin folds and is capable of binding and activating CDK (Brown *et al.*, 1995; Kobayashi *et al.*, 1992). Cyclin A3 was constructed by using the polymerase chain reaction to introduce an N-terminal Nco I site at residue 170 (thus changing the residues V170 and S171 to M170 and G171 respectively) and a C-

Figure 4-1. Sequence and diagramatical representations of full length cyclin A2 and the ΔN170 mutant cyclin A3.

(A) Sequence of full-length human and bovine cyclin A2.

```
                10          20          30          40
                |           |           |           |
  1 MLGNSAPGPA TREAGSALLA LQQTALQEDQ ENINPEKAAP VQQPRTRAAL 50
    ========== ========== ====EFQEDQ ENVNPEKAAP AQQPRTRAGL
                |           |           |           |
 51 AVLKSGNPRG LAQQQRPKTR RVAPLKDLPV NDEHVTVPPW KANSKQPAFT 100
    AVLRAGNSRG PAP=QRPKTR RVAPLKDLPI NDEYVPVPPW KANNKQPAFT
                |           |           |           |
101 IHVDEAEKEA QKKPAESQKI EREDALAFNS AISLPGPRKP LVPLDYPMDG 150
    IHVDEAE=EI QKRPTESKKS ESEDVLAFNS AVTLPGPRKP LAPLDYPMDG
                |           |           |           |
151 SFESPHTMDM SIVLEDEKPV SVNEVPDYHE DIHTYLREME VKCKPKVGYM 200
    SFESPHTMEM SVVLEDEKPV_SVNEVPDYHE DIHTYLREME VKCKPKVGYM
                |           |           |           |
201 KKQPDITNSM RAILVDWLVE VGEEYKLQNE TLHLAVNYID RFLSSMSVLR 250
    KKQPDITNSM RAILVDWLVE VGEEYKLQNE TLHLAVNYID RFLSSMSVLR
                |           |           |           |
251 GKLQLVGTAA MLLASKFEEI YPPEVAEFVY ITDDTYTKKQ VLRMEHLVLK 300
    GKLQLVGTAA MLLASKFEEI YPPEVAEFVY ITDDTYTKKQ VLRMEHLVLK
                |           |           |           |
301 VLTFDLAAPT VNQFLTQYFL HQQPANCKVE SLAMFLGELS LIDADPYLKY 350
    VLAFDLAAPT INQFLTQYFL HQQPANCKVE SLAMFLGELS LIDADPYLKY
                |           |           |           |
351 LPSVIAGAAF HLALYTVTGQ SWPESLIRKT GYTLESLKPC LMDLHQTYLK 400
    LPSVIAAAAF HLALYTVTGQ SWPESLVQKT GYTLETLKPC LLDLHQTYLR
                |           |           |           |
401 APQHAQQSIR EKYKNSKYHG VSLLNPPETL NL 432
    APQHAQQSIR EKYKNSKYHG VSLLNPPETL NV
                |           |           |           |
                10          20          30          40
```

(B) Domain and motif structure of full-length cyclin A2.



(C) Construct A3 ΔN170

Figure 4-1. Primary sequence and motifs of cyclin A2 and the truncation mutant A3. (**A**) Primary sequence of full length human cyclin A2 showing the sequence encompassing the destruction box(es) red and cyclin motifs I and II (green and blue respectively). The bold typeface indicates the sequence of cyclin A3 and conserved residues are boxed in grey. The underlined residues (VS) were changed to MG during the cloning of the cDNA into pET21d. (**B**) Domain structure of full length cyclin A2. (**C**) Domain structure of cyclin A3 with a C-terminal hexahistidine tag.

terminal Xho I site to remove the stop codon (see figure 4-1). The resulting cDNA was cloned into the expression plasmid pET21d (Novagen) in such a way as to add a C-terminal hexahistidine tag. When introduced into a suitable strain of *E. coli* (harbouring a DE3 lysogen encoding the T7 RNA polymerase) pET21d:cyclin A3 can direct the transcription and translation of hexahistidine-tagged cyclin A3.

4.1.2. Purification of cyclin A3 using Ni^{2+}-NTA agarose affinity chromatography.

Cyclin A3 was expressed to high levels in *E. coli* and was a soluble protein with average yields in the range of 10-30 mg/L of culture. The C-terminal hexahistidine tag allowed efficient purification of full-length cyclin A3 from high-speed lysates of *E. coli* harbouring a pET21d:cyclin A3 construct using Ni^{2+}-NTA affinity chromatography (figure 4-2). The conditions of induction of cyclin A3 have been shown to be important for the efficient expression and purification of the protein (Brown *et al.*, 1995; Kobayashi *et al.*, 1992). To maximise the expression levels and yields of purified protein the cells were grown for 30 minutes at 37°C in the presence of 0.1 mM IPTG before cooling to 30°C and continued growth for 3 hours. Growth at the higher temperature allowed the efficient synthesis of large amounts of RNA encoding cyclin A3. Maintaining growth at 37°C resulted in poor yields of cyclin A3 and this is likely due to the very high levels of cyclin A3 synthesis and a concomitant reduction in protein solubility at very high concentrations. Reducing the temperature to 30°C limits the expression of cyclin A3 and prevented the protein accumulating at too high concentrations.

A single purification step using Ni^{2+}-NTA affinity chromatography yielded cyclin A3 approximately 95% pure, with the major contaminants having similar molecular weights as the large and small subunits of the *E. coli* chaperone GroE (60 and 12 kDa respectively). These contaminants were largely removed by the inclusion of 5 mM ATP in the wash buffers. In cases where GroEL is known to co-purify with proteins overexpressed in *E. coli*, it has been reported that the inclusion of ATP in the wash buffers can help to remove these contaminants (Thain *et al.*, 1996). This suggests that the main contaminants which co-purified with cyclin A3 were the GroEL subunits. Attempts to express full length cyclin A2 in *E. coli* have been unsuccessful and yielded little soluble protein and it appears that the successful expression of cyclin A3 is dependent on the deletion of the N-

Figure 4-2. Hexahistidine tagged cyclin A3 can be purified by Ni^{2+}-chelate affinity chromatography.

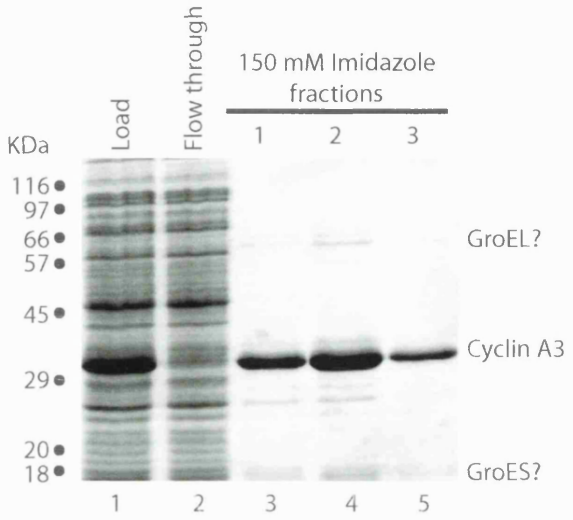


Figure 4-2. *E. coli* BL21(DE3)pLysS cells harboring pET21d:cyclin A3 were induced with 1 mM IPTG for 30 minutes at 37°C and then shifted to 30°C for 3 hours. Cells were lysed by freeze-thaw, and the cleared lysate (lane 1) was loaded onto a 1 ml Ni^{2+}-NTA agarose column. Hexahistidine tagged cyclin A3 was retained on the column and the flow through (lane 2) contained no detectable cyclin A3. The column was washed with increasing concentrations of imidazole and cyclin A3 was eluted using 150 mM imidazole (lanes 3-5). The major contaminants have similar molecular weights as the large and small subunits of GroE. Molecular weight standards were run in an adjacent lane.

terminal 170 amino acids. Expression of full-length cyclin A2 has been achieved by infecting recombinant baculovirus containing the full-length cyclin A2 cDNA under the control of the polyhedrin promoter into insect (*Sf*9) cells. It is possible that the poorly ordered N-terminus of cyclin A is recognised by a mechanism in *E. coli* which recognises misfolded proteins and targets their degradation. Such a mechanism, in the presence of an incorrectly or partially folded protein, could lead to poor yields upon purification. Similar strategies involving the deletion of the N-termini of B-type and other cyclins has not met with the same success as for cyclin A2, and so it seems that this approach is unique for cyclin A2.

Unfortunately the cyclin A3 which was eluted from the Ni^{2+}-NTA agarose column was unstable and when the eluates were stored at 4°C a visible precipitate rapidly formed in as little as one hour when the protein concentrations were high (15 mg/ml). The precipitation could be slowed if the protein was diluted considerably (0.15 mg/ml). It is possible that this precipitation was due to the aggregation of the protein which could have been caused by the exposed surfaces of cyclin A where CDKs normally interact with one another. These large hydrophobic surfaces would normally be buried when cyclin is complexed with CDK and it is energetically unfavourable for these to remain solvent exposed (Russo *et al.*, 1996). The interaction of these hydrophobic surfaces with the exclusion of water may be more energetically favourable and could cause aggregation which could eventually lead to precipitation.

Further to these *in vitro* precipitation problems, if expression of the protein is performed for greater than 3 hours at 30°C, or for more that 30 minutes at 37°C, yields were seen to be greatly reduced. This was at least in part due to partitioning of cyclin A3 into an insoluble fraction in the bacteria and further precipitation during purification. These observations suggested that cyclin A3 at high concentrations aggregates and precipitates. To eliminate aggregation and precipitation of cyclin A3 following purification on $Ni2^{+}$-NTA agarose it was essential to further purify the protein using gel filtration.

4.1.3. Purification of monomeric cyclin A3 by gel filtration.

I wanted to use cyclin A3 to make an affinity chromatography resin and this required that ideally the ligand be monomeric as some protocols require that the coupling occur at 4°C for up to 12 hours. Were I to use the material derived from the $Ni2^{+}$-agarose column this would have resulted in the precipitation of cyclin A3 during coupling. I used gel filtration

to purify monomeric cyclin A3 from the heterogeneous population which eluted from the $Ni2^{+}$-agarose column. To begin, I pooled the cyclin A3 containing fractions from the $Ni2^{+}$-agarose column and concentrated them by ultrafiltration to approximately 15 mg/ml using a Centricon device. During the concentration step however some of the cyclin A3 further aggregated and precipitated. Two ml of concentrated cyclin A3 was then loaded onto a 36 cm long x 1.5 cm wide (64 ml packed volume) Aca 34 gel filtration column. The column was run at 0.1 ml/min overnight at 4°C and 64 1ml fractions were collected.

A small aliquot of each fraction was analysed by 15% SDS-PAGE and the resulting gel was stained using Coomassie brilliant blue (figure 4-3). Fraction 1 was defined as the end of the excluded volume and as the start of the included volume. Aggregates of cyclin A3 appeared early in the included volume (fraction 1) and continued to elute through to fraction 15. A peak of material absorbing at A_{280} and containing a protein corresponding in molecular weight to that of cyclin A3 eluted as a broad peak across approximately 12 fractions beginning in fraction 24. A sharp peak of material eluting with a long retention time in the column contained no protein and this is likely to correspond to imidazole which was present in the elutions from the $Ni2^{+}$-agarose column, and which is known to absorb at 280 nm.

A greater amount of protein corresponding to cyclin A3 monomer eluted from the column than did protein corresponding to aggregated cyclin A3. This suggests that this small amount of aggregated cyclin A3 is responsible for causing further aggregation of the monomeric population eventually leading to the precipitation of the protein. Cyclin A3 preparations obtained from bacteria which were allowed to express for greater than 3 hours, or at temperatures above 37°C had significantly larger populations of aggregated cyclin A3 than monomeric cyclin A3 when analysed by gel filtration. This suggests that the aggregation and precipitation of cyclin A3 occurs as a function of expression level and concentration. Monomeric cyclin A3 was stable for at least one month and did not precipitate when stored at 4°C.

Cyclin A3 was analysed by gel filtration through a Sephadex-200 column using an AKTA FPLC. The elution profile of cyclin A3 (figure 4-4B) was compared with the elution profile of known gel filtration size standards (figure 4-4A). Cyclin A3 eluted with a Stokes radius corresponding to a molecular weight of approximately 18-20 kDa. This is somewhat smaller than the predicted molecular weight and than the apparent molecular

Figure 4-3. Cyclin A3 elutes as multiple molecular weight species on gel filtration by AcA 34.

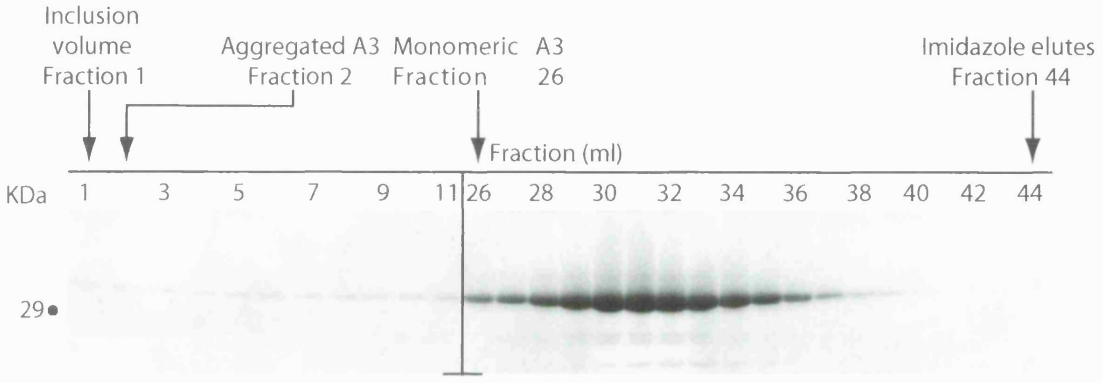
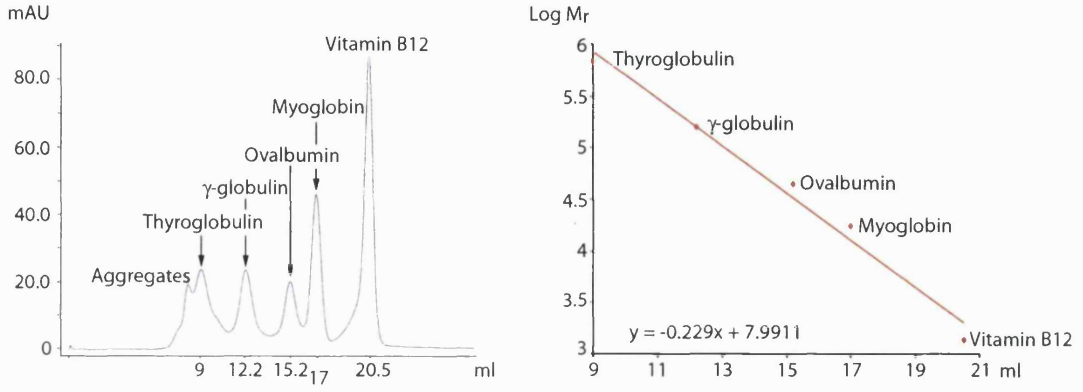


Figure 4-3. Coomassie blue stained SDS-polyacrylamide gel of fractions from a 64 ml (length 36 cm, radius 0.75 cm) AcA 34 gel filtration column loaded with cyclin A3. Hexahistidine tagged cyclin A3 was purifed using Ni^{2+}-chelate affinity chromatography and fractions containing cyclin A3 were pooled and concentrated to roughly 15 mg/ml. Two ml were loaded onto an AcA 34 gel filtration column and 1 ml fractions were collected and analysed by SDS-PAGE. The excluded volume is defined as ending at fraction 1 and high molecular weight species of cyclin A3 appear immediately. The major peak of A_{280} absorbing material contained cyclin A3 which corresponds to a molecular mass slightly lower than that of cyclin A3. A sharp peak of material absorbing at A_{280} appeared starting in fraction 44 and this is presumably imidazole.

Figure 4-4. Purification of monomeric cyclin A3 using gel filtration through Sephadex-200.

A. Chromatogram and plot of log M_r vs retention volume of size standards.



B. Chromatograms of purified cyclin A3.

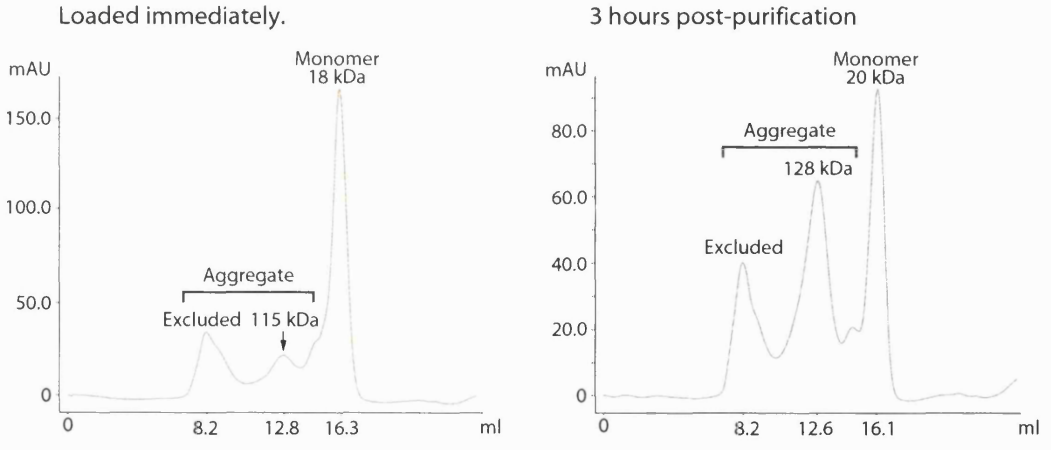


Figure 4-4. Chromatograms of size standards and cyclin A3 resolved using gel filtration on an AKTA FPLC. **(A) Left panel:** BIO-RAD gel filtration standards were separated on a Sephadex-200 gel filtration column. **Right panel:** The logs of the molecular weights of the standards were plotted against their retention volumes. Regression analysis yielded an equation describing the slope of the line of best fit which was used to calculate the molecular weight of cyclin A3 based on its retention volume. **(B) Left panel:** Cyclin A3 was expressed for 3 hours in *E. coli* and purified as described before and 3 mg of cyclin A3 was immediately loaded onto the column in a volume of 500 µl . Cyclin A3 eluted with a retention volume calculated to contain a protein of 18 kDa. Peaks of protein also appeared at approximately 115 kDa and in the excluded volume of the column. **Right panel:** 3mg of cyclin A3 which had been left at 4°C for three hours after the Ni^{2+}-NTA agarose purification was separated using the same column.

weight seen by SDS-PAGE. There were two populations of cyclin A3 which appeared to be aggregated species. One population eluted with a retention volume of approximately 12.8 ml which corresponds to a calculated molecular weight of 115 kDa. Given the molecular weight of the monomeric cyclin A3 it is possible that this aggregate is hexameric. A second population eluted in the excluded volume. This elutes in the excluded volume of the column. Analysis by gel filtration of cyclin A3 that had been stored at 4°C for 3 hours post Ni^{2+}-NTA agarose purification showed a decrease in the amount of cyclin A3 monomers and a concomitant increase in the amount of aggregated material present (figure 4-4B). Similar data were obtained when cyclin A3 purified from cells expressing the protein for longer than 3 hours was analysed by gel filtration (not shown).

4.2. Defining an appropriate affinity chromatography matrix.

Once I had prepared the protein to be used as ligand in the affinity chromatography screen, the next challenge was to identify a suitable matrix on which I could immobilise cyclin A3. Resins which are going to be used in affinity chromatography experiments should have very low non-specific protein binding capacities while being able to be specifically coupled to the ligand of interest. Initially I hoped to re-associate the hexahistidine tagged cyclin A3 with Ni^{2+}-NTA agarose and use this as an affinity chromatography reagent to identify protein interactors using the small-pool expression screening technique.

4.2.1. Cyclin A3 Ni^{2+}-NTA agarose is a poor affinity chromatography reagent for use in eukaryotic cell lysates.

I hoped to be able to re-bind the cyclin A3 to Ni^{2+}-NTA agarose and to use this an affinity chromatography reagent. To test this I transcribed and translated proteins known to interact with cyclin A3 (CDK1, CDK2 and p27) *in vitro* in the presence of [^{35}S]-methionine using the rabbit reticulocyte lysate system. These reactions were diluted in 200 μl buffer containing 150 mM NaCl and 0.05% Nonidet P-40 to which 5 μg of monomeric cyclin A3 was added. Following an incubation on ice for one hour the cyclin A3 was harvested using 10 μl of Ni^{2+}-NTA agarose and this resin was washed extensively in buffer containing 250 mM NaCl and 0.05% Nonidet P-40. Proteins bound specifically were eluted by boiling the resin in SDS-sample buffer (data not shown). Although the capture of cyclin A3 and

CDK1 or CDK2 or p27 was efficient, the coomassie blue stained gel of the proteins captured on the resin showed that a large number of proteins from the reticulocyte lysate bound to the Ni^{2+}-NTA agarose even after extensive washing. In particular it was never possible to remove the haemoglobins from the resin. It has been observed that Ni^{2+}-NTA agarose binds to many eukaryotic proteins when used with lysates of eukaryotic cells, due to the greater incidence of hexahistidine tracts in eukaryotic proteins relative to those *of E. coli*. It would seem therefore that Ni^{2+}-NTA agarose is a poor reagent for affinity chromatography when the target proteins are in a lysate of eukaryotic cells.

4.2.2. Biotinylation of cyclin A3.

I next explored the possibility of using a biotinylated form of cyclin A3 which could be immobilised on streptavidin or avidin agarose and used as an affinity chromatography resin. Historically the biotinylation of proteins has been performed chemically which causes the non-specific modification of lysine residues in the target protein. I was concerned that such an approach would generate a protein where the surface lysines would become modified and as such block sites of potential protein-protein interactions. Sequences derived from *in vivo* biotinylated proteins have been used to tag recombinant proteins which will become biotinylated in the presence of a biotin ligase and biotin, either *in vivo* or *in vitro* (Chapman-Smith *et al.*, 1994; Leon-Del-Rio and Gravel, 1994; Yamano *et al.*, 1992). Shorter sequences have been defined using phage display technology which can be specifically and efficiently biotinylated *in vitro* and *in vivo* (Stolz *et al.*, 1998). The gene encoding the enzyme which catalyses the biotinylation of target sequences in *E. coli* is birA (biotin holoenzyme synthetase or biotin ligase) which is expressed at low levels *in E. coli* (Chapman-Smith and Cronan, 1999). The high levels of expression achieved using the pET system seems to be especially prone to under-biotinylation of the target protein.

It has been reported that proteins containing a biotinylation signal which are expressed at low to moderate levels in cells expressing endogenous levels of the birA gene are modified at 10-30% efficiency. Tagged proteins expressed at high levels in the presence of endogenous levels of birA are biotinylated with as little as 2-6% efficiency (Chapman-Smith and Cronan, 1999). This problem can be overcome by overexpressing biotin ligase in the cells expressing the target protein to be biotinylated. Due to the

relatively high levels of biotinylation obtained it is also necessary to include biotin in the growth media.

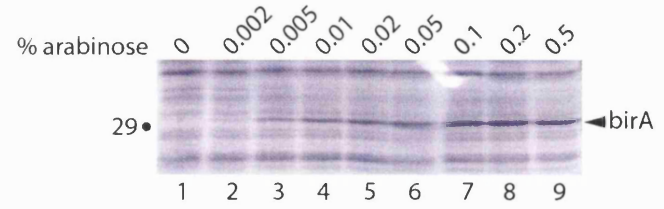
To introduce a biotinylatable sequence into cyclin A3 I cloned a linker coding for the sequence ENLYFQGMSGLNDIFEAQKIEWHIE between the cyclin A3 coding sequence and the hexahistidine tag. The sequence in blue is the hepta-amino acid sequence recognised by the Tobacco Etch Virus protease, which could be used to cleave tags from the cyclin A3, and consequently from the supporting matrix (Parks *et al.*, 1995). The sequence in red is the minimal sequence reported to act as an efficient biotin accepting substrate in *E. coli* (Stolz *et al.*, 1998). The plasmid pCY216, a construct containing *E. coli* biotin ligase under the control of the arabinose promoter, the origin of replication from plasmid p15A, which is compatible with the pBR322 derived origin of replication present in the pET series of vectors, and the gene for chloramphenicol acetyl transferase, was a kind gift of Dr. John Cronan (University of Illinois, Urbana, IL, USA).

When the biotinylatable cyclin A3 (cyclin A3-biotin) was expressed in the absence of biotin ligase I found that there was a very low level of biotinylation. The biotinylation of cyclin A3 is easily detected by a mobility shift in SDS-PAGE, and that this shift was caused by biotinylation was shown by far-western blotting with streptavidin-HRP (figure 4-5). When biotin ligase was co-expressed with cyclin A3 and when biotin was added to the growth media however, cyclin A3 became completely biotinylated. (figure 4-5). Unusually however, cyclin A3-biotin was expressed comparably well in the absence or presence of IPTG. BCCP, a 17 kDa protein, is the single protein in *E. coli* which is biotinylated, and this is easily distinguished from cyclin A3-biotin which migrates with an apparent molecular weight of 31 kDa.

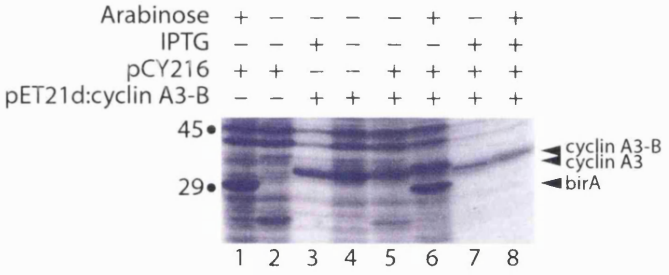
The initial experiments showing the efficient biotinylation of cyclin A3 were encouraging but when I attempted to purify the biotinylated cyclin A3 from *E. coli* cell lysates the yields were very low, in the region of 0.1 mg/L of culture and most of the cyclin A3 was insoluble. Reducing the temperature at which the culture was grown did not improve solubility and it appeared that the cyclin A3-biotin would not be amenable to purification in a native form.

Figure 4-5. Expression and *in vivo* biotinylation of cyclin A3.

A. Expression of birA in *E. coli*.



B. *in vivo* biotinylation of cyclin A3.



C. Western and far-western blots of *in vivo* biotinylation of cyclin A3.

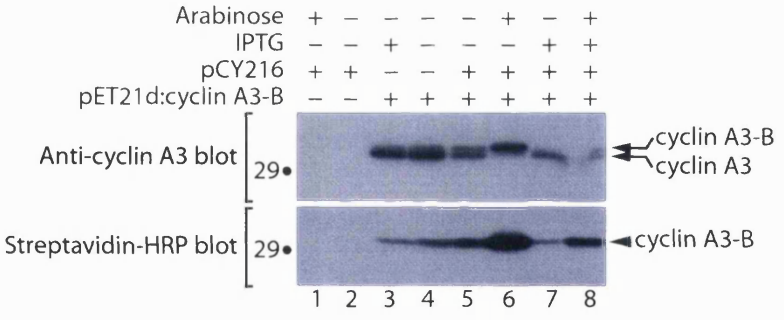


Figure 4-5. A biotinylatable form of cyclin A3 can be efficiently biotinylated *in vivo* by co-expressing birA. **(A)** Expression of birA in nine cultures of *E. coli* containing pCY216 were induced by the addition of increasing amounts of arabinose to the culture media. Equivalent amounts of cells were harvested and lysed by boiling in SDS-sample buffer. Equal amounts of each lysate were analsyed by SDS-PAGE. **(B)** *E. coli* transformed with pCY216 alone (lanes 1-2), pET21d:cyclin A3-B alone (lanes 3-4) or both plasmids (lanes 5-8) were grown in cultures and treated with either 0.5% arabinose to induce expression of birA (lanes 1, 6 and 8), 1 mM IPTG to induce expression of cyclin A3-B (lanes 3 and 7-8) or both arabinose and IPTG (lanes 1, 3 and 7-8. Cyclin A3-B was slightly larger than the WT form and birA has an apparent molecular weight of 29 kDa by SDS-PAGE. **(C) Upper panel:** Samples corresponding to 1/100th of those shown in (B) were separated by SDS-PAGE and transferred to nitrocellulose membranes. The extent of biotinylation was determined by detecting cyclin A3 with different mobilities through SDS-PAGE by western blotting using a monoclonal antibody (E23). **Lower panel:** Biotinylated proteins were detected on a duplicate blot using streptavidin-HRP.

4.2.3. Cyclin A3-Affigel resin is a good affinity chromatography reagent.

I next decided to try a more conventional approach to generate an affinity chromatography resin. Any matrix used for affinity chromatography should be inherently inert, have a low non-specific protein binding capacity once coupled with ligand, have a high-capacity for ligand binding and not leach coupled ligand. Affigel resins fulfil these requirements and are available in forms which allow coupling to be performed at physiological pH.

Affigel resins are activated with N-hydroxysuccinimide and coupling is achieved using primary amine groups in the ligand via stable amide bonds, therefore the buffer containing the ligand must be free of any primary amine containing buffer such as Tris. I removed the 50 mM Tris present in the preparation of monomeric cyclin A3 by passage through a G25-Sephadex column equilibrated in PBS. Affigel 10 is suitable for coupling proteins with an isoelectric point in the range of 6.5-11, and maximal coupling is achieved at pHs close to or below their isoelectric point. Affigel 15 however couples proteins with isoelectric points in the range 3-6.5 at pHs at or above their isoelectric points. The difference in efficiency of coupling to Affigel 10 and Affigel 15 for acidic and basic proteins is attributed to the difference between the charge on the protein and the charge on the gel. Hydrolysis of some of the active esters during aqueous coupling imparts a slight negative charge to the gel and this will attract basic proteins (those buffered at pHs below their isoelectric point) thus increasing their coupling efficiency. Conversely, acidic proteins (those buffered at pHs above their isolectric point) will be repulsed by the negative charge and coupling efficiency will be lower. Affigel 15 incorporates a tertiary amine into the spacer arm which imparts a slight positive charge and therefore the effects on basic and acidic proteins are reversed with basic proteins coupling poorly and acidic proteins coupling well. The predicted isoelectric point of bovine cyclin A3 is 6.57, and this would theoretically couple equally well to Affigel 10 or 15.

4.2.4. Cyclin A3-Affigel can pull-down CDK2.

I coupled monomeric cyclin A3 to Affigel 15 at a concentration of 5 mg per ml resin, and blocked any unreacted coupling sites using 1 M Tris-Cl pH 7.4 The efficiency of coupling was tested by pelleting the resin and using a small amount of the supernatant from the coupling reaction in a Bradford assay. This showed that there was undetectable amounts of protein in the supernatant and suggested that the efficiency of coupling was close to 100%.

This resin did not precipitate significant quantities of any component in the rabbit reticulocyte lysate (not shown), although it was able to precipitate *in vitro* transcribed-translated [^{35}S]-methionine labelled CDK2 (figure 4-6). As described before CDK2 was transcribed and translated in the presence of [^{35}S]-methionine and this was diluted to 200 µl in a buffer containing 150 mM NaCl and 0.05% Nonidet P-40. To the diluted reaction increasing amounts of cyclin A3-Affigel were added, and this was incubated at 4°C for 1 hour.

The resin was pelleted by gentle centrifugation and washed in buffer containing 250 mM NaCl and 0.05% Nonidet P-40 before specifically bound proteins were eluted by boiling in SDS-sample buffer. The labelled CDK2 did not bind blank resin which had been blocked with Tris-Cl pH 7.4 prior to use, suggesting that the capture of CDK2 by cyclin A3 resin is a specific characteristic of cyclin A3, and not due to the resin. As the amount of cyclin A3-Affigel used in the pull-down was increased so did the amount of CDK2 recovered. A similar experiment using transcribed and translated casein kinase II showed that the interaction between cyclin A3 and CDK2 in this assay is specific. CKII did not bind to the cyclin A coupled Affigel. It appeared that only approximately 20% of the CDK2 input into the pull-down was recovered by the cyclin A3 resin.

It is not especially surprising that the capture is not more efficient as although the interaction between CDK2 and cyclin A3 is a particularly strong one, the coupling of cyclin A3 to the Affigel is via any primary amine resulting in the random orientation of molecules on the matrix. This means that many of the cyclin A molecules could be oriented inappropriately to interact with CDK2. This could be beneficial however when screening for proteins whose interaction with cyclin is not characterised. By presenting the cyclin to potential interactors in many orientations, the possibility of identifying proteins is greater than if the cyclin were presented in a uniform fashion on the surfaces of a resin (as would be the case for coupling via hexahistidine tags.), as this orientation may be favourable for binding to CDK2 for example, but not to other unknown proteins.

Another explanation for the low levels of binding of CDK2 to cyclin A3 Affigel could be explained by the poor folding of CDK2 when it is transcribed and translated in rabbit reticulocyte lysate. It is possible that the capture of CDK2 by the cyclin A3 Affigel is highly efficient but that only 20% of the translated protein is folded appropriately and

that the remaining 80% is inappropriately folded and cannot interact with cyclin A3. Reticulocytes are a highly specialised cell type, with globins being the predominant protein synthesised. It is possible therefore that they are not particularly competent to fold other exogenous' proteins, and the folding of some proteins has been reported to be enhanced significantly if a 'cell extract' from the same source as the cDNA encoding the translated protein is added to the translation reaction. If only a small amount of each protein present in a pool were correctly folded then only a small percentage of the total translated material would be available to interact with cyclin A3.

To further establish that the cyclin A3-Affigel could capture known binding partners of cyclin A I used the resin in affinity chromatography experiments with extracts of human cells. Extracts of HeLa S3 (suspension) cells were prepared by lysing cells which had been pelleted and washed twice in PBS. Cells from a 1 L culture were resuspended in 4 ml buffer containing 50 mM HEPES-KOH pH 7.4, 150 mM KCl, 0.1% Nonidet P-40 and protease inhibitors. The cells were then ruptured using 20 strokes of a Dounce homogeniser. Cellular debris was pelleted at 100000 g and the supernatant was retained as S100 extract. Ten µl of cyclin A3-affigel (corresponding to 50 µg cyclin A3) was incubated batch-wise with 15 mg total protein (0.5 ml of extract) and after extensive washing any bound proteins were eluted by boiling in SDS-sample buffer. Western blotting showed that both p21 and p27, although undetectable in the S100 extract were captured and concentrated by the cyclin A3-Affigel. CDK2 and CDK1 were also efficiently captured by the resin. We were not in possession of an anti-HsCdc6 antibody so were unable to assess the efficiency of capture of this protein, although antibodies against p107 and pRb were unable to detect p107 or pRb in the cyclin A3-affigel eluates.

4.3. Cyclin A3:CDK3 complexes may be more suitable reagents for affinity chromatography.

Although the cyclin A3 resin appeared as though it would be an adequate reagent to begin screening, some proteins (such as p27) interact not only with cyclin, but also extensively with the cognate CDK partner. It occurred to me at this point that to maximise the possibility of identifying proteins that were substrates of, or regulators of CDK activity it would be prudent to use a complex of cyclin A3 and CDK2. A construct was donated to me by Jane Endicott (University of Oxford) from which GST-CDK2 and the *S. cerevisiae*

CDK-activating kinase ortholog Civ1 could be co-expressed in *E. coli*. Co-expression of both proteins in *E. coli* generated threonine 160 phosphorylated CDK2 (Brown *et al.*, 1999). I prepared lysates of *E. coli* expressing cyclin A3 alone and GST-CDK2 and Civ1 and mixed them together. From this I was able, through a simple three step purification, to obtain essentially pure, stoichiometric complexes of CDK2:cyclin A (figure 4-7). The mixed lysates were incubated with Ni^{2+}-NTA agarose for one hour and the resin was extensively washed with imidazole increasing from 10-50 mM.

Due to the strong interaction between CDKs and cyclins; CDK2 efficiently co-purified with cyclin A3. Unfortunately, when these proteins were eluted from the Ni^{2+}-NTA agarose using 150 mM imidazole they were unstable and rapidly precipitated. Initially I thought that this was due to an excess of cyclin A3 in the preparations, but when the cyclin A component was made to be limiting the precipitation still occurred. It was possible that the precipitation was due to an excess of imidazole, and removing it would prevent the precipitation. Initially, I dialysed the complexes against a buffer lacking imidazole, but as the buffer exchange could take several hours precipitation inevitably occurred. An alternative way to remove the imidazole was by gel filtration. Indeed, after passage through a G25-Sephadex column there was no precipitation, and the complexes were stabilised. Cyclin A3 expressed significantly better than GST-CDK2 and so by purifying the complexes on GSH-Sepharose I obtained stoichiometric complexes of GST-CDK2:cyclin A3. Typical yields from one litre cultures expressing cyclin A3 or GST-CDK2 and Civ1 were 3 mg.

4.3.2. Cyclin A3:CDK2 complexes co-purified from *E. coli* in the presence of *S. cerevisiae* Civ1 are as active as those purified from *Sf*9 cells.

To check if the GST-CDK2:cyclin A3 complexes purified from *E. coli* were as active as those obtained from baculovirus infected insect cells I obtained small samples of CDK2:cyclin A and CDK7:cyclin H (CAK) which had been purified from insect cells. (a kind gift of Jane Endicott, University of Oxford). Figure 4-8 shows a comparison of histone H1 kinase activities of the complexes purified from the different expression systems. The enzymes were assayed on an equimolar basis with respect to the CDK2 moiety. As has been previously reported monomeric CDK2 alone had little histone H1 kinase activity irrespective of its phosphorylation state and CDK2 complexed with cyclin

Figure 4-8. Active complexes of CDK2:cyclin A3 can be purified from *E. coli* using a co-expression strategy.

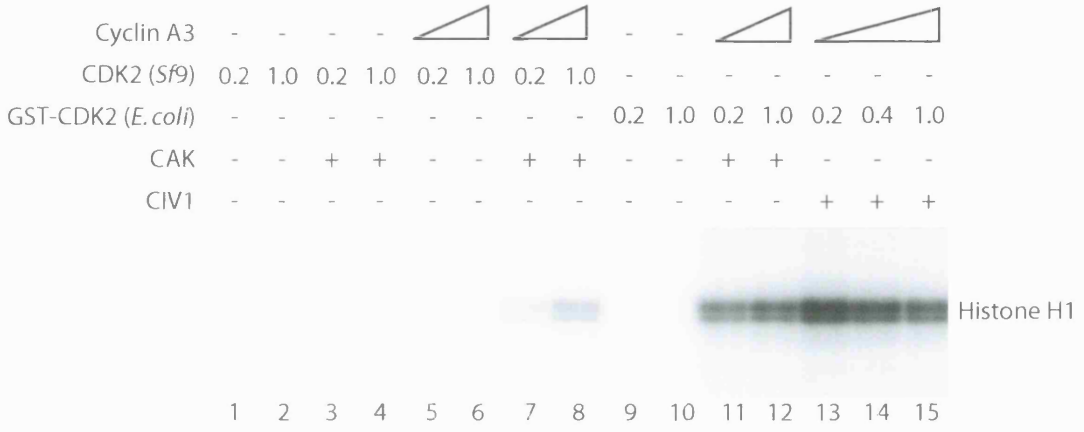


Figure 4-8. Histone H1 kinase assays using recombinant CDKs and CDK:cyclin complexes purified from either insect (*Sf9*) cells or bacteria (*E. coli*). Kinase assays using 50 ng/ul histone H1 and 400 μM ATP containing 50 nCi/μl $\gamma[^{32}P]$-ATP as substrates were performed for 15 minutes at 30°C using CDK2 derived from *Sf9* cells in the absence (lanes 1 and 2) or presence of CDK7:cyclin H (lanes 3 and 4), cyclin A3 (lanes 5 and 6), or both cyclin A3 and CDK7:cyclin H (lanes 7 and 8). For comparison the histone H1 kinase activity of GST-CDK2 purified from *E. coli* (lanes 9 and 10), GST-CDK2 co-purified with cyclin A3 in the presence of CDK7:cyclin H (lanes 11 and 12), or GST-CDK2 co-expressed with CIV1 and co-purified with cyclin A3 (lanes 13-15) was also tested. Equimolar amounts of kinase with respect to the CDK2 moeity were used.

A, but not phosphorylated on threonine 160 by CAK, also had minimal activity. Activation of non-phosphorylated CDK2 by incubation with CDK7:cyclin H purified from baculovirus infected *Sf*9 cells showed greatly stimulated kinase activity. Interestingly however, the *E. coli* derived GST-CDK2 which was phosphorylated *in vitro* by CDK7:cyclin H had greater histone H1 kinase activity than the complexes purified from insect cells and fully activated with CAK. CDK2 derived from *Sf*9 cells is possibly phosphorylated on residues threonine 14 and tyrosine 15, by a kinase activity endogenous to the insect cells. These inhibitory phosphorylations could not occur in *E. coli* however, and this possibly explains the apparent higher specific activity of complexes derived from *E. coli* relative to those from *Sf*9 cells. The CDK2 phosphorylated *in vivo* by co-expression of Civ1, and complexed with cyclin A3, had a higher specific activity than complexes activated *in vitro* using CDK7:cyclin H.

4.3.3. Cyclin A3:CDK2 complexes efficiently bind to p27, but not to the substrate pRb, although pRb is efficiently phosphorylated by the enzyme.

As I wanted to screen as many protein pools as possible for proteins which interacted with CDK2:cyclin A3 I wanted to develop an efficient method for capturing the GST-CDK2:cyclin A3 and any labelled interacting proteins. I experimented with performing the resin harvesting by centrifugation, but found that it was easy to accidentally remove the GSH-Sepharose during the wash steps when processing multiple samples. I devised a micro-column format, where siliconised P200 tips were plugged with a small amount of glass wool, allowing a small amount (10 µl) of GSH-Sepharose to be retained in the tip.

The diluted transcription-translation reactions to which GST-CDK2:cyclin A3 had been added were allowed to pass through the tips over the GSH-Sepharose. The GST-fusion protein was captured along with cyclin A3 and any interacting proteins by the GSH-Sepharose which was retained in the tip by the glass wool. The tips were allowed to empty by gravity and the siliconisation of the tips meant that the tip emptied in 4-5 minutes. The micro-columns were then washed three times. To harvest the GSH-Sepharose and bound proteins, 1 cm from the tapered end of the tip was cut off and placed wide end down in a 0.5 ml Eppendorf tube and a three second pulse in a benchtop microcentrifuge expelled the resin from the tip into the bottom of the tube. Bound proteins were eluted by boiling in SDS-sample buffer. It was a relatively simple and quick procedure to prepare 96 such

Figure 4-9. Comparison of the binding of p27 and pRB to GST-CDK2:cyclin A3.

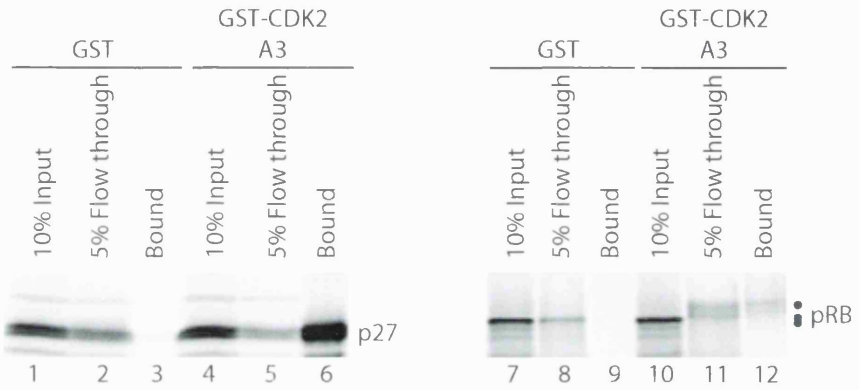


Figure 4-9. Autoradiogram of SDS-PAGE of *in vitro* transcribed and translated p27 or pRB captured with GST or GST-CDK2:cyclin A3 complexes. Duplicate 10 µl transcription and translation reactions of p27 and pRB were perfomed in the presence of [^{35}S]-methionine. One µl of labelled p27 and pRB retained as 10% of the input (lanes 1 and 7 respectively). The remaining 9 µl of reaction products were diluted and incubated in the presence of bacterially produced GST or GST-CDK2:cyclin A3. The bacterially produced proteins and any interacting proteins were harvested withGSH-Sepharose. GST-CDK2:cyclin A efficiently captured p27 (lane 6) but GST alone did not (lane 3). GST-CDK2:cyclin A3 captured pRb with lower efficiency (lane 12) relative to p27, but did not interact with GST alone.

Figure 4-10. AMP-PNP does not affect the kinase activity of GST-CDK2:cyclin A3 towards, or enhance its interaction with, pRB.

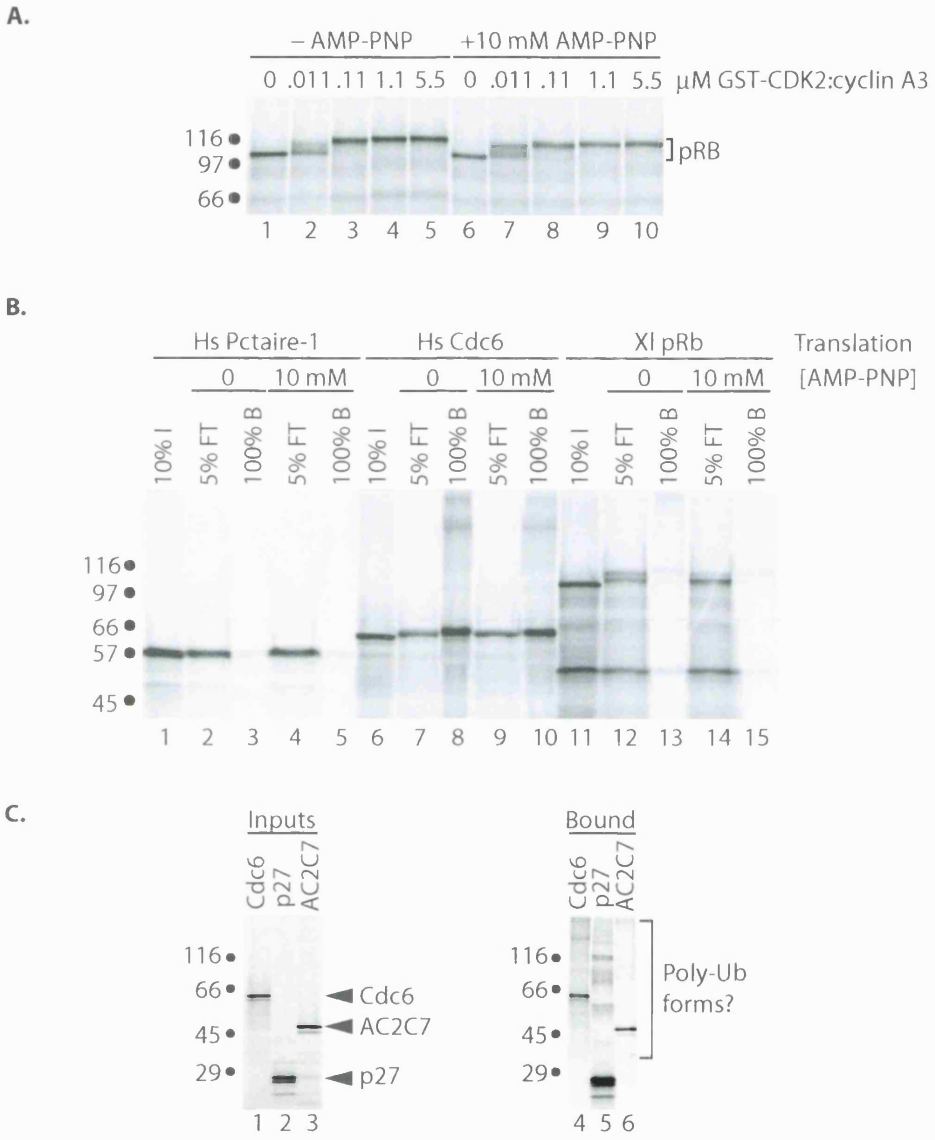


Figure 4-10. Analysis of the effects of AMP-PNP on the kinase activity of GST-CDK2:cyclin A3 for *in vitro* transcribed and translated pRb and of AMP-PNPs effects on the affinity of GST-CDK2:cyclin A3 for Pctaire-1, Cdc6 and pRb. **(A)** To test the effects of AMP-PNP on the kinase activity of GST-CDK2:cyclin A3, pRb was transcribed and translated in the presence of [^{35}S]-methionine, and 1 μl of the reaction was used in kinase reactions containing increasing amounts of kinase in the absence (lanes 1-5) or presence of 10 mM AMP-PNP (lanes 6-10). **(B)** Pctaire-1, Cdc6 and pRb were transcribed and translated in the presence of [^{35}S]-methionine (lanes 1, 6 and 11 respectively). Pctaire, Cdc6 or pRb translation reactions were diluted to 200 μl in the absence (lanes 2-3, 7-8 and 12-13 respectively) or presence of 10 mM AMP-PNP (lanes 4-5, 9-10 and 14-15 respectively), and incubated in the presence of 5 μg GST-CDK2:cyclin A3. GST-fusion proteins were harvested with GSH-Sepharose, after which the resin was washed. Finally, specifically bound proteins were eluted by boiling in SDS-sample buffer. (**C**) Cdc6, p27 and clone AC2C7 were captured using GST CDK2:cyclin A3 and GSH-Sepharose. Interacting proteins were analysed by SDS-PAGE and fluorography. Lanes 1-3 show the 10% of the input to the pull-down assay, and lanes 4-6 show the material captured by GST-CDK2:cyclin A3 including the high molecular weight forms. Abbreviations: I=Input, FT=Flow Through, B=Bound.

at inhibiting phosphorylation of substrate relative to identical reactions lacking AMP-PNP. This assay also suggested that the minimal amount of kinase that could induce phosphorylation of approximately 50% of a translated substrate was in the region of 75 nM. To ensure that kinase reactions could proceed to completion I planned to use 750 nM CDK2:cyclin A3 for the small-pool expression screen.

4.4.2. An explanation for the lack of inhibition of CDK2 activity by AMP-PNP in this assay.

The reason for the phosphorylation of pRb in the presence of 100000 fold molar excess of AMP-PNP over kinase is unclear. The concentration of ATP in the reticulocyte lysate is approximately 0.5 mM, and as this is diluted 1:10, AMP-PNP is in 200 fold molar excess relative to ATP. Under these conditions it is predicted that AMP-PNP would be used preferentially and that this would inhibit the phosphorylation of pRb. It is conceivable however that there is sufficient ATP bound to the GST-CDK2 which permits the phosphorylation of the very small amounts of pRb that are used as substrate in the kinase reactions. For a construct which transcribes and translates with high efficiency (luciferase) it is possible to obtain up to 200 ng of translated product from a 50 µl coupled transcription and translation reaction using the commercially available Promega TNT® Quick Coupled Transcription/Translation system. If pRb were translated with at least comparable efficiency then this would equate to 4 ng/µl or 39 pM (390 atto mol total) of translated Xl pRb based on a predicted molecular weight of 103 kDa for Xl pRb.

At the lowest concentration of kinase used (11 nM), if 100% of CDK2 molecules had bound ATP which could be hydrolysed (also 11 nM or 110 femto mol total), then assuming negligible exchange of bound ATP in the absence of hydrolysis, this substrate ATP concentration (excluding that present in the reticulocyte lysate) would be in 282 fold molar excess relative to the protein substrate. In theory this could lead to the efficient phosphorylation of pRb that was seen. To observe an effect of AMP-PNP on the phosphorylation of pRb would require that the protein substrate be in molar excess of the ATP bound to CDK2 (11 nM or 110 femto mol total) at the lowest concentration of enzyme used, and this is generally in the region of 1 µM. Under these conditions of 10 mM AMP-PNP (100 nano mol total), 50.01 µM ATP (500.1 pico mol total), 1 µM pRb (10 pico mol total), and 10 nM kinase (100 femto mol total) it could be predicted that 100 femto mol of

pRb (or 1% of total) would be phosphorylated due to kinase bound ATP, and that approximately a further 553 atto mol of pRb (or 0.006% of total) could be phosphorylated before only approximately 41 molecules of CDK2 remained that were not AMP-PNP bound.

Alternatively, it is also possible that AMP-PNP cannot be bound by CDK2. In this case AMP-PNP would not be able to act as a competitive inhibitor. This could easily be tested by adding an excess of AMP-PNP to a standard histone H1 kinase assay in the presence of ATP and $\gamma[^{32}P]$-ATP. If AMP-PNP could not be bound by the enzyme then there would be no inhibition and incorporation of $\gamma[^{32}P]$ would be equivalent to reactions lacking AMP-PNP. When crystals of CDK2:cyclin A3 were soaked in AMP-PNP however, AMP-PNP bound in the active site of the kinase, mimicking ATP; so it seems likely that *in vitro*, GST-CDK2 would be able to bind AMP-PNP in kinase assays (Brown *et al.*, 1999).

4.4.3. AMP-PNP does not affect the affinity of GST-CDK2:cyclin A3 for pRb or Cdc6.

There appeared to be no difference in the efficiency of capture of either Cdc6 or pRb in the absence or presence of AMP-PNP. Interestingly, the pRb that was captured was predominantly the phosphorylated form in both the presence and absence of AMP-PNP, suggesting that possibly this form has a stronger affinity for CDK2:cyclin A3, which could cause end-product inhibition of the kinase. Affinity chromatography in the absence of AMP-PNP results in approximately 50% phosphorylation of pRb (slow-migrating form), of which only a small percentage is retained on the CDK2:cyclin A3 resin. During chromatography in the presence of AMP-PNP however, approximately only 10% of the pRb is in the phosphorylated form and yet still only a small percentage of the input is retained by CDK2:cyclin A3 (figure 4-10B). Therefore it seems as though AMP-PNP has no effect on the retention of substrate by CDK2:cyclin A3. As a negative control for non-specific binding of proteins to CDK2:cyclin A3 or the GSH-Sepharose resin, I included translated human PCTAIRE, which is a CDK-like kinase and which is known to not interact with, or be a substrate of CDK2 or cyclin A3.

4.4.4. Capture of Cdc6 by CDK2:cyclin A3 resin revealed that Cdc6 was possibly polyubiquitinylated in reticulocyte lysate.

CDK2:cyclin A3 binds CDC6 with greater affinity than it does pRb as evidenced by the retention of more than 20% of the CDC6 input products on the affinity resin. This could be attributed to the presence of a well conserved Cy-motif in the N-terminus of CDC6. It is also interesting to note from these experiments that there are a number of high molecular weight species which are precipitated along with CDC6 by the CDK2:cyclin A3 complexes (figure 4-10C). p27 and Cdc6 are synthesised predominantly as single species in coupled transcription and translation reactions, but when these products are captured by GST-CDK2:cyclin A3 and GSH-Sepharose a number of high molecular weight forms become visible. It is possible that these are polyubiquitinylated forms of p27 and Cdc6 and that they are present in the translation reactions, but at such low levels that they are not visualised by autoradiography and that they only become apparent when they are concentrated by affinity chromatography. Alternatively, it could be that these forms are generated as a result of the interaction with GST-CDK2:cyclin A3.

I attempted to show that the high molecular weight forms of p27 and Cdc6 captured by GST-CDK2:cyclin A3 were polyubiquitylinated by addition of [^{125}I]-labelled ubiquitin or hexahistidine tagged ubiquitin to the affinity chromatography reactions. I was unable to detect the incorporation of either of these forms of ubiquitin into the [^{35}S]-labelled p27 and Cdc6. Also, addition of N-ethylmalemide, which inhibits ubiquitination by covalently modifying the ubiquitin-accepting cysteine residue in E2 enzymes, to the affinity chromatography reactions seemed to have little effect on the appearance of the high molecular weight species. The use of a K48R ubiquitin mutant which permits the addition of single ubiquitin molecules to substrates by the E3 enzymes, but which prevents polyubiquitinylation, also seemed to have no effect on the appearance of the high molecular weight species of p27 and Cdc6 (data not shown). These experiments could be interpreted as meaning that the high molecular weight species of p27 and Cdc6 which bind to GST-CDK2:cyclin A3 are not due to ubiquitinylation, however it is also possible that there are sufficient amounts of ubiquitin-charged E2 present in the reticulocyte lysate. It could be argued that this charged E2 is sufficient to efficiently modify the low amounts of p27 and Cdc6 that are present in the translation reactions. Although I was not able to prove directly,

it is possible that these high molecular weight species are polyubiquitinylated forms of CDC6 which are only visible when CDC6 is concentrated by chromatography. Alternatively, the ubiquitinylation of CDC6 could be a direct result of a phosphorylation event catalysed by CDK2. Recent reports have confirmed that human CDC6 is in fact ubiquitinylated *in vivo* (Drury *et al.*, 1997; Petersen *et al.*, 2000; Sanchez *et al.*, 1999), and that this leads to its proteolysis.

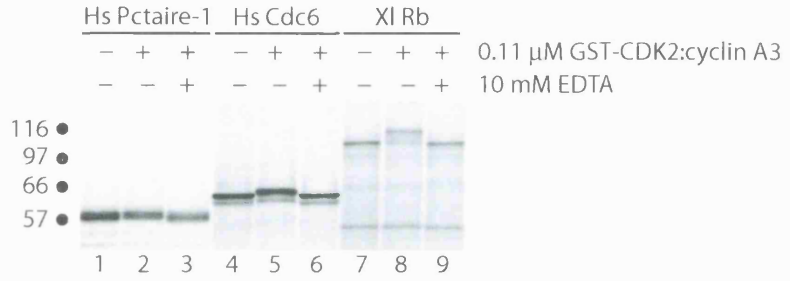
4.4.5. EDTA does not affect the affinity of GST-CDK2:cyclin A3 for pRb or Cdc6.

Since AMP-PNP had no detectable effect on the affinities of Cdc6 or pRb for GST-CDK2:cyclin A3 I next tested the Mg^{2+}-chelating agent EDTA to try and achieve an increase in affinity of these substrates for GST-CDK2:cyclin A3. EDTA chelates the essential Mg^{2+} ions that CDK2 requires for efficient hydrolysis of ATP and should cause a complete loss of kinase activity. The experiments were performed exactly as described before except that the 10 mM AMP-PNP was replaced with 10 mM EDTA. The concentration of Mg^{2+} in reticulocyte lysates is 1.6-1.8 mM, although there are other compounds present in the lysates which can act as chelators of Mg^{2+} and the actual concentration may vary slightly.

Figure 4-11A shows that as expected the inclusion of 10 mM EDTA into the kinase reactions results in complete loss of CDK2 activity as determined by the loss of electrophoretic retardation of Cdc6 and pRb in the presence of kinase. The same effects would be expected were the kinase reactions performed using a D145R mutant of CDK2. Such a mutant cannot co-ordinate the active site Mg^{2+} ion and would be unable to catalyse phosphotransfer. Unfortunately I observed no significant increase in the affinity of Cdc6 or pRb for GST-CDK2:cyclin A3 when the affinity chromatography was performed in the presence versus the absence of 10 mM EDTA (figure 4-11B). These observations suggest that a 'substrate-trapping' approach such as that which has been useful in the identification of substrates of phosphatases may not be successful when applied to CDK2:cyclin A3, although an approach where essential catalytic residues such as D145 or K33 are mutated may still be worth exploring.

Figure 4-11. EDTA eliminates the kinase activity of GST-CDK2:cyclin A3 towards, but doesn't affect its affinity for, Cdc6 or pRb.

A. Kinase reactions in the presence of 10 mM EDTA.



B. Affinity chromatography reactions in the presence of 10 mM EDTA.

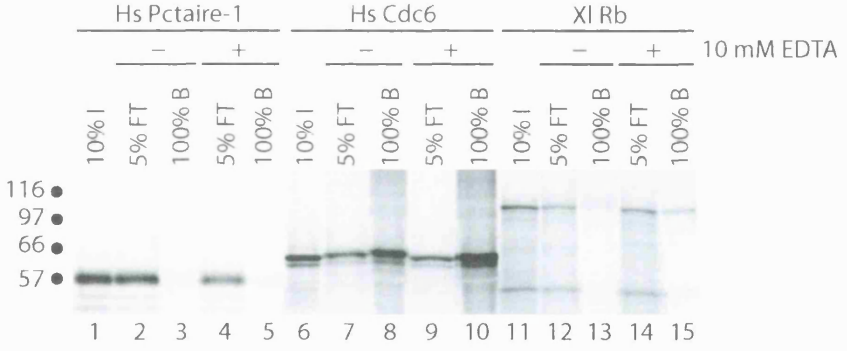


Figure 4-11. Analysis of the effects of EDTA on the kinase activity of GST-CDK2:cyclin A3 towards, and affinity for, *in vitro* transcribed and translated Pctaire-1, Cdc6 or pRb. **(A)** To test the effects of EDTA on the kinase activity of GST-CDK2:cyclin A3, Pctaire-1, Cdc6 and pRb were transcribed and translated in the presence of [^{35}S]-methionine (lanes 1, 4 and 7). One µl of the translated Pctaire-1, Cdc6 and pRb were used in kinase reactions containing 0.11 µM GST-CDK2:cyclin A3 in the absence (lanes 2, 5 and 8 respectively) or presence of 10 mM EDTA (lanes 3, 6 and 9 respectively). **(B)** Pctaire-1, Cdc6 and pRb were transcribed and translated in the presence of [^{35}S]-methionine (lanes 1, 6 and 11 respectively). Pctaire-1, Cdc6 or pRb translation reactions were diluted to 200 µl in the absence (lanes 2-3, 7-8 and 12-13 respectively) or presence of 10 mM EDTA (lanes 4-5, 9-10 and 14-15 respectively), and incubated in the presence of 0.275 µM GST-CDK2:cyclin A3. GST-fusion proteins were harvested with GSH-Sepharose, after which the resin was washed. Finally, specifically bound proteins were eluted by boiling in SDS-sample buffer. Abbreviations: I=Input, FT=Flow Through, B=Bound.

kinase Civ1 (Thuret *et al.*, 1996). Large amounts of active complexes were purified from lysates expressing GST-CDK2 and Civ1 or cyclin A3 alone, by mixing them together and purifying firstly via the hexahistidine tag on cyclin A3, secondly by gel filtration and finally via GSH-Sepharose. The complexes were unstable in the presence of imidazole and the gel filtration step to remove this was essential. More recently I have purified the complexes using a single step protocol which requires making the GST-CDK2 moiety limiting, mixing the cyclin A3 and GST-CDK2 and Civ1 lysates, and simply purifying via GSH-Sepharose. These highly active complexes were able to capture transcribed and translated p27 and Cdc6, although pRb bound much more poorly.

While I was examining the ability of the GST-CDK2:cyclin A3 complexes to capture transcribed and translated pRb I found that pRb underwent a significant mobility shift in the presence of the recombinant kinase during affinity chromatography. Further experiments showed that transcribed and translated Cdc6 could also be used as substrate in kinase reactions. This suggested to me that while I was screening for binding partners of CDK2:cyclin A3 I would also be able to screen directly for substrates that underwent a mobility shift when they were phosphorylated by CDK2:cyclin A3. I have also described in this chapter how multiple small-pools of protein can be screened for binding partners of CDK2:cyclin A3 using a 96 micro-column format.

While I was establishing the conditions under which I could capture transcribed and translated Cdc6 and p27 with GST-CDK2:cyclin A3 I observed that multiple high molecular weight species of Cdc6 and p27 were also captured. I investigated whether these were polyubiquitinylated species, but was unable to demonstrate directly that this was true. Shortly after performing these experiments the ubiquitin-dependent destruction of human Cdc6 was reported, so it is possible that the high molecular weight forms of Cdc6 that I saw were due to ubiquitinylation.

I was concerned that the efficiency of capture of interacting partners by GST-CDK2:cyclin A3 would be low and that I would not be able to detect their interaction with GST-CDK2:cyclin A3 above the background due to non-specific binding of all the other labelled proteins in a transcribed and translated pool of proteins. In an attempt to increase the tightness of binding to GST-CDK2:cyclin A3 of known substrates and binding partners I attempted to mimic a strategy which had been successful for the isolation of substrates of

Chapter Five

Small-pool expression cloning for substrates of, and interactors with, CDK2:cyclin A3.

I modified the method of small-pool *in vitro* expression cloning (IVEC) originally described by Lustig *et al.* (1997) to identify proteins which could interact with, or be phosphorylated by, complexes of CDK2:cyclin A (Lustig *et al.*, 1997). In their original description they screened a *Xenopus laevis* cDNA library by comparing pools of proteins translated either in the presence of frog egg extracts arrested at metaphase of meiosis II (CSF arrest), or interphase egg extracts. Proteins which underwent a mobility shift when translated in the presence of the CSF extract but not in the interphase extracts were considered potential substrates of CDK1:cyclin B. CSF extract contains a number of other protein kinases however which may have been responsible for the phosphorylations they observed. Due to this they also had to check that the clones they identified as undergoing mobility shifts in CSF extracts could be phosphorylated *in vitro* using recombinant CDK1:cyclin B. As a further check they also showed that phosphorylation in CSF extract generated MPM2 reactivity in the translated protein.

5.1. Preparing pools of plasmid DNA.

In the previous chapter, I showed that *in vitro* transcribed and translated pRb and Cdc6 could be used as substrates for GST-CDK2:cyclin A3 in *in vitro* kinase reactions and that substrate phosphorylation could be detected indirectly due to changed electrophoretic mobility through SDS-PAGE. I had also developed a method for efficiently screening small-pools of *in vitro* transcribed and translated proteins by affinity chromatography. My next challenge was to prepare a large number of pools of plasmid DNA which could be used as template in coupled transcription and translation reactions. Aliquots of unamplified plasmid library corresponding to approximately 100 colony forming units were plated onto 488 LB plates containing 50 μg/ml carbenicillin. I used carbenicillin because it is more resistant to inactivation by the action of β-lactamase than is ampicillin. The inactivation of ampicillin by β-lactamase results in an increase in the number of satellite colonies. Carbenicillin prevents the formation of satellite colonies and so the number of plasmid

bearing colonies is kept high relative to when ampicillin is used. Colonies were harvested by adding 10 ml LB broth to each plate and by gently rotating them for 20 minutes. Miniprep DNA was prepared from the harvested bacteria using an automated Qiagen robot at the equipment park of ICRF.

When the library was constructed I titred it so that I knew that each aliquot of library contained approximately 100 CFU, although it appeared that on average there were only 70 colonies per plate. I plated 488 pools of bacteria each containing 70 independent cfus corresponding to 34 160 independent library clones. Not all of these clones would contain full length cDNAs, however, and not all would direct the transcription and translation of full length proteins. Moreover, not all cDNAs transcribe and translate efficiently *in vitro* and given that multiple templates are being used per reaction, some proteins probably translated so poorly that they were undetectable using autoradiography.

The pools of plasmid DNA could be used in any small-pool screen using T7 RNA polymerase driven coupled transcription and translation. They have been distributed to a number of laboratories who are using them to identify proteins involved in transcription, as well as proteins which are substrates of kinases and proteases.

5.2. Affinity chromatography and substrate screening can be performed in parallel.

Having prepared the pools of plasmid DNA I next began to screen for potential interactors with and substrates of GST-CDK2:cyclin A3. Figure 5-1 shows a representation of how the two screens can be performed in parallel. Initially one round of screening comprised of 48 translation reactions, 48 kinase reactions and 48 affinity chromatography columns. To increase throughput, later rounds of screening comprised 96 translation reactions and 96 kinase reactions. A 10 μl coupled transcription-translation reaction containing [^{35}S]-methionine which had been programmed with 1 μl of a pool of plasmid DNA was split so that 1 μl was retained as the input into the kinase reaction/affinity chromatography and 1 μl was used as substrate in an *in vitro* kinase reaction with GST-CDK2:cyclin A3. The remaining 8 μl were diluted and used for affinity chromatography.

Figure 5-1. Screening for substrates of, and binding partners for, CDK2:cyclin A3

A. Pool preparation and screening by electrophoretic mobility shift in SDS-PAGE.

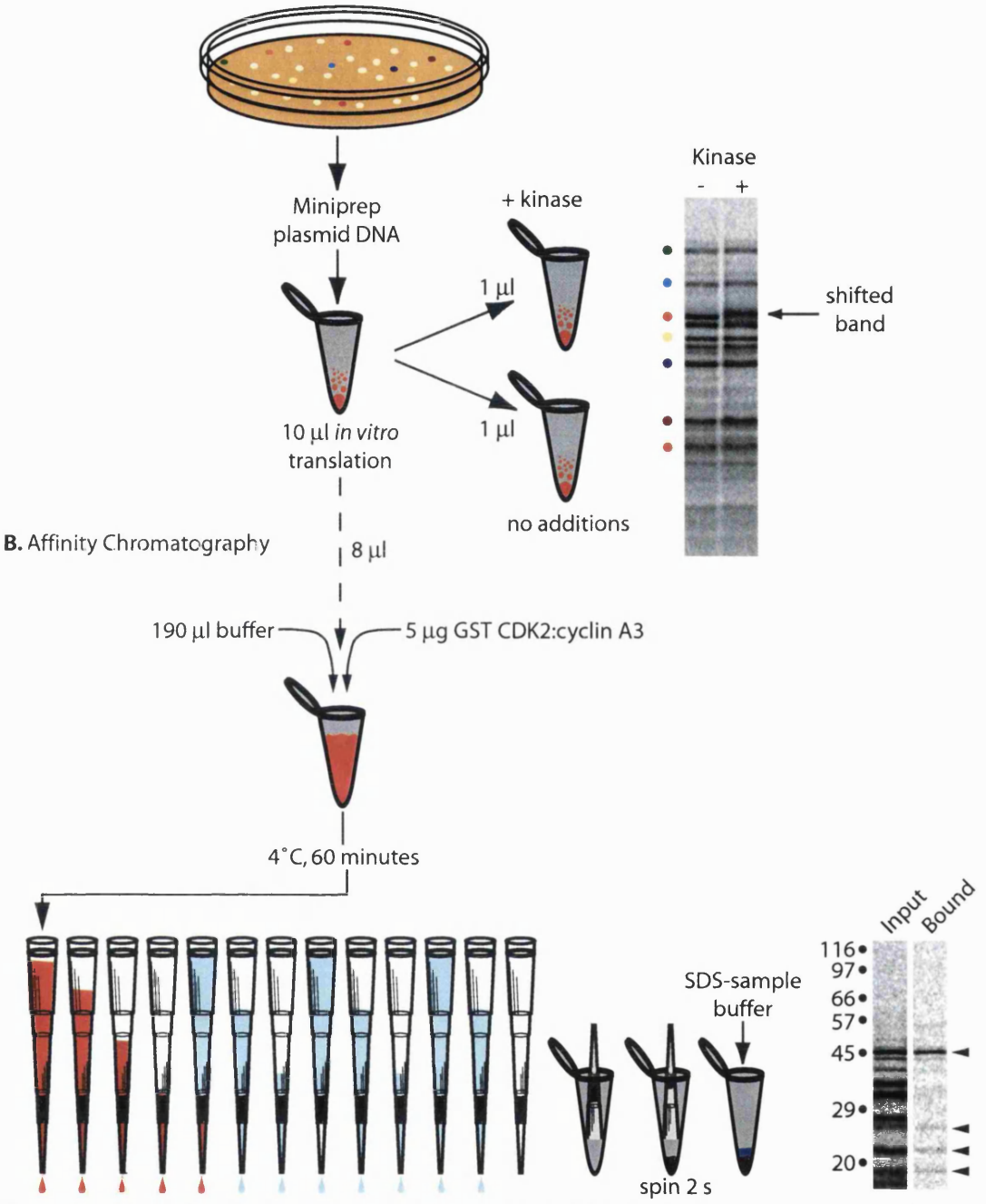


Figure 5-1. Plasmid pool preparation and screening using (**A**) a change in electrophoretic mobility through SDS-PAGE or (**B**) by affinity chromatography. (**A**) The titred library was plated on carbenicillin containing LB agar plates at a density of 100 CFU/plate. Colonies were harvested into LB and a Qiagen robot prepared pools of miniprep plasmid. A pool of plasmid DNA was transcribed and translated *in vitro*. One µl aliquots of the translation reaction were incubated in the presence or absence of GST CDK2:cyclin A3. Reaction products were analysed by SDS-PAGE and autoradiography. cDNAs encoding proteins which underwent a mobility shift in the presence of kinase were isolated and the sequenced. (**B**) The remaining 8 µl of translation reaction was diluted to 200 µl with a buffer containing 5 µg GST CDK2:cyclin A3. GST-fusion proteins and any interacting proteins were harvested using GSH-Sepharose held in the tapered end of a p200 tip with a plug of glass wool. The micro-column was washed and the tapered end of the tip containing the GSH-Sepharose was cut off and inverted in a 0.5 ml Eppendorf tip. The GSH-Sepharose was expelled from the tip by a brief pulse in a microcentrifuge. Proteins bound to the GSH-Sepharose were eluted with SDS-sample buffer and analysed by SDS-PAGE and autoradiography.

5.2.1. Screening with mobility shifts as the selection criteria.

To detect proteins which were potential substrates of GST-CDK2:cyclin A3, 1 µl of each coupled transcription-translation reaction was incubated in the presence or absence of 100 ng of CDK2:cyclin A3 complex in a volume of 10 µl. Following a 30 minute incubation at 30°C, the reactions were analysed by SDS-PAGE. The products of the reactions were visualised by autoradiography, and the electrophoretic mobilities of the translation products in the presence or absence of GST CDK:cyclin A3 were compared. Differences in electrophoretic mobilities of proteins in the pools were generally easy to spot. Figure 5-1A demonstrates that the identification of a pool containing a protein displaying an altered mobility in the presence of GST-CDK2:cyclin A3 could be simple. In the example shown, a single protein in the pool had an altered mobility following incubation with GST-CDK2:cyclin A3 and it is likely that this altered mobility is due to the phosphorylation of this protein by the added kinase. It is important to remember that the autoradiography is visualising [^{35}S] and that there is no γ[^{32}P]-ATP added to these reactions. The selection criteria is altered electrophoretic mobility, not the incorporation of [^{32}P].

Not all phosphorylation events result in altered mobilities however and so not all substrates can be identified using this selection criteria. Moreover, some phosphorylation events can increase the electrophoretic mobility rather than retard it. It is unclear which physical or chemical properties of a protein or of the phosphorylation causes these changes in the electrophoretic mobility of a protein. While some proteins undergo discrete changes in mobility through SDS-PAGE, it is also possible that a kinase reaction will not proceed to completion and incomplete phosphorylation of a protein would cause it to become smeared through the gel. The detection limitations of autoradiography could mean that a smeared band may not visualised and would seem to disappear. Although this could be attributed to actual degradation, this would have to be induced directly by CDK2:cyclin A3 with degradation being the action of an endogenous protease. Alternatively, one of the other proteins present in the pool would have to be a protease activated by CDK2:cyclin A3. If the 'disappearance' of a protein is actually the result of generating multiple electrophoretic species due to phosphorylations, then treatment of that pool with phosphatase should result in the re-appearance of the original [^{35}S]-labelled protein. If the disappearance of the band was not due to the generation of multiple phosphorylation forms but was instead due to

degradation, then treatment with phosphatase would not cause reappearance of the [^{35}S]-labelled band.

Because some proteins do not alter their mobility on SDS-PAGE I attempted to also identify proteins which were phosphorylated by GST-CDK2:cyclin A3 by including γ[^{32}P]-ATP in the kinase reactions and to use incorporation of [^{32}P] as the selection criteria for substrates. I tested whether this would work for translated Cdc6, but found that the incorporation of [^{32}P] into Cdc6 could not be detected due to the background levels of incorporation of [^{32}P] into endogenous reticulocyte lysate proteins (data not shown). Consequently, no further efforts were made with this approach.

5.2.2. Screening by affinity chromatography.

Each 8 μl which remained from the translation reactions was diluted with 190 μl of buffer to which had been added 5 μg of GST-CDK2:cyclin A3. Following an incubation of one hour on ice the GST-CDK2:cyclin A3 and any bound [^{35}S]-labelled translation products were harvested using GSH-Sepharose. Figure 5-1B shows diagrammatically how this is achieved using a GSH-Sepharose micro-column format. The diluted translation reaction containing GST-CDK2:cyclin A3 was passed through a siliconised P200 tip plugged with glass wool containing 10 μl of GSH-Sepharose. The tip is supported in a P200 tip box which is used to catch the flow through and subsequent washes. The columns were washed three times and the GSH-Sepharose was harvested by cutting one cm from the tapered end of the tips and inverting each one in a 0.5 ml Eppendorf tube and a brief spin in a microcentrifuge. The proteins captured by the GSH-Sepharose were eluted by heating in SDS-sample buffer.

The captured proteins were separated by SDS-PAGE and visualised using autoradiography. In the example shown in figure 5-2, there were four proteins in the pool which appeared to interact with GST-CDK2:cyclin A3, although the protein with an apparent molecular weight of 45 kDa seems to have been recovered with much higher efficiency. This first round of screening does not contain controls for binding of [^{35}S]-labelled proteins to either GST or to the GSH-Sepharose. Inclusion of these controls would require that the coupled transcription translation reactions be doubled in size and would demand that twice as many columns be prepared and used. It seemed more efficient

Figure 5-2. Identifying single plasmids encoding proteins of interest from the primary screen.

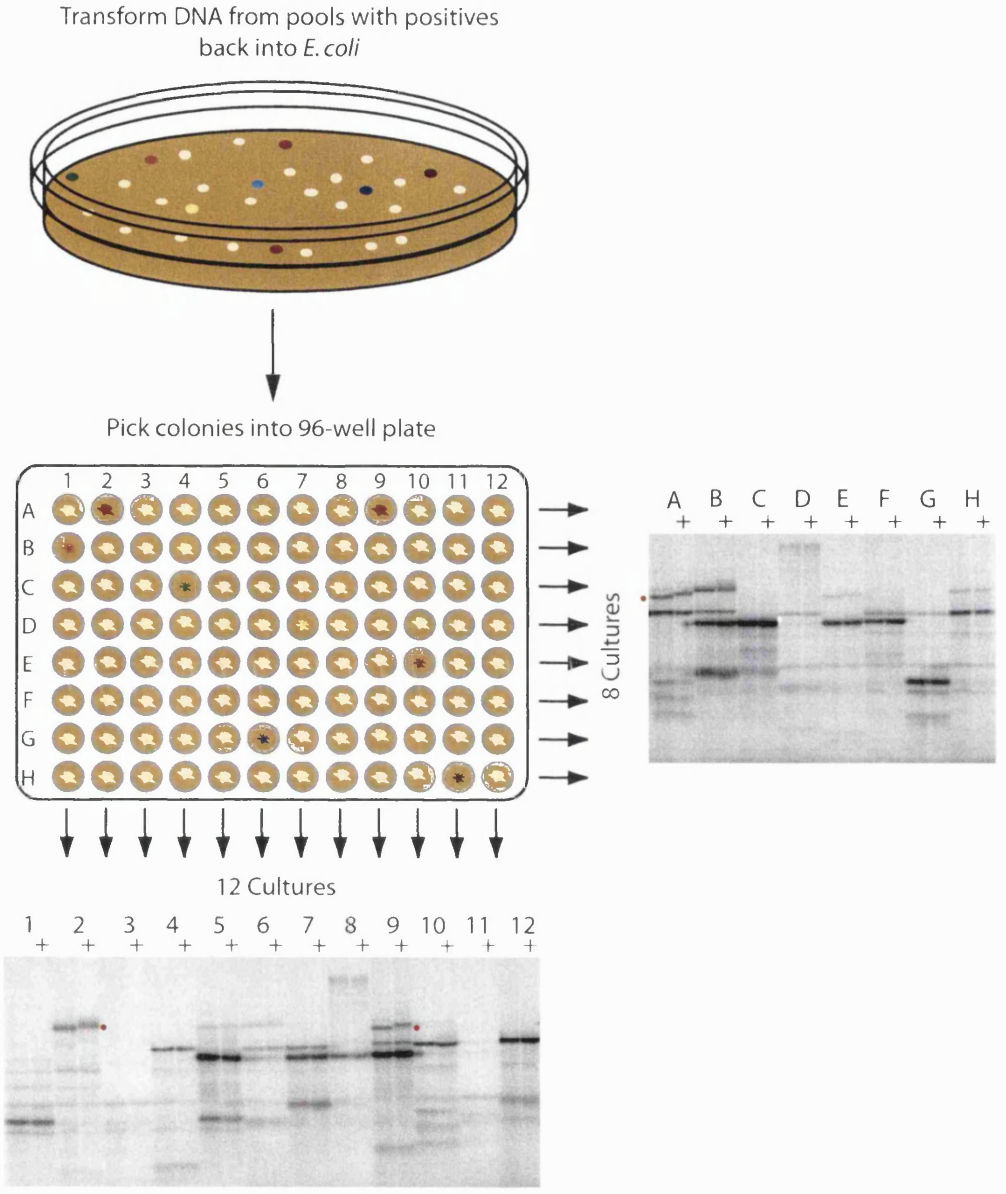


Figure 5-2. Scheme to identify clones encoding proteins of interest. A pool containing a plasmid encoding a protein of interest is transformed into *E. coli* and plated at a colony forming density of approximately 200 CFU/plate. Single colonies are inoculated into the wells of a 96-well microtitre plate which are filled with 150 µl of LB agar containing ampicillin. Each of the wells in a column on the plate are then used to inoculate a single culture and this is repeated until 12 cultures corresponding to the 12 columns (1-12) have been prepared. This is repeated for the wells in the rows (A-H). Thus, 12 cultures containing 8 different inoculates and 8 cultures containing 12 different inoculates were generated. The 20 cultures were grown overnight and used to prepare miniprep DNA. The screen is repeated as shown in figure 5-1. If the colony harbouring the plasmid of interest was picked into the microtitre dish then it wil be present in a culture from a row and a culture from a column. Cross referencing these will give the well in the microtitre dish inoculated with the colony harbouring the plasmid containing the cDNA of interest.

to wait until a pool containing a protein interacting with GST-CDK2:cyclin A3 was identified and then to perform the necessary controls for binding of translation products to GST or GSH-Sepharose. In the example shown in figure 5-1B, the protein appeared to bind GST-CDK2:cyclin A3 relatively tightly. Further tests showed, however, that it interacted strongly with GSH-Sepharose without bound GST-CDK2:cyclin A3. This suggested that the translated protein was not interacting specifically with GST-CDK2:cyclin A3, but was a screening artefact. In fact, I was unable to identify any strongly binding proteins using this screen.

5.3. Isolating single plasmids containing the cDNA encoding the protein of interest.

An important requirement of a screen is that individual cDNAs encoding proteins of interest be easily identifiable. In the original screen described by Kirschner and Stukenberg, identification of a single clone encoding a potential substrate was achieved by progressively reducing the size of the pool in which the clone of interest was present (King *et al.*, 1997). This required at least three or four further rounds of screening. Instead of reducing the complexity of the pool until a single clone was present, I adopted an approach which meant that the single cDNA encoding the protein of interest could theoretically be obtained in a single further round of screening.

Figure 5-2 shows how a single clone encoding a protein of interest can be isolated from a complex pool containing up to 96 different plasmids. The relevant plasmid pool was re-transformed into *E. coli* and plated at a density of about 200 CFU. Each individual colony was then picked from the plate into an LB agar filled well of a 96-well microtitre dish until each well of the dish contained an independent colony. The 12 colonies in each of the eight rows were picked into 8 cultures (A-H), to give 8 cultures containing 12 colonies each, for a total of 96 picks. The eight colonies in the each of the 12 columns were picked into 12 cultures (1-12), to give 12 cultures containing 8 colonies each, for a further 96 picks. The cultures were grown overnight and plasmid DNA was prepared from them. Twenty coupled transcription-translation reactions were performed using these plasmids as template. In the example shown, the substrate screening using GST-CDK2:cyclin A3 was repeated and the kinase reactions products were analysed by SDS-PAGE and autoradiography.

Figure 5-3. A typical round of small-pool screening selecting on the basis of altered electrophoretic mobility in the presence of GST CDK2:cyclin A3.

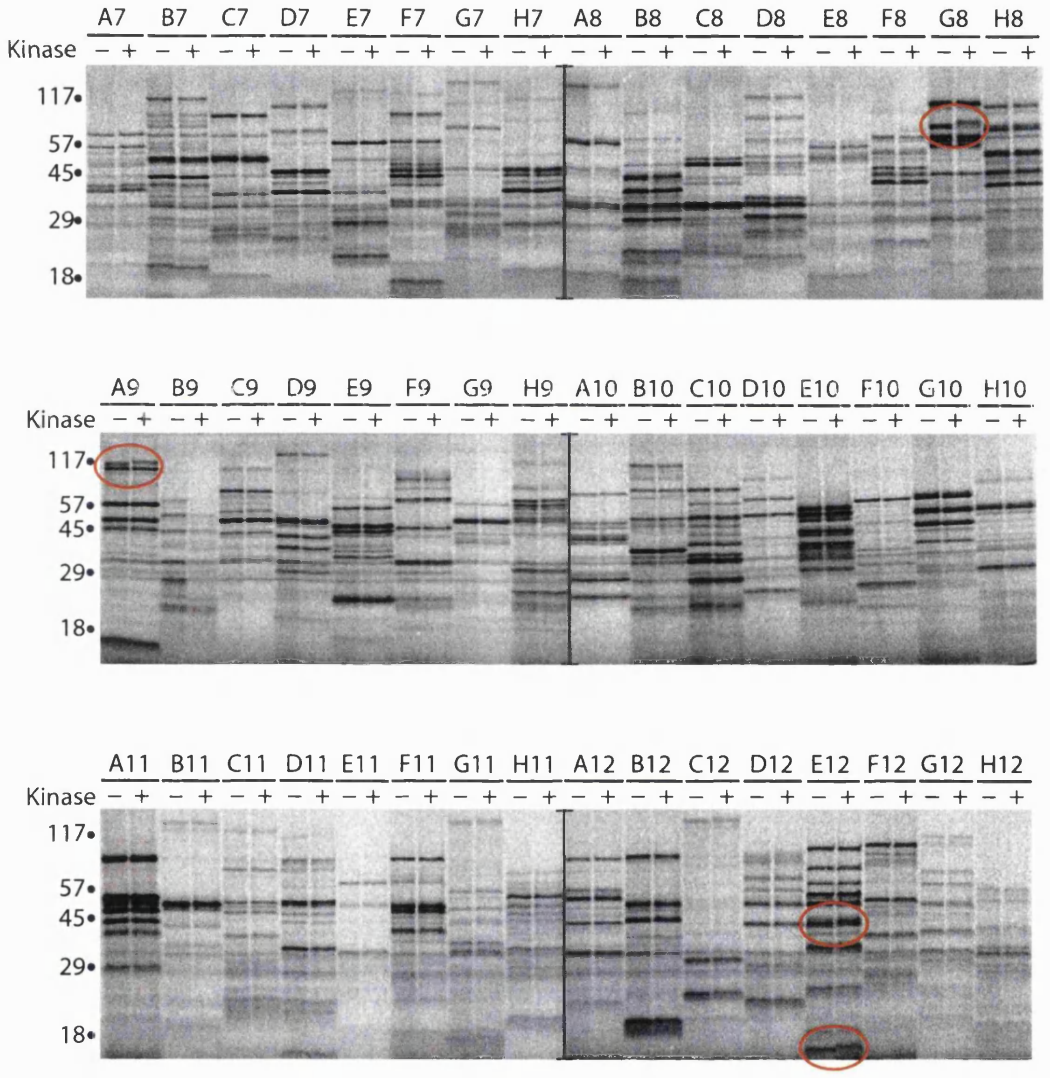


Figure 5-3. Autoradiograms of SDS-PAGE analysis of transcribed and translated [^{35}S]-methionine labelled proteins in the presence or absence of GST CDK2:cyclin A3. Small-pools of plasmid DNA were transcribed and translated *in vitro* in the presence of [^{35}S]-methionine. One µl of translation products were incubated in the presence or absence of 111 nM GST CDK2:cyclin A3 for 30 minutes at 30°C. Reaction products were analysed by SDS-PAGE and autoradiography. Lanes are analysed pairwise, with differences in electrophoretic mobilities between proteins in adjacent lanes of a pair being easily identified. Proteins undergoing an electrophoretic mobility shift in the presence of GST CDK2:cyclin A3 are circled in red.

The lanes of the gel were labelled according to whether they contained translation products synthesised using plasmid prepared from a culture generated from the columns (lane pairs 1-12), or from the cultures prepared from the rows of the microtitre dish (lane pairs A-H). In the example shown, lane pair 2 and lane pair 9 from the lanes of translation products generated from the plasmid from the 12 column cultures contained a translation product which underwent a mobility shift in the presence of GST-CDK2:cyclin A3. Lane pair A from the lanes of translation products generated from the plasmid from the 8 row cultures contained a translation product of the same molecular weight as that in lane pairs 2 and 9 which underwent a mobility shift in the presence of GST-CDK2:cyclin A3. Columns 2 and 9 both contained the clone of interest, as did row A. Wells 2A and 9A therefore both contained a bacterial colony harbouring the plasmid containing the cDNA of interest. This was checked by culturing the colonies in 2A and 9A, preparing miniprep DNA from the cultures and repeating the kinase assay using the translated products as substrates. When a pool of plasmid containing 96 different plasmids is transformed into *E. coli* which are then plated at a density of 96 CFU/plate and 96 independent colonies are picked into the wells of a microtitre dish, there should on average be one copy of each clone per dish. Generally however, the clone of interest appeared with greater frequency suggesting that the pools of plasmid contained less than 96 clones.

5.3.1. Screening is rapid.

Once the pools of library had been prepared a primary screen of 48 plasmid pools containing approximately 3360 independent library clones could be performed in a single day. To obtain a single cDNA encoding a protein of interest takes a further day to transform and grow the cultures corresponding to the rows and columns from the microtitre dish. The second round of screening then takes another day to perform. A single clone encoding a potential substrate or binding partner could therefore be obtained in as little as three days. Should the clone of interest not be represented in the second round of screening, then it can be repeated until the clone is identified, adding a further two days to the screening process for each round required.

5.4. A typical round of screening for substrates of GST-CDK2:cyclin A3.

Figure 5-3 shows the results of a typical round of screening using altered mobility of translated proteins in the presence of GST-CDK2:cyclin A3 as the selection criteria. The lanes are paired: one lane containing 10% of the 10 µl translation reaction incubated in the absence of kinase; the other containing 10% of the translation reaction which had been incubated in the presence of 100 ng GST-CDK2:cyclin A3. Proteins which appeared to undergo a shift in electrophoretic mobility are circled in red. In this round of screening there were four proteins which appeared to shift, although the two circled in AE12 could possibly be the same protein with the lower band representing an internal initiation or premature termination product of the higher molecular weight protein. The differences in mobility are generally easily spotted and require that little time be spent analysing the autoradiograms. This first round of screening confirmed the presence of many large cDNAs in the library, although on average only 10-20 proteins were produced per pool of plasmid. So although in theory each pool contained on average 70 library clones, only approximately 21-22% of these clones produced proteins. This has serious consequences for the number of proteins actually screened and reduced the approximate number from 3 360 to 722. In this particular round of screening therefore, four potential substrates were present from a potential 722 candidates, or approximately 0.6%.

5.5. A typical round of screening using affinity chromatography.

Figure 5-4 shows the results from the first round of screening using affinity chromatography as the selection criteria for clones of interest. The same pools were used as in figure 5-3. From a 10 µl translation reaction, two µl are removed as inputs into the kinase reactions, and so the remaining 80% was used for affinity chromatography with GST-CDK2:cyclin A3. Each translation reaction was diluted to 200 µl in a buffer containing 5 µg GST-CDK2:cyclin A3 complexes. The kinase complexes and any interacting proteins were captured using GSH-Sepharose which had been immobilised in the thin end of P200 pipette tip. The columns were washed, the resin harvested and any interacting proteins were eluted using SDS-sample buffer. Each lane of figure 5-4 contains 100% of the material bound to CDK2:cyclin A3. Although I had shown that these complexes, even in massive molar excess would not completely deplete a translation

Figure 5-4. A typical round of screening using interaction with GST CDK2:cyclin A3 as the selection criteria.

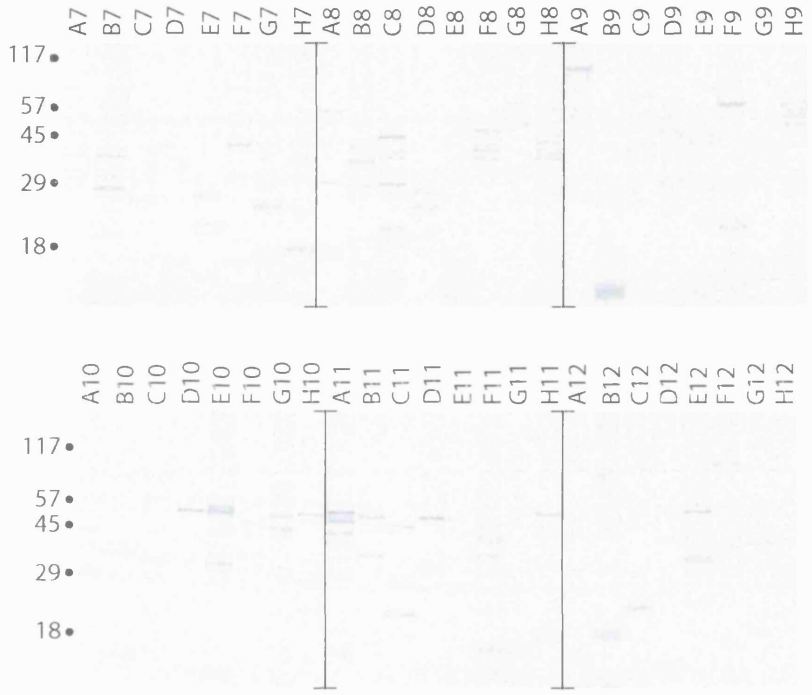


Figure 5-4. Autoradiogram of SDS-PAGE analysis of [^{35}S]-labelled proteins that interacted with GST CDK2:cyclin A3. Forty-eight pools of plasmid DNA were transcribed and translated *in vitro* in the presence of [^{35}S]-methionine. The scheme outlined in figure 5-1(B) was then used to select proteins which interacted with GST CDK2:cyclin A3. Captured proteins were eluted using SDS-sample buffer.

Figure 5-5. Bands undergoing an electrophoretic mobility shift in the presence of GST CDK2:cyclin A3.

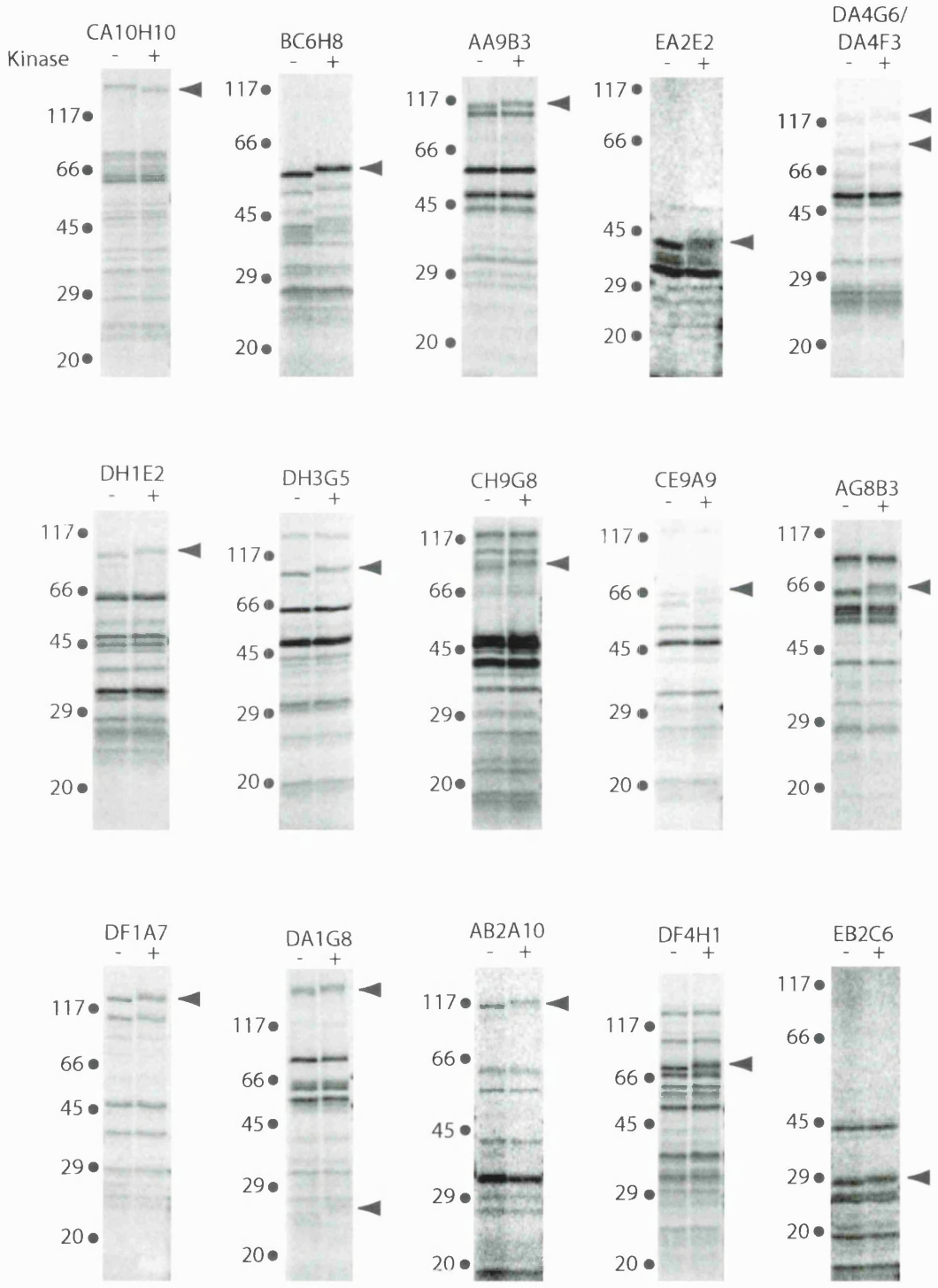


Figure 5-5. Multiple substrates of CDK2:cyclin A3 can be identified using small-pool expression screening and a selection criteria of electrophoretic mobility shift. Pools of plasmid DNA were transcribed and translated in the presence of [^{35}S]-methionine and 1µl aliquots were incubated in the absence (-) or presence (+) of 111 nM GST CDK2:cyclin A3 for 30 minutes at 30°C. Reaction products were analysed by SDS-PAGE and autoradiography. Differences in the mobilities of the labelled proteins are easily spotted in adjacent lanes. Arrowheads indicate positions of the proteins undergoing shifts.

reaction of a known interacting protein such as p27 (figure 4-9) it was surprising to see that there were so few proteins which were not captured by GST-CDK2:cyclin A3. Having shown that as much as 40% of translated Cdc6 could be captured using 5 μg of complex it was surprising that not a single protein from these 722 interacted sufficiently well to allow the capture of at least 10% if the input into the affinity chromatography reaction.

5.5.1. Analysing the affinity chromatography screening data was more complex than the substrate screening data.

The results from the affinity chromatography require significantly more analysis than those from the kinase screening. Each of the weak signals on the gels containing the bound material must be compared with the input lanes to determine whether the retained material represents a significant capture of the input. In this particular round none of the bands in figure 5-4 represents a capture greater than a few percent of the input. Given that affinity chromatography should serve as an enrichment step, it appeared that none of the proteins tested here show a significant interaction with GST-CDK2:cyclin A3. Based on the weakness of the interaction between some substrates (such as pRb) and CDK2:cyclin A3 it is possible however that there may be a relevant interaction between some of these proteins and GST-CDK2:cyclin A3. The work required to test each of the 60 weak interactions shown here would be considerable however, and that time could be better spent focusing on the kinase screening, and testing more clones by affinity chromatography for proteins interacting more strongly with GST-CDK2:cyclin A3.

I screened a further 96 pools using both the kinase reaction and affinity chromatography approaches. From these pools representing 6720 library clones and 1444 actual proteins, I found only one protein which showed a significant interaction with CDK2:cyclin A3 (figure 5-1B). As stated previously this protein interacted with GST bound to GSH-Sepharose equally well however, and this suggested that it was unlikely to represent a functional interactor with CDK2 or cyclin A3.

5.6. How many proteins were screened for substrates of CDK2:cyclin A3, and how many 'substrates' were identified?

At this point, I decided to stop screening using the relatively time intensive affinity chromatography approach, and instead focus on the simpler, quicker and more successful

screen has reached saturation as only 34200 clones or 7300 proteins have been screened. The recent completion of the human genome project suggests that there are some 30000 coding sequences in the human genome, and therefore even if a normalised library were used then these 7300 screened proteins would only represent one quarter of the total proteins needed to be screened to reach saturation. It is unusual therefore that the same protein has been identified in this type of screen more than once. Since I established the small-pool screening technique in the Hunt Laboratory, my colleague Yoshimi Tanaka has performed a similar small-pool screen for substrates of CDK1:cyclin B and during the course of her screening identified the *Xenopus laevis* ortholog of MARK2 suggesting that it is an *in vitro* substrate for both cyclin A and B associated kinase activities. (Y. Tanaka, personal communication). The independent isolation of this clone also supports the reliability of the screening method, although my own screening failed to identify other known substrates of CDK2:cyclin A such as Cdc6 and p107.

The known functions of the clones isolated using small-pool expression cloning for substrates of CDK2:cyclin A are diverse (figure 5-6) and a detailed description of all their individual functions is beyond the scope of this thesis. Briefly however, a number of proteins were identified which have roles in transcription (general transcription factor 3C α, TFE3 and MTA1) or translation (heme-regulated initiation factor 2α kinase, eIF4G-1, eRFS and calnexin). Others have clear roles in signal transduction (MARK2, BCAR3/NSP2 and OS-9), and some have roles in cell cycle events (MCM2, Ect2 and nucleophosmin).

5.8. Proof of principle: Identification of MCM2 and nucleophosmin as substrates of CDK2:cyclin A3.

Although most of the proteins identified by substrate screening underwent a retardation of mobility following phosphorylation, the protein identified in pool CA10H10 clearly migrated faster following incubation with GST-CDK2:cyclin A3. Isolation of this clone and sequencing of the cDNA insert identified this protein as MCM2, a protein involved in the initiation of DNA replication which is a component of the putative replication helicase, the hetero-hexameric MCM complex (Labib and Diffley, 2001). MCM2 has previously been reported to undergo an electrophoretic mobility shift during the cell cycle, and has

been recently reported to be a substrate of both CDK2 and dbf4 (Masai *et al.*, 2000). In this respect the identification of an already characterised substrate of CDK2 demonstrates that a screen such as this can successfully be used to identify substrates of CDK2:cyclin A3.

5.8.1. Nucleophosmin is an *in vitro* substrate of CDK2:cyclin A3.

Further to the identification of MCM2 by small pool expression cloning, clone EA2E2 was identified as nucleophosmin. Nucleophosmin was originally identified as a major nucleolar phosphoprotein and has been implicated in regulating nucleolar structure and function, ribosome assembly, DNA polymerase α function, cytoplasmic/nuclear trafficking and has been reported as having protein chaperone and ribonuclease activity (Herrera *et al.*, 1995; Szebeni and Olson, 1999). Shortly after I identified nucleophosmin as a substrate of GST-CDK2:cyclin A3 using small-pool expression cloning, it was also identified as an *in vitro* substrate of CDK2:cyclin E using an entirely different approach (Okuda *et al.*, 2000). Centrosomes were isolated through sucrose density gradients and incubated with CDK2:cyclin E purified from insect cells. It was found that nucleophosmin was present in the centrosomal material and that it was phosphorylated by CDK2:cyclin E. Further work using antibodies raised against nucleophosmin showed that it localised at the centrosomes during G1 but not during S-phase or G2 and that CDK2:cyclin E induced displacement of nucleophosmin from centrosomes was essential for centrosome duplication late in G1. Fukasawa *et al* have since shown that nucleophosmin is phosphorylated on T199 by CDK2:cyclin E, that CDK2:cyclin A is also capable of phosphorylating T199 with similar efficiency and that phosphorylation on T199 occurs *in vivo*. Further to this it was also shown that CDK1:cyclin B is capable of phosphorylating nucleophosmin on T234 and T237, but not on T199 (Tokuyama *et al.*, 2001). The independent identification of nucleophosmin as a substrate of CDK2 demonstrated that small-pool expression cloning using purified, recombinant kinases can be used to identify novel, physiologically relevant substrates of CDKs.

5.8.2. Ect2 is an *in vitro* substrate of CDK2:cyclin A3.

Another protein identified as a substrate of CDK2:cyclin A using small-pool expression cloning was Ect2 (AA9B3). Ect2 was originally identified in a screen for mitogenic signal transducers in epithelial cells. Injection of N-terminally truncated Ect2 (*ect2*) into nude mice efficiently induced tumour formation and showed that Ect2 is an oncogene (Miki *et*

al., 1993). Ect2 is a an exchange factor for Rho-GTPases and was shown to be phosphorylated during late G2 and mitosis, but not during S-phase or G1. This cell cycle dependent phosphorylation was shown to greatly enhance the exchange activity of Ect2 on Rho proteins (Tatsumoto *et al.*, 1999). It was possible, although not directly proven that the cell cycle dependent phosphorylation could be catalysed by CDK1 complexed either with cyclins A or B. Ect2 will be discussed in greater detail in the following chapter.

5.9. General Discussion.

This chapter has described a small-pool expression screen for substrates of CDK2:cyclin A3. This thesis began with the idea of performing a small-pool expression screen using affinity chromatography, however during the optimisation of this screening procedure I found that it would also be possible to select directly for substrates of CDK2:cyclin A3 using changes in the electrophoretic mobility through SDS-PAGE of a translated protein as the selection criteria. The two techniques of affinity chromatography and kinase-substrate selection were initially employed in parallel. After having screened through 144 pools of translated proteins corresponding to 10080 primary library clones or roughly 2200 translated proteins it became apparent that there were few (if any) proteins which interacted sufficiently strongly with CDK2:cyclin A3 during affinity chromatography to allow easy identification.

At this point I stopped screening by affinity chromatography and concentrated on the direct substrate selection method. I screened a total of 488 pools of plasmid corresponding to a total of approximately 7300 translated proteins and found 16 proteins which underwent mobility shifts through SDS-PAGE after incubation in the presence of CDK2:cyclin A3. I have omitted one protein from the table in figure 5-6 as the 5' and 3' cDNA sequences from this clone did not match. The 5' sequence was identified by BLAST search as high mobility group protein 14 (HMG14), a chromatin associated protein (Graziano and Ramakrishnan, 1990; Nissen *et al.*, 1991) and the 3' sequence was identified as thrombospondin-1 (TSP1), an extracellular matrix glycoprotein secreted by many cell types that inhibits tumour cell growth and metastasis (Clezardin *et al.*, 1993). It is possible that the translation product encoded by this clone was generated by the artefactual ligation of cDNA fragments during library construction and is in fact a chimeric cDNA. HMG14 (recently renamed HMGN), is the only protein currently known to bind directly to the

nucleosomal structure and is dissociated from chromatin by phosphorylation during mitosis (Prymakowska-Bosak *et al.*, 2001). It has also been reported however that 'Cdc2-type' Polo-like kinases are responsible for HMG14s phosphorylation. If the mobility shift induced by CDK2:cyclin A3 in the fusion protein was found to be due to phosphorylation of the HMG14 moiety of the protein, then this may warrant further investigation.

Some of the identified *in vitro* substrates of CDK2:cyclin A are unlikely to represent physiological targets. For example, clone DA4F3 was found to encode OS-9 a gene frequently co-amplified with CDK4 in osteosarcomas. When OS-9 was analysed by the SignalP programme which identifies secretory signal peptides (http://www.cbs.dtu.dk/services/SignalP/) (Nielsen *et al.*, 1997) it was found to contain a signal peptide between residues 2-26, with cleavage of the signal peptide occurring between residues 26 and 27. Virtually all proteins containing a signal peptide enter into the secretory pathway and it seems unlikely that OS-9 would ever be in an environment where CDKs are present. This illustrates a further flaw in small-pool expression screening for substrates of CDK2:cyclin A3.

When a cDNA library is constructed and effectively expressed *in vitro*, many of the expressed proteins will have been removed from their usual cellular context. I have briefly commented on the consequences on protein folding of expressing cDNAs in a heterologous expression system such as rabbit reticulocyte lysates but it also has consequences in terms of mis-localisation. A secreted protein, such as OS-9, may represent an excellent *in vitro* substrate of CDK2:cyclin A3 and indeed it has a good consensus sequence for phosphorylation by CDKs (SPTK) and a number of other SP/TP dipeptides. In its usual cellular or extracellular context OS-9 may never come into contact with CDK2 however. It is certainly possible that some of the clones identified in this screen are never in the same physiological compartment as CDK2:cyclin A3. In this respect, small-pool expression cloning for substrates will identify some false positives, and when the functions of the cloned cDNAs are known it is possible to make educated guesses at the identities of the false positives. When novel sequences are identified however, the utility of a stringent secondary screen becomes apparent.

Chapter Six.

Characterising the clones identified as substrates of CDK2:cyclin A3.

The small-pool screening method that I have used identified proteins which are *in vitro* substrates of CDK2:cyclin A3. Whether they are *bona fide in vivo* substrates is not a criterion this screen selects for and consequently any *in vitro* substrate identified also needs to be shown to be a physiologically relevant substrate. A physiologically relevant substrate is one that is phosphorylated in a cell cycle dependent manner at a time when the CDK:cyclin pair suspected of catalysing the phosphorylation is active. Also it should be shown that phosphorylation of the substrate causes a change in its behaviour, such as activity, localisation or abundance, and that the change(s) have an effect on a cell cycle event.

CDKs are promiscuous enzymes with few identified specificity determinants at the level of the primary amino acid sequence of the substrate. The consensus CDK phosphorylation motif is minimally known to be S/TPXK/R where X is any amino acid and the lysine or arginine in the +3 position (relative to the phosphorylated residue) is in some examples not required (Gurley *et al.*, 1995). Adams *et al* described the identification of a motif in substrates which targets them to CDK:cyclin complexes (Adams *et al.*, 1996). Schulman *et al* analysed the structure of cyclin A and identified a conserved hydrophobic patch on the surface of cyclin A which was shown to be involved in the targeting of CDK:cyclin complexes to substrates (Schulman *et al.*, 1998). This hydrophobic patch on cyclin was initially shown to interact with a short amino acid sequence found in a number of characterised substrates of CDKs which is called the Cy motif or RXL motif.

Sequence analysis of novel substrates of CDKs can be performed to identify consensus CDK phosphorylation and Cy motifs. The presence of an S/TP motif is an essential requirement for a CDK substrate, but as previously indicated this is also the hallmark of other proline directed kinases such as MAPK. The presence of a good Cy consensus motif would also support the suspicion that a protein is a substrate of

CDK:cyclin although neither of these primary sequence elements are sufficient to establish that a protein is a *bona fide* substrate and further characterisation is required to establish this.

6.1. A good secondary screen is hard to find.

The small-pool expression screen described in this thesis identified 15 *in vitro* substrates of CDK2:cyclin A3 of which two have been independently verified as physiologically relevant substrates of CDK cyclin complexes (MCM2 (Lei *et al.*, 1997; Masai *et al.*, 2000) and nucleophosmin (Okuda *et al.*, 2000; Tokuyama *et al.*, 2001)). These proteins are important mediators of cell cycle events, and moreover their phosphorylation by CDKs has also been shown to cause a change in their behaviour which causes a cell cycle event. Ect2, a third *in vitro* substrate of CDK2:cyclin A identified in this screen shows a phosphorylation-dependent change in activity and a dramatic re-localisation as cells progress through mitosis (Tatsumoto *et al.*, 1999). The characterisation of both MCM2 and nucleophosmin as *bona fide* substrates of CDKs represents a significant body of work and similar efforts using each of the 13 other clones identified in the small-pool expression clone are unrealistic. It is beneficial therefore to be able to identify clones which are likely to represent substrates of CDK2:cyclin A. A comprehensive and quick secondary screen is difficult to conceive however as it should generally comprise a functional assay such as the detection of a change in the activity of a protein upon phosphorylation.

For a protein such as Ect2 which has a well defined activity (GTP-GDP exchange activity) (Kimura *et al.*, 2000; Miki *et al.*, 1993) this is not such a challenge, however when 13 different proteins with 13 different activities need characterising, some of which may be unknown, this is a more daunting challenge. CDK2:cyclin A complexes are believed to have roles during S-and G2-phases of the cell cycle, while the activity of CDK1:cyclin A is believed to be important during late G2 and through mitosis until cyclin A becomes degraded just prior to the metaphase-anaphase transition (Geley *et al.*, 2001). A sensible prediction might be that substrates of CDK2:cyclin A should be nuclear during S and G2-phases of the cell cycle. If cyclin is the main mediator of substrate selection and recruitment to CDK then it is possible that this type of screen could also select substrates of CDK1:cyclin A. In this case it is possible also that substrates may be localised throughout the cytoplasm when the nuclear envelope breaks down.

6.1.1. Cellular localisation as a selection criterion for substrates of CDK:cyclin A3.

Normally, one would expect substrates of CDKs to be located in the same subcellular compartment as the kinase when the kinase of interest acts (Draviam *et al.*, 2001; Pines and Hunter, 1991). A conventional means of establishing cellular localisation of a protein of interest is using immunofluorescence. Immunofluorescence requires that a suitable antibody recognising the protein of interest be available, and for many of the proteins identified in this screen that was not the case. An alternative to immunofluorescence is to express the protein of interest in tissue culture cells as a fusion to yellow-, green-, red- or cyan-fluorescent protein (YFP, GFP, RFP or CFP respectively) and to then use conventional fluorescence microscopy to visualise the expression pattern of the fusion protein in the cell.

6.2. KIAA0144 as a potential novel substrate of CDK2:cyclin A.

The first clone identified as an *in vitro* substrate of CDK2:cyclin A3 was AB2A10. Sequence homology searches of the publicly available databases showed that this clone was identical to the sequence KIAA0144. KIAA0144 was identified as part of a project dedicated to the isolation of long cDNAs (Nagase *et al.*, 1995). KIAA0144 is a hypothetical protein of 104 kDa and to date no information is available about its cellular function, localisation or expression pattern. As a protein of no known function it was entirely possible that it could represent a novel *in vivo* substrate of CDK2:cyclin A. As I had only a partial cDNA clone of KIAA0144 I obtained a full length from Dr N. Nomura, Kazusa DNA Research Institute, Chiba, Japan.

To gain some insight into the potential function of KIAA0144 I performed a conceptual translation of the full length cDNA and submitted the deduced amino acid sequence to the Simple Modular Architecture Research Tool (SMART) programme at WWW.EMBL-HEIDELBERG.SMART.DE (Schultz *et al.*, 1998). SMART is able to detect domains in proteins and may give a clue to the function of the protein. The SMART algorithm detected the presence of an N-terminal ubiquitin-associated domain (UBA) domain which has been reported to bind ubiquitin chains (Bertolaet *et al.*, 2001). A diverse range of proteins possess UBA domains including the DNA damage inducible proteins

hRad23A and B which have been shown to be involved in the Pds1-dependent mitotic-checkpoint in yeast, a number of Ser/Thr kinases including the CDC25C-associated kinase C-TAK and many proteins in the ubiquitination pathway (Bertolaet *et al.*, 2001). The presence of a UBA domain in KIAA0144 suggested the possibility that this protein may play a role in cell cycle control or checkpoint regulation. This made it an exciting candidate to be a novel substrate of CDK2:cyclin A3.

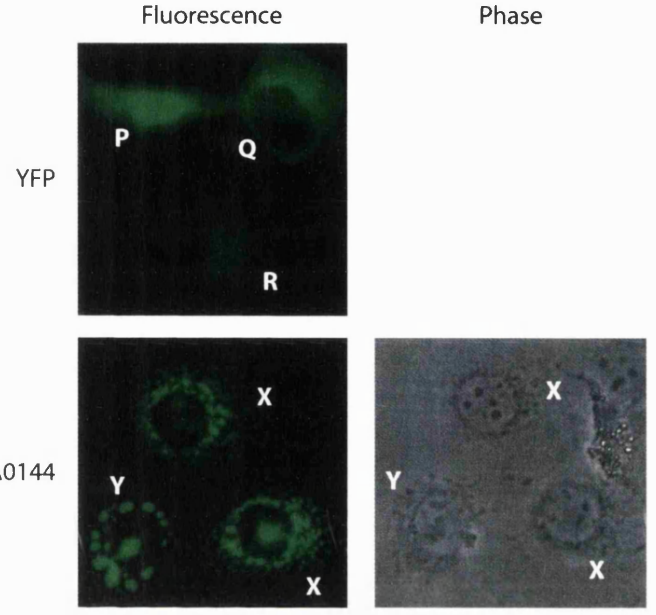
6.2.1. Localisation of YFP-KIAA0144 in HeLa cells.

I wanted to determine whether KIAA0144 was present in the nucleus of the cell. I constructed a plasmid containing KIAA0144 fused to the C-terminal of YFP under the control of the EF2α promoter. To facilitate the cloning process I reconstructed the multiple cloning site of a plasmid (pEFplink) donated by Dr. Ralph Graeser and renamed it pEF MCS YFP. pEF MCS YFP-KIAA0144 was transfected into Cos-1 cells using TransitLT transfection reagent. As a control for transfection pEF MCS YFP was transfected into duplicate wells of cells. The cells were cultured overnight and expression of YFP-KIAA0144 in the living cells was analysed using a Zeiss Axiovert 10 inverted microscope and a Coolsnap Raper colour CCD camera. Figure 6-1 shows images captured using the colour camera in black and white mode. Colour was added to the fluorescence images using Adobe Photoshop. Transfection efficiencies approached 50% with both pEF MCS YFP and pEF MCS YFP:KIAA0144. As is shown in the upper panel of figure 6-1A, YFP was excluded from nuclei in some cells; localising diffusely throughout the cytoplasm (cell Q). In other cells YFP accumulated in the nuclei while also localising to the cytoplasm (cell P) and was also seen to appear faintly diffuse throughout both cytoplasm and nucleus. There was no specific concentration of YFP at cellular structures or organelles. These differences in localisation could be due to the different expression levels of the protein in the cells and is likely a direct consequence of the amount of plasmid delivered to individual cells.

YFP-KIAA0144 was excluded from nuclei and accumulated in the cytoplasm of living cells. Figure 6-1A shows that YFP-KIAA0144 localised to distinct speckles throughout the cytoplasm (cells X). In some cells however the speckles were larger and

Figure 6-1. Localisation of YFP-KIAA0144 in Cos-1 and HeLa cells.

A. YFP-KIAA014 localised to cytoplasmic speckles in transfected Cos-1 cells.



B. YFP-KIAA0144 localised to cytoplasmic speckles in microinjected HeLa cells.

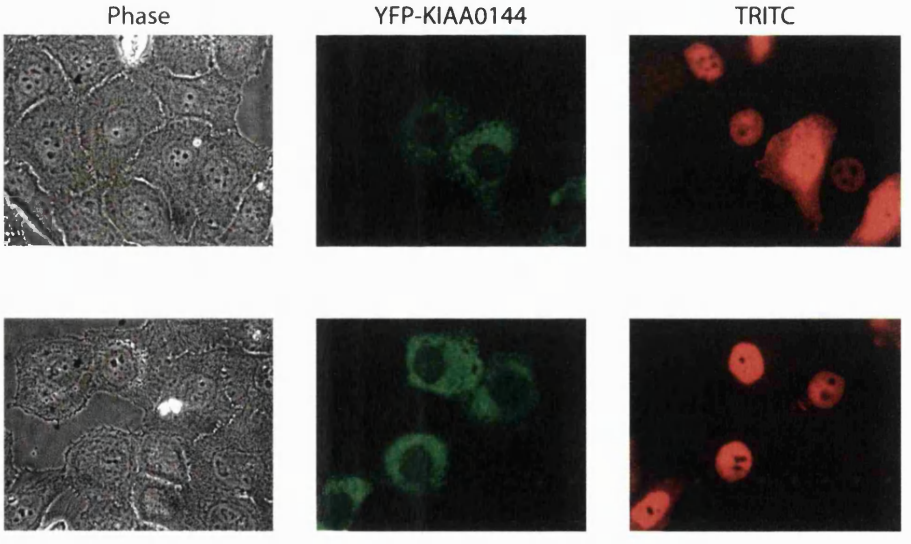


Figure 6-1. Overexpressed YFP-KIAA0144 localises cytoplasmically in Cos-1 and HeLa cells. **A.** Cos-1 cells were cultured to 50% confluence in 6 well dishes to 50% confluence and were transfected with 1 µg/well of plasmid encoding YFP-KIAA0144 using TransitLR transfection reagent. Twenty-four hours post-transfection live cells were imaged using a Zeiss Axiovert 10 inverted microscope and a Coolsnap Raper colour CCD camera. **B.** HeLa cells were cultured to 50% confluence before being subjected to a double thymidine block. Cells were released into fresh media and nuclei were microinjected using an Eppendorf automatic microinjection system mounted on an Zeiss Axiovert 35. Images were obtained using an Axiovert 135TV inverted microscope equipped with a Princeton Instruments MicroMax1300 camera driven by AQM200 image acquisition software.The microscope was adapted for live cell imaging (37°C and CO_2 supply). The injection marker was TRITC labelled high molecular weight dextran.

become localised peripheral to the nuclei (cell Y). In the phase images there are also a large number of peri-nuclear particles present, although if the fluorescence and phase images are merged the fluorescence does not overlay the dense speckles seen in the phase image (not shown). In some cells it appears that there is a nuclear localisation of YFP-KIAA0144, although it is likely that this is due to its localisation in cytoplasm above or below the nucleus. Confocal imaging would determine whether this is true nuclear localisation or not. The localisation of YFP-KIAA0144 was also determined by microinjection of pEF MCS YFP:KIAA0144 into living HeLa cells which had been released from a double-thymidine block. Figure 6-1B shows the localisation of YFP-KIAA0144 in living HeLa cells. As was seen in Cos-1 cells, there was an accumulation of YFP fluorescence in cytoplasmic speckles. Unlike in the Cos-1 cells the speckles in the HeLa cells were distributed in a uniform fashion throughout the cytoplasm. TRITC conjugated high molecular weight dextran was co-injected with the plasmid and acted as an injection marker.

6.2.2. Advantages and disadvantages of using YFP-fusion proteins.

An advantage of using YFP-fusions to study the localisation of a protein of interest is that the fluorescent fusion protein can be directly visualised in living cells. The drawbacks however are that the plasmids which are generally available express the fusion protein from strong promoters such as the EF2α promoter and this leads to constitutive overexpression that can lead to mis-localisation of the fusion protein within the cell. Another problem is that the overexpression can saturate the cell and lead to the masking of the true localisation. In some cases this can be overcome by fixing the cells using conventional methods and extracting them with 0.1-1% detergents such as NP-40. It is beneficial to be able to support the localisation of YFP-fusion proteins by detecting the localisation of the endogenous protein using immunofluorescence, although in this case that was not an option as a suitable antibody recognising KIAA0144 was not available.

YFP-KIAA0144 localised cytoplasmically in fine speckles and at no point during the cell cycle did this change (not shown). When I analysed the KIAA0144 primary sequence by eye I found that it contained an RGD motif, which is characteristic of proteins with cell adhesion functions or which are generally extracellular (D'Souza *et al.*, 1991). This motif was not identified by the SMART algorithm. It seemed unlikely therefore that

KIAA0144 is a physiological substrate of CDK2:cyclin A. It is possible that KIAA0144 could be exposed to CDK1:cyclin A or CDK1:cyclin B once the nuclear membrane had dissolved and cells had entered mitosis or that the localisation changed in response to a specific stimulus or signal generated in response to an event such as DNA damage. Hexahistidine tagged KIAA0144 expressed in *E. coli* and purified on Ni^{2+}-NTA agarose was not efficiently phosphorylated by GST-CDK2:cyclin A3 (not shown). This would suggest that it is not a substrate of CDK2:cyclin A although it was possible that KIAA0144 was incorrectly folded and not able to bind to CDK2 or cyclin A.

6.2.3. Sequences showing homology to KIAA0144.

At the time of writing this thesis I re-examined GenBank for proteins homologous to KIAA0144 and found that a partial cDNA sequence encoding a protein called NICE-4 had been deposited. Sequence alignments of the conceptual translation of NICE-4 and KIAA0144 using the homologous protein fragments showed that they were 46% identical across the homologous sections. The NICE-4 sequence extended beyond the predicted termination codon in KIAA0144 however. NICE-4 was identified as a gene potentially included in the Epidermal Differentiation Complex which is a family of 27 genes believed to be involved in the maturation of the epidermis (Marenholz *et al.*, 2001). Sequence database searches found that the NICE-4 mRNA was present in a number of tissue types and casts some doubt on its involvement in epidermal maturation.

KIAA0144 contained 15 S/TP motifs of which one is a good consensus CDK phosphorylation motif (SPQK). KIAA0144 is a highly serine/threonine rich protein however; containing 155 serines and 102 threonines, representing 14.8% and 10.2% of the amino acid composition respectively, which is unusually high. One concern with using small-pool expression cloning for the identification of protein kinase substrates is that *in vitro* translation removes the protein from its usual cellular context and exposes it to a kinase which it may never be displayed to *in vivo*. It is possible to conceive how this may have lead to the identification of a serine/threonine rich protein such as KIAA0144 as an *in vitro* substrate of CDK2:cyclin A3.

6.3. AA9B3 is homologous to murine Ect2.

The sequences derived from clone AA9B3 were found to be orthologous to a murine cDNA encoding a protein called Ect2 (Miki *et al.*, 1993). At the time I cloned AA9B3 there was no human ortholog of murine Ect2 present in any of the publicly available databases. The cDNA fragment I cloned appeared to encode the full-length ortholog of Ect2. The nucleotide and primary amino acid sequences of human Ect2 (hEct2) are shown in Appendix 1. The deduced primary amino acid sequence of hEct2 was submitted to SMART to detect domains and the detected domains are shown in the schematic representation of hEct2 in figure 6-2. The murine and human orthologs of Ect2 are 78% identical, although hEct2 possesses an N-terminal extension of 143 amino acids not present in mEct2. A *Drosophila melanogaster* ortholog of Ect2 appears to be Pebbles (Lehner, 1992); hEct2 and Pebbles are 34% identical, while mEct and Pebbles are 32% identical.

6.3.1. hEct2 possesses two BRCT1 domains, a Rho-GEF domain and a pleckstrin homology domain.

The SMART algorithm identified two N-terminal BRCT1 domains, a Rho-GEF domain and a C-terminal pleckstrin homology domain. Although hEct2 contains 12 S/TP motifs, fine primary sequence analysis identified a good consensus CDK phosphorylation site in the C-terminal of the protein (KTPKR). I cloned hEct2 in October 1999 and in November 1999 Miki *et al* independently reported the cloning and characterisation of the human ortholog. Tatsumoto *et al* showed that the guanine nucleotide exchange activity of Ect2 was stimulated by a phosphorylation which occurred during G2 and mitosis and that the activity of Ect2 was required for successful completion of cytokinesis (Tatsumoto *et al.*, 1999).

6.3.2. YFP-hEct2 is a nuclear protein in G1 and G2 but associates with the spindle, the midzone and the plasma membrane during mitosis.

I cloned the full length hEct2 cDNA into the pEF MCS YFP plasmid and used this in microinjection experiments to investigate the localisation of hEct2 during the cell cycle. HeLa cells which had been released from a double thymidine block were injected in the nuclei with approximately 1-10 fl of plasmid at a concentration of approximately 50 ng/μl. Fluorescence and phase images were collected at intervals of 5 minutes one hour after the

Figure 6-2. Domain structure of Ect2.

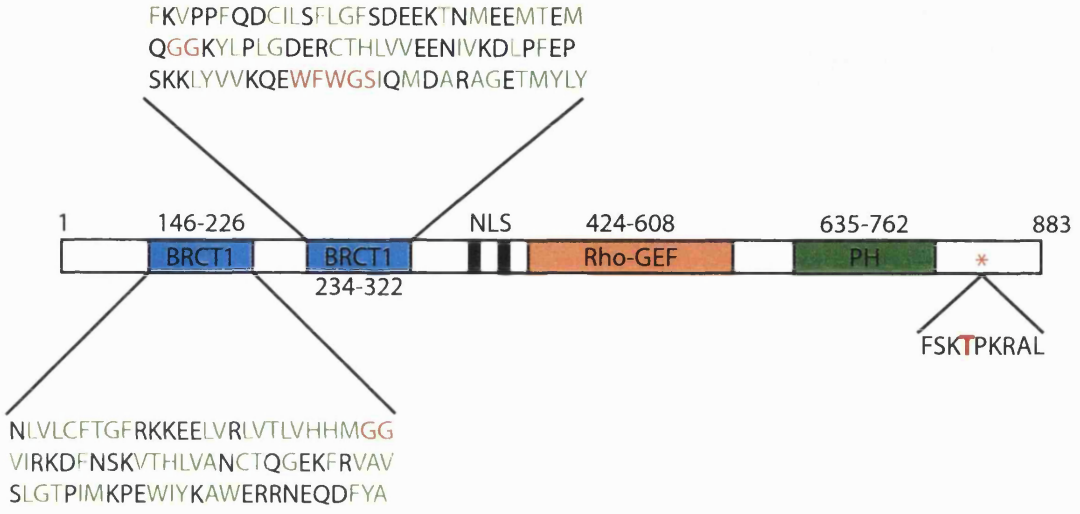


Figure 6-2. Schematic representation of domain structure of Ect2. A conceptual translation of the sequence of the full length cDNA was analysed using the Simple Modular Architecture Research Tool (SMART) at EMBL. This software predicted the presence of two N-terminal BRCT1 domains (blue), a bipartite NLS (black), a Rho-GEF domain (orange) and a pleckstrin homology domain (green). There is a good CDK phosphorylation consensus site in the C-terminus.

injections were performed. Figure 6-3 shows the localisation of YFP-Ect2 as cells traversed G2 and entered and exited mitosis.

During G2 YFP-hEct2 accumulated in the nuclei of injected cells and appeared to become slightly concentrated in nucleoli. The injected cells took an abnormally long time to complete G2 and enter mitosis. By 13 hours the nuclear envelope has solubilised and chromosomes have begun to condense and approximately 20 minutes later the cell had completely rounded up and chromosomes have aligned at the metaphase plate. The positions of the chromosomes can be seen as bleaching of fluorescence in the fluorescent images and are indicated by white arrowheads in the phase images. In prophase YFP-hEct2 is localised diffusely throughout the cytoplasm and is clearly not chromatin bound. As mitosis proceeds and cells reach metaphase, although YFP-hEct2 appears to remain diffusely cytoplasmic, there appears to be a faint staining of YFP-hEct2 on a spindle-like structure. As the chromosomes begin to move poleward YFP-hEct2 remains diffusely cytoplasmic although there is distinct fluorescent signal at the site of the metaphase plate where cytokinesis will occur. Between 14h21m and 14h35m after the beginning of image capture YFP-Ect2 also appeared to become localised at the plasma membrane while remaining concentrated at the metaphase plate. As the cell began to divide at 14h39m after the beginning of the movie there was an increase in the strength of the fluorescent signal at the plasma membrane especially in the region where cytokinesis was to occur.

There was also a 'knot' of fluorescence close to the plasma membrane in one of the newly forming daughter cells which is marked with a yellow arrowhead in the fluorescent image. As cells entered G1 YFP-hEct2 remained diffusely cytoplasmic, although there was still a distinct fluorescent signal at the plasma membranes of the newly formed daughter cells. YFP-hEct2 rapidly accumulated in the nuclei of the daughter cells. Miki *et al* used immunofluorescence to analyse the localisation of endogenous Ect2 in HeLa cells and reported the localisation of endogenous Ect2 at the spindle and midbody during mitosis. I observed the same localisation using microinjected plasmid encoding YFP-hEct2 and time-lapse video fluorescence and phase microscopy.

Figure 6-3. YFP-Ect2 is a nuclear protein in interphase and localises to the spindle, midbody and plasma membrane during mitosis.

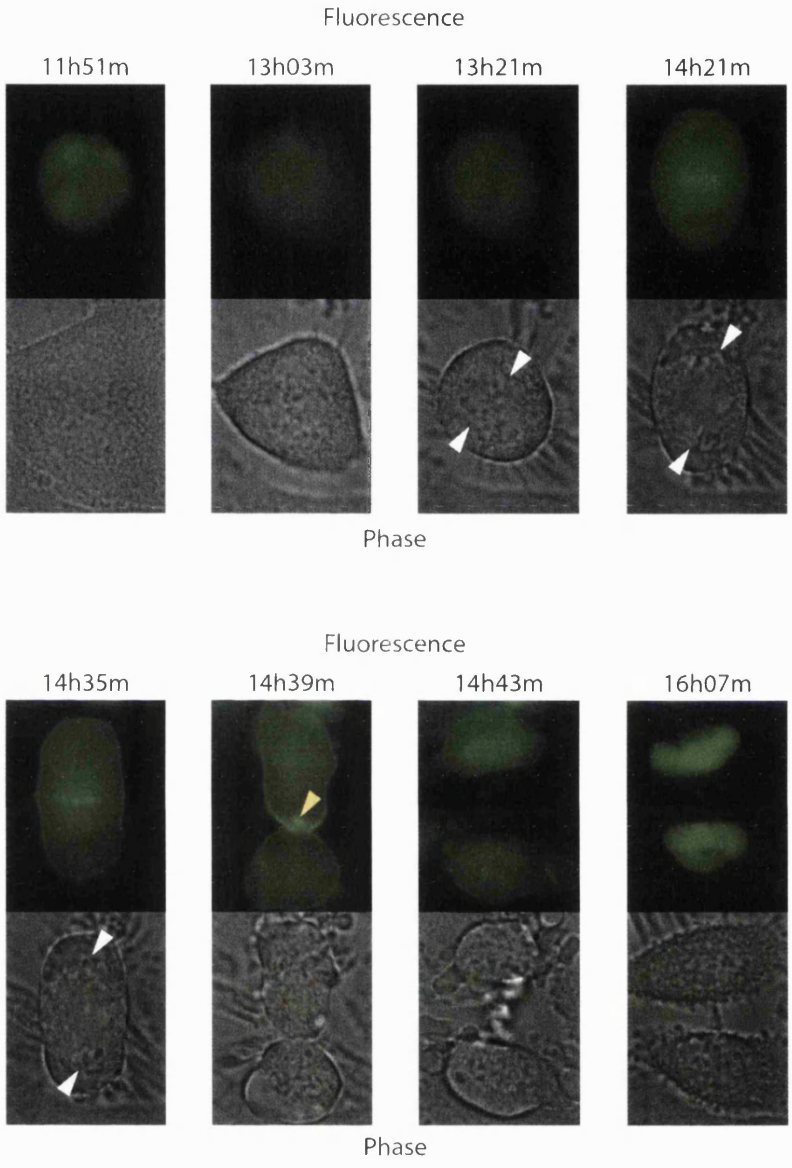


Figure 6-3. Cell cycle dependent localisation of YFP-Ect2 in HeLa cells. HeLa cells were arrested in late G1 using a double-thymidine block. Individual cells were injected with plasmid expressing YFP-Ect2 from the EF2α promoter. Cells were incubated under 7% CO_2 at 37°C while images were collected at 5 minute intervals using an Axiovert 135TV microscope. YFP-Ect2 was localised diffusely throughout the nucleus and appeared to be concentrated in the nucleoli during G2. As the cell entered mitosis YFP-Ect2 became distributed throughout the cytoplasm. At t=13h21m this cell was at metaphase and YFP-Ect2 could be seen to be localised on the spindle. As the cell progressed to telophase YFP-Ect2 became increasingly concentrated at the site of cytokinesis and just prior to cytokinesis at t=14h35m there was a distinct localisation of YFP-Ect2 at the plasma membrane which persisted until the two daughter cells had entered G1 and reformed their nuclei. During cytokinesis YFP-Ect2 became predominantly localised at the cleavage furrow and a foci of flurescence consistently appeared in one daughter cell. As the daughter cells re-formed their nuclei YFP-Ect2 remained diffuse in the cytoplasm and localised at the plasma membrane. As cells progressed through G1 YFP-Ect2 became nuclear once again and began to accumulate in nucleoli.

Miki *et al* also reported that expression of the N-terminal section of hEct2 in HeLa cells caused a defect in cytokinesis and resulted in 60% of transfected cells becoming multinucleated. The N-terminal fragment that they expressed lacked the C-terminal Rho-GEF catalytic and PH domains (Tatsumoto *et al.*, 1999). They attribute the failed cytokinesis to a lack of catalytic activity through the loss of the Rho-GEF domain in their mutant and although this is a reasonable explanation for their results, it is also possible that the failure in cytokinesis is due to a failure of their mutant to localise appropriately to the plasma membrane. This may have not been an explanation that occurred to them as they were unable to observe the plasma membrane localisation that I did. Further experiments investigating the role of the PH domain in regulating cellular localisation of hEct2 would be required.

The localisation data shown here would support a model in which Ect2 function is required for correct cytokinesis, although it would also suggest that the localisation of hEct2 at the plasma membrane may be important. It is possible to perceive a model where the cell cycle dependent activation of the Rho-GEF activity of hEct2 is required and that hEct2 then requires to be localised to the plasma membrane where cytoskeletal rearrangements are required to form an acto-myosin contractile ring for cytokinesis. It seems unlikely that this re-localisation would be by passive diffusion once the nuclear envelope has been dissolved because the plasma membrane localisation shown here did not occur until late in mitosis, suggesting that this is a controlled event.

6.3.5. Why does hEct2 contain two BRCT domains?

The SMART algorithm identified two BRCT domains in the N-terminus of hEct2. The BRCT domain is a protein-protein and protein-DNA interaction module found in a diverse number of proteins involved in DNA damage repair and DNA metabolism (Dulic *et al.*, 2001; Makiniemi *et al.*, 2001; Moore *et al.*, 2000; Taylor *et al.*, 1998; Thornton *et al.*, 2001). The role of the BRCT domain in hEct2 is perplexing as hEct2 appears to act in regulating cytokinesis and may play an important role at the plasma membrane to achieve this. That hEct2 is sequestered in the nuclei, and possibly nucleoli, of interphase cells could be to prevent it from acting prematurely at the plasma membrane, and this would be supported by recent reports showing that nucleoli act as sites of sequestration of important cell cycle regulators (Visintin *et al.*, 1999). Alternatively, hEct2 could have other

cytokinesis independent functions that require its sequestration in the nucleus during interphase.

BRCT domains are found predominantly in proteins involved in cell cycle checkpoint functions responsive to DNA damage and this prompted me to investigate whether hEct2 localisation was responsive to DNA damage. To achieve this I transfected HeLa cells cultured to 50% confluence on cover slips with plasmid pEF MCS YFP:hEct2 and 24 hours post-transfection γ-irradiated them with 10 Gy using a ^{137}Cs source. The cells were incubated at 37°C with 7% CO_2 for an hour before I fixed the cells and processed them for immunofluorescence using an affinity purified anti-Rad51 antibody (a gift of Dr. Steve West, ICRF, Clare Hall.) Rad51 is involved in the repair of double-strand breaks and localises to those breaks, forming nuclear foci within an hour of the damage occurring and acting as a good marker for DNA damage (Haaf *et al.*, 1995; Liu and Maizels, 2000; Tashiro *et al.*, 2000). In cells which had undergone irradiation and had a large number of double-strand breaks as visualised by the formation of Rad51 foci, there was no difference in the localisation of YFP-hEct2 than in cells which were not irradiated and had few Rad51 foci (data not shown). These data suggested that YFP-hEct2 did not relocalise in response to double-strand DNA breaks.

6.4. Clone EA2E2 is nucleophosmin.

Sequence analysis identified clone EA2E2 as nucleophosmin. Nucleophosmin, also known as numatrin, B23 and NO38 is a multifunctional nucleolar phosphoprotein of M_r 35-40 kDa with a pI of 5.1 (Feuerstein *et al.*, 1988). Nucleophosmin has been shown to bind nucleic acids (Dumbar *et al.*, 1989), possess molecular chaperone activity (Szebeni and Olson, 1999), be associated with maturing pre-ribosomal ribonucleoprotein particles (Yung *et al.*, 1985), possess a nuclear localisation signal, engage in nucleocytoplasmic shuttling (Borer *et al.*, 1989), self-associates into oligomers and possesses ribonuclease activity (Herrera *et al.*, 1995; Savkur and Olson, 1998). These activities have suggested that nucleophosmin is involved in the assembly of pre-ribosomes. Nucleophosmin has also been reported to be phosphorylated by casein kinase II in interphase and by CDK1 in mitosis (Jiang *et al.*, 2000), and has been shown to localise to centrosomes during mitosis (Zatsepina *et al.*, 1999). Nucleophosmin has also been shown to contain a sequence which is a consensus CDK phosphorylation site, and which is phosphorylated *in vivo* (Jones *et al.*, 1981). Some

of these properties suggest that nucleophosmin could be involved in the regulation of cell cycle events.

Shortly after my identification of nucleophosmin as an *in vitro* substrate of CDK2:cyclin A3 a report by Okuda *et al* showed that nucleophosmin was a protein present in centrosome preparations and which was phosphorylated by CDK2:cyclin E (Okuda *et al.*, 2000). They also showed that nucleophosmin was present in G1 but not G2-phase centrosomes, and that the CDK2:cyclin E dependent phosphorylation induced the loss of nucleophosmin from centrosomes. This loss of nucleophosmin was shown to be necessary for centrosome duplication although nucleophosmin reassociated with centrosomes during mitosis, an observation also made by Zatsepina *et al* (Zatsepina *et al.*, 1999).

6.4.1. YFP-nucleophosmin localises to nucleoli during G1 and G2, and is rapidly accumulated in newly forming nucleoli after mitosis.

I wanted to repeat the localisation experiments of Okuda *et al* showing nucleophosmin present at G1 centrosomes. I was unable to obtain an antibody against nucleophosmin so instead cloned the full length cDNA encoding nucleophosmin into pEF MCS YFP and transfected this construct into HeLa cells. As a control I transfected the empty pEF MCS YFP plasmid into cells in duplicate wells. Twenty-four hours following the transfection I analysed the living cells to determine the cellular localisation of YFP-nucleophosmin. Figure 6-4A shows the localisation of YFP (left panel) and of YFP-nucleophosmin (right panel). As described before YFP localises diffusely throughout the cell although begins to accumulate in cells expressing the protein to high levels. YFP-nucleophosmin localises in large spots within the nuclei of transfected cells. The phase image clearly shows the position of the nucleoli (black arrows) and the signals seen in the fluorescent image clearly localise to the nucleoli (whit arrows). It appears that in interphase cells YFP-nucleophosmin localises correctly to nucleoli.

The pEF MCS YFP:nucleophosmin plasmid was also injected into cells released from a double thymidine block. In injected cells traversing G2, YFP-nucleophosmin was also clearly a nucleolar protein. As cells entered mitosis and the nuclear envelope dissolved and nucleoli were disassembled YFP-nucleophosmin became dispersed in increasingly small particles throughout the cell. After cytokinesis when the daughter cells

entered G1 nucleophosmin was rapidly imported into the nucleus. YFP-nucleophosmin accumulated in nucleoli as cells progressed through G1 as these structures reformed (not shown). Overexpression of YFP-nucleophosmin appeared to have no adverse affects on the cell cycle and the fusion protein behaved as the endogenous protein is reported to (Ochs *et al.*, 1983).

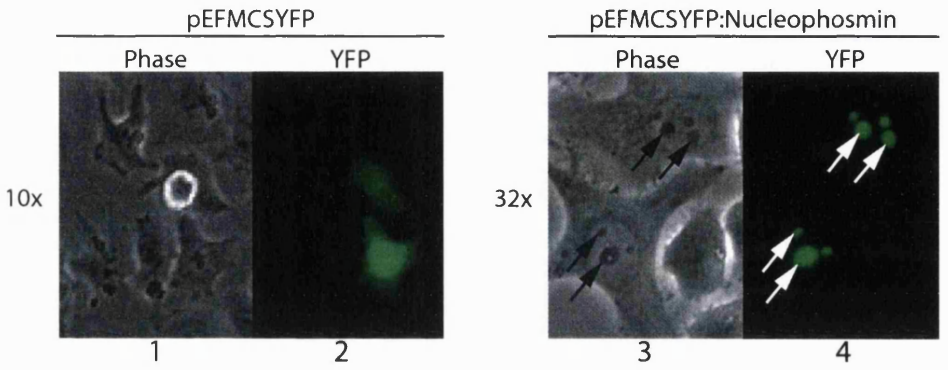
6.4.2. YFP-nucleophosmin does not localise to G1 centrosomes.

It was often observed that a small dot of fluorescence was observed peripheral to the nucleus in cells expressing YFP-nucleophosmin. I suspected that this was possibly centrosome associated YFP-nucleophosmin. To investigate this I again transfected HeLa cells with pEF MCS YFP:nucleophosmin and 24 hours after the transfection processed the cells for immunofluorescence to detect γ-tubulin. Figure 6-4B shows the localisation of both YFP-nucleophosmin and γ-tubulin in G1 HeLa cells. The phase images in panels 5 and 9 of figure 6-4B clearly show the nuclei and the presence of two discreet nucleoli (upper and lower panels). DNA was stained with Hoechst and showed the position of the nuclei and nucleoli (panels 6 and 10). The fluorescence images in panels 7 and 11 show the localisation of YFP-nucleophosmin to the nucleoli and also to a faint dot on the periphery of the nuclei (yellow arrows). The secondary antibody used to detect the anti-γ-tubulin antibody was Cy3 conjugated and it can be seen clearly that the fluorescent signals in panels 8 and 12 do not co-localise with YFP-nucleophosmin.

It is difficult to determine whether a cell is in G1 or G2, and the usual approach to determine this is to perform an anti-cyclin A immunostain. I was unable to use this approach here as I was unable to use neither a FITC, TRITC or Cy3 conjugated secondary antibody for the detection of cyclin A due to the use of YFP-nucleophosmin and Cy3 already. G2 cells have duplicated centrosomes however and this can be a useful marker for this cell cycle stage. When visualised directly at 100x magnification it was possible to determine that two centrosomes were present but had not yet separated. Unfortunately, the quality of the images captured by the camera was not sufficiently high to allow resolution of duplicated centrosomes. In the images shown here there was clearly a single centrosome present when compared to G2 cells containing duplicated centrosomes. In over 100 cells analysed I was unable to find a single cell in which the YFP-nucleophosmin signal could be overlaid with the Cy3 signal from the γ-tubulin detection.

Figure 6-4. Localisation of YFP-Nucleophosmin in HeLa cells.

A. YFP-Nucleophosmin localised to nucleoli during interphase in HeLa cells.



B. YFP-Nucleophosmin did not localise to centrosomes in G1 HeLa cells.

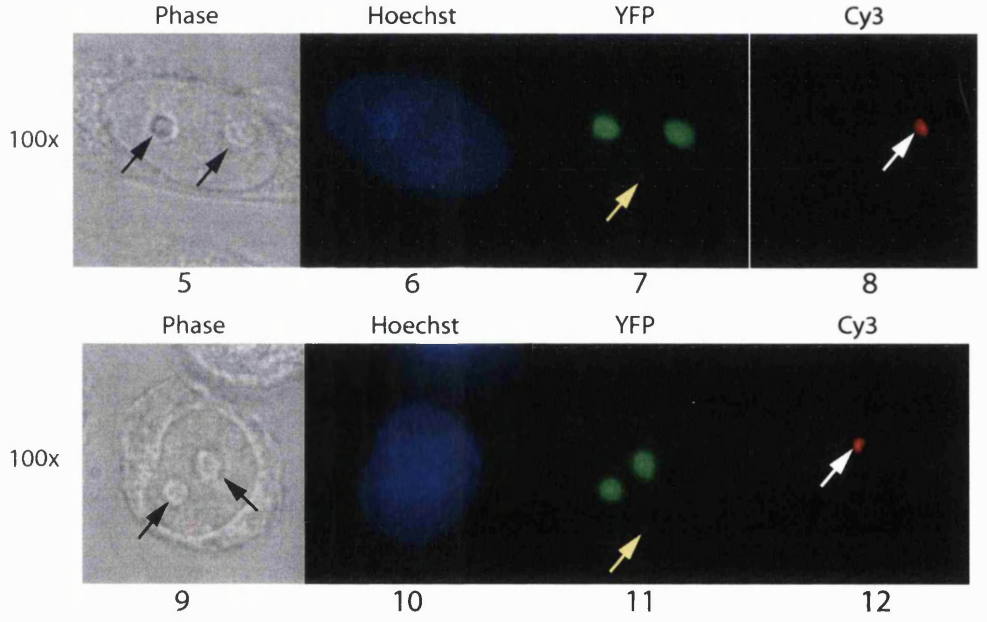


Figure 6-4. YFP-Nucleophosmin localises to nucleoli but not to centrosomes in HeLa cells. **(A).** HeLa cells were grown to 50% confluence in 6 well dishes and transfected with2 µg/well of a plasmid encoding YFP-Nucleophosmin (pEFYFPMCS:nucleophosmin) using TransitLT reagent. Twenty four hours post-transfection live cells were imaged using a Zeiss Axiovert 10 inverted microscope and a Coolsnap Raper colour CCD camera. Transfection efficiencies were approximately 40% and two representative cells are shown. Arrows mark nucleoli. **(B).** Cells were grown on coverslips in 6 well dishes to 50% confluence and transfected with 2 µg/well of pEFMCSYFP:nucleophosmin using TransitLT reagent. Twenty four hours post-transfection cells were processed for immunofluorescence using an anti-γ tubulin monoclonal antibody. Primary antibody was deteced using Cy3-coupled rabbit anti-mouse secondary antibody. Cells were mounted on slides and imaged using a Zeiss Axioplan microscope and a Hamamatsu black and white CCD camera. Cells incubated with secondary antibody alone showed no Cy3 staining.

6.5. The MTA1 protein possesses Bromo-adjacent homology, SANT, ELM2 and SH3 domains as well as Zinc-finger and leucine zipper motifs.

Clone BC6H8 was identified as Metastasis-associated protein 1 (MTA1) by homology search using the BLAST algorithm at NCBI. The clone I had obtained lacked the N-terminal 808 nucleotides, so I attempted to obtain a full length cDNA using RT-PCR. I was unable to obtain a full length cDNA from the library I had constructed and so obtained a full length cDNA from Dr. Yasushi Toh, Kyushu University, Fukuoka, Japan. The deduced amino acid sequence of the MTA1 cDNA was analysed using SMART to detect protein domains and motifs. Figure 6-5 shows a schematic representation of the domains identified using the SMART algorithm and also those suggested by other groups.

The BAH domain was initially identified due to its proximity to the bromodomains present in the polybromo; protein a component of a large nuclear complex. The BAH domain has been identified in a number of proteins involved in gene regulation and repression (Callebaut *et al.*, 1999). Although many of the proteins which have been found to contain BAH domains are of unknown function, the BAH domain is usually found in proteins containing other domains implicated in transcriptional regulation (SET, PHD and bromodomains) or DNA binding motifs (AT hooks, HMG boxes and Zinc fingers) (Goodwin and Nicolas, 2001). It is believed that BAH domains are likely to act as protein-protein interaction modules. In agreement with this observation it is not surprising to find that MTA1 also contains a leucine zipper and a GATA type Zinc finger, both of which have been implicated in DNA binding, although both have also been found to be involved in promoting protein dimerization (Alber, 1992; Chevray and Nathans, 1992).

MTA1 also contains a SANT domain; so called due to it also being found as a DNA binding domain in SWI3, ADE2, N-CoR and TFIIIB. SMART also identified a poorly defined domain called ELM2 because of its conservation between MTA1 and Egl-27. Given the abundance of domains and motifs involved in either DNA binding or protein-protein interactions which have been identified in MTA1 it is possible that MTA1 is part of a multi-protein complex which is involved in gene regulation.

Figure 6-5. Domain structure of MTA1.

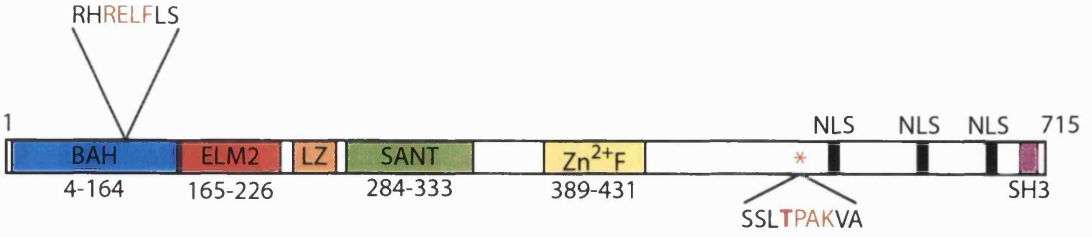


Figure 6-5. Schematic representation of domain structure of MTA1. A conceptual translation of the sequence of the full length cDNA was analysed using the Simple Modular Architecture Research Tool (SMART) at EMBL. This software predicted the presence of an N-terminal Bromo-adjacent homology domain (BAH, blue), an Egl-27,/MTA1 homology 2 domain (red), a SWI3 /ADA2,/N-CoR/TFIIIB DNA binding domain (SANT, green) and a GATA Zinc finger (Zn^{2+}, yellow). There is a Cy motif in the BAH domain and good consensus CDK phosphorylation site in the C-terminus. Other groups have identified an n-terminal leucine zipper motif (orange), a C-terminal SH3 domain (pink) and three putative NLSs (black).

6.5.1. Localisation of overexpressed YFP-MTA1 in HeLa cells.

At the time I cloned MTA1 there was no data available on its cellular function. It was originally identified as part of a subtractive hybridisation screen searching for genes upregulated in metastatic rat mammary adenocarcinoma cells (Toh *et al.*, 1994). MTA1 possesses three putative NLSs although none have been tested empirically, and as there is such an abundance of DNA binding motifs in MTA1 it seemed likely that it would be a nuclear protein. It was an exciting candidate to be an *in vivo* substrate of CDK2 as it was also implicated in transcriptional control. To test whether MTA1 was a nuclear protein and to investigate whether its localisation was regulated in a cell cycle dependent manner I cloned the full length cDNA into pEF MCS YFP and used this construct to analyse the localisation of MTA1 in HeLa cells.

HeLa cells were transfected with pEF MCS YFP:MTA1 and 24 hours post-transfection they were analysed to determine the localisation of YFP-MTA1 which is shown in figure 6-6. YFP-MTA1 had two distinct types of localisation in HeLa cells and there was an even distribution between the two localisation patterns. In the upper panel of figure 6-6, YFP-MTA1 is seen to localise diffusely throughout the nuclei of transfected cells and is clearly excluded from the nucleoli. There is no accumulation of YFP-MTA1 in these transfected cells. The second localisation pattern that I repeatedly observed is shown in the lower panel of figure 6-6. In these cells YFP-MTA1 was seen to be localised in large nuclear foci and was not distributed diffusely throughout the nucleus. The foci were clearly seen not to be nucleoli when the phase and YFP images were merged (not shown). I suspected that these two different types of localisation may be due to cell cycle differences, although this would be curious as most cells in an asynchronous culture are in G1 and I would not have expected to have seen equal numbers of transfected cells displaying the two localisation patterns.

To investigate whether these differences in expression patterns were due to cell cycle differences, plasmid pEF MCS YFP:MTA1 was microinjected into cells which had been released from a double-thymidine block. These cells were then imaged as they progressed through the cell cycle. As was observed with the transfection experiments there were both types of localisation pattern, although the diffusely nuclear pattern was predominant. As the cells progressed through the cell cycle there was no change in either

Figure 6-6. YFP-MTA1 localises to the nuclei of HeLa cells.

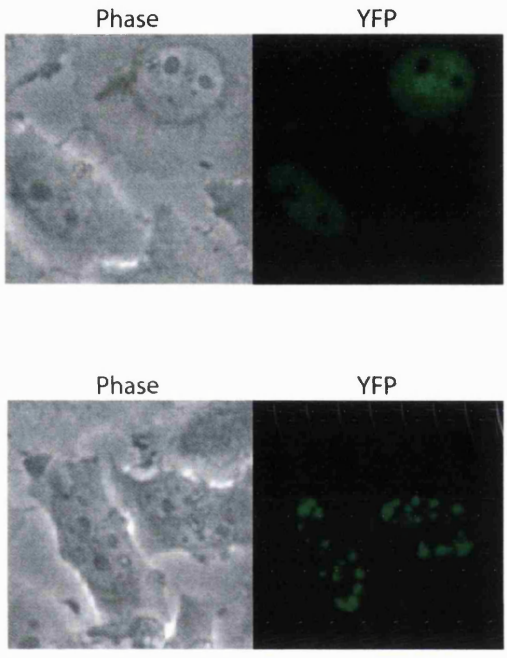


Figure 6-6. YFP-MTA1 localises diffusely through the nuclei or in large nuclear speckles of transfected cells. HeLa cells were grown to 50% confluence before being transfected with pEF MCS YFP:MTA1. Twenty-four hours post-transfection cells were imaged to visualise YFP-MTA1. Upper panel: In some transfected cells YFP-MTA was localised diffusely through the nucleus but excluded from the nucleoli. Lower panel: In some cells YFP-MTA1 was localised in large nuclear speckles distributed throughout the nucleus but appeared to not be associated with nucleoli.

of the localisation patterns however (not shown). This experiment showed that there appeared to be no change in the localisation of YFP-MTA1 that was dependent on cell cycle stage. It is possible that the different expression patterns which I observed were due to differences in the expression levels of YFP-MTA1. The delivery of plasmid to cells by microinjection is much more controlled than in transfections and it is possible that the lower level of plasmid delivered to the cells by microinjection lead to lower expression of YFP-MTA1 and the resulting diffusely nuclear localisation. In contrast, during transfection some cells will receive more plasmid than others and therefore these higher levels of plasmid could lead to higher levels of expression and cause the accumulation of YFP-MTA1 in nuclear speckles. Although this is not a hypothesis I tested by adjusting the amount of plasmid delivered during microinjection experiments, it raised the interesting possibility that localisation of MTA1 may be regulated by its expression level, which in turn could be controlled at the level of transcription, translation, or mRNA or protein stability.

6.5.2. Production of polyclonal antibodies to MTA1.

When I began the analysis of the localisation of YFP-MTA1 there was no anti-MTA1 antibody available. To generate a reagent to investigate whether MTA1 was cell cycle regulated I wanted to generate polyclonal antibodies against this protein. I attempted to express both full length and an N-terminally truncated MTA1 (MTA1ΔN269) as hexahistidine tagged fusion proteins in *E. coli* which could be purified and used as immunogens. Neither full length MTA1 or MTA1ΔN269 protein expressed well in *E. coli* and the material which was expressed proved to be insoluble (not shown). As an alternative to using MTA1 purified from *E. coli* I decided to have two peptides synthesised which were present in the C-terminus of MTA1 and which were rich in aromatic residues. The first peptide spanned the residues 614-631 of MTA1 (HMGPSRNLLLNGKSYPTK) and terminated in a lysine residue to allow easy coupling via dimethylpimelidate (DIMPIM). The second peptide spanned the residues 645-662 (RRRMNWIDAPGDVFYMPK) and terminated in a lysine residue to allow coupling to KLH via DIMPIM. The peptides were synthesised and purified by the Peptide Synthesis Laboratory at 44 Lincolns Inn Fields, London. I dissolved the peptides in water and

coupled them directly to KLH, and these conjugates were used to immunise rabbits by the antibody production facility at Clare Hall.

I received the final sera from the rabbits immunised with the peptides three months after the first immunisations. The sera generated from the rabbit immunised with the peptide spanning residues 614-631 of MTA1 was called TD1 and the sera from the rabbit immunised with peptide spanning residues 645-662 was called TD2. I tested the sera in an ELISA assay for their ability to recognise the immunogen peptides. Both TD1 and TD2 recognised the peptides which they were raised against although the affinity of the antibodies generated was not particularly high (not shown). TD1 did not recognise the peptide spanning residues 645-662 of MTA1 and TD2 did not recognise the peptide spanning residues 614-631. As these antibodies did not react with their antigens with particularly high affinity I was advised to affinity purify them (Julian Gannon, personal communication).

I coupled both peptides to CnBr-activated Sepharose and used the peptide-coupled Sepharose to prepare affinity columns. Antibodies recognising the peptides were purified from 30 ml of sera TD1 and TD2. The affinity purified antibodies were used to detect MTA1 in cell extracts prepared from HeLa cells. Unfortunately, the affinity purified antibodies recognised at least 15 bands on Western blots and it was difficult to determine which of these might correspond to MTA1. I experimented with different buffers in an attempt to reduce the number of bands that both TD1 and TD2 recognised but to no avail. These antibodies would be of little use in characterising the cell cycle regulation of MTA1.

6.5.3. Immunolocalisation of endogenous MTA1 in HeLa cells.

While I was localising overexpressed YFP-MTA1 in HeLa cells and waiting for the sera from the rabbits immunised with the two MTA1 peptides to become available, two commercially available antibodies became available from Santa Cruz biotechnology (SC-9445 and SC-9446). I obtained these antibodies and used them to determine the cellular localisation of endogenous MTA1. HeLa cells were grown on coverslips to 50% confluency before being fixed for immunofluorescence to detect endogenous MTA1 using both SC-9445 and SC-9446. DNA was stained with Hoechst 33258.

Figure 6-7. Immuno-localisation of endogenous MTA1 in HeLa cells.

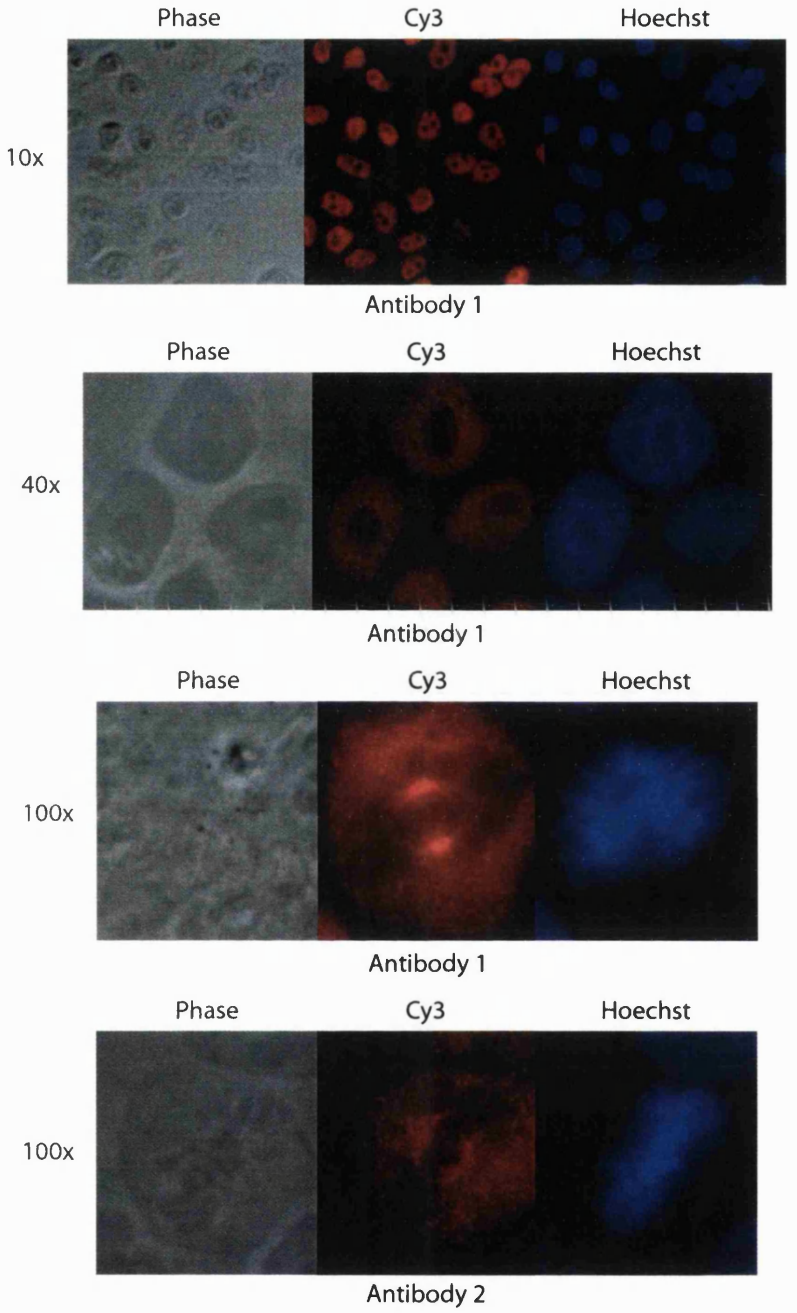
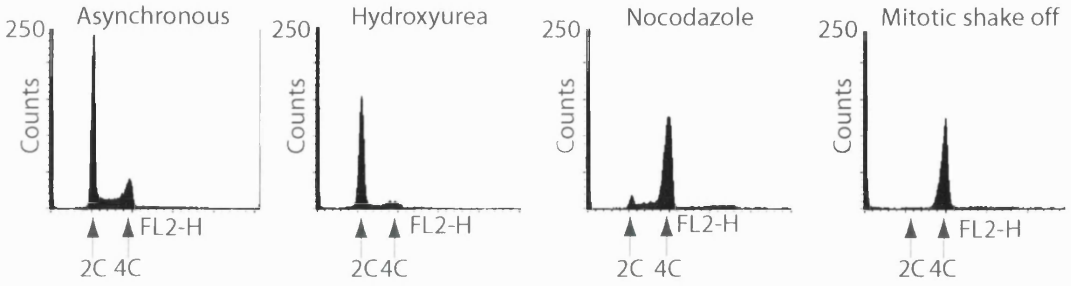


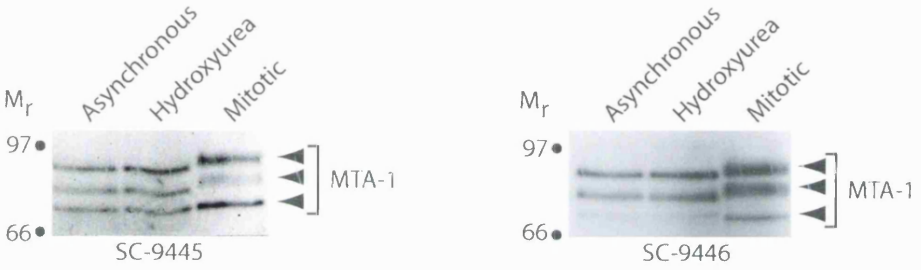
Figure 6-7. MTA1 is a nuclear protein in interphase and localises to a spindle like structure during mitosis. HeLa cells were grown on coverslips in E4 medium containing 10% FCS and were fixed with 2% paraformaldehyde. Cells were washed three times and blocked with 10% FCS, washed three times and incubated with anti-MTA1 affinity purified goat antibody SC-9445 (upper three panels) or SC-9446 (lower panel) for one hour at room temperature. Cells were washed three times and primary antibodies were detected using Cy3 coupled donkey anti-goat secondary antibodies. DNA was stained with Hoechst 33258, and coverslips were mounted onto slides for microscopy.

Figure 6-8. MTA1 is a mitotic phosphoprotein.

A. FACS analysis of cell cycle arrested HeLa cells.



B. Anti-MTA1 Western blots of extracts prepared from arrested HeLa cells.



C. Western blot of phosphatase treated extracts prepared from cell cycle arrrested HeLa cells.

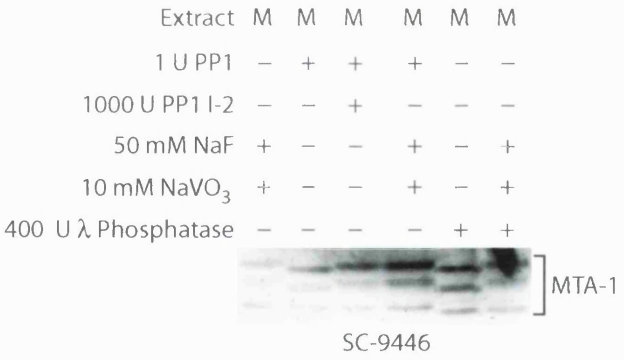


Figure 6-8. MTA1 is phosphorylated during mitosis. **(A).** Cells were treated with no drug, 2 mM hydroxyurea or 1 μM nocodazole for 16 hours. Mitotic cells were shaken from a duplicate plate of nocodazole treated cells. Cells for FACS analysis were fixed with methanol and treated with RNAse A for 30 mintues at 37°C. **(B).** Extracts were prepared by lysing cells in a buffer containing 1% NP-40 on ice for 30 minutes. The lysed cells were spun in a microcentrifuge at 14 krpm for 15 minutes at 4°C. 20 μg of soluble protein was resolved by 12.5% SDS-PAGE and transferred to nitrocellulose membrane. Western blots were performed using antibodies SC-9445 and SC-9446 from Santa Cruz Biotechnology according to their directions. **(C).** Extracts of mitotic cells were prepared as before and incubated in the presence of the indicated phosphatases or phosphatase inhibitors. PPI is protein phosphatase 1, PP1 I-2 is a specific inhibitor of protein phosphatase 1. 20 μg of protein was resolved by12.5% SDS-PAGE and transferred to nitrocellulose membrane. A western blot was performed using antiobody SC-9446 as described before.

cells containing a low number of G1 and S-phase cells showed that MTA1 in these extracts had an electrophoretic mobility similar to MTA1 in extracts prepared from asynchronously growing or hydroxyurea arrested cells. These observations suggested that MTA1 was a mitotic phosphoprotein and that there was a potent phosphatase present in G2 or S-phase cell extracts which efficiently removed mitotic phosphorylations from MTA1. The physiological role of the mitotic phosphorylation described here and the kinase catalysing the phosphorylation remain undetermined.

6.6. A streptavidin-biotin capture technique to identify substrates of CDK2:cyclin A3.

I wanted to be able to show that the proteins identified in the primary screen were directly phosphorylated by CDK2:cyclin A3. It was unrealistic to expect that I would be able to express all of the proteins I had identified using a simple expression system such as *E. coli*. I had also met with limited success when I attempted to express KIAA0144 and MTA1 in *E. coli*. Instead I developed a novel method to show that CDK2:cyclin A3 directly incorporated [^{32}P] from γ[^{32}P]-ATP into purified substrates. I had seen that biotinyl-lysyl tRNA was available (Promega corporation), and that this could be incorporated into proteins using the rabbit reticulocyte transcription-translation systems.

I transcribed and translated the clones I had identified as substrates of CDK2:cyclin A3 in the presence of biotinyl-lysyl tRNA. I purified the translation reaction products on monomeric streptavidin agarose. I washed the affinity resin extensively before incubating it in the presence of buffer containing CDK2:cyclin A3 and γ[^{32}P]-ATP. The proteins captured on the streptavidin agarose were eluted by boiling in SDS-sample buffer and separated using SDS-PAGE. The kinase reaction products were visualised by autoradiography. Figure 6-9 shows the autoradiogram of the kinase reaction products. As a positive control I used Cdc6 translated in the presence of biotinyl-lysyl tRNA. Figure 6-9 shows that [^{32}P] was incorporated into a protein of molecular weight identical to phosphorylated Cdc6, but that this protein was not phosphorylated in a control translation reaction purified on streptavidin agarose and which had been incubated with CDK2:cyclin A3 and γ[^{32}P]-ATP.

Figure 6-9. Phosphorylation of *in vitro* transcribed and translated biotinylated clones using CDK2:cyclin A3.

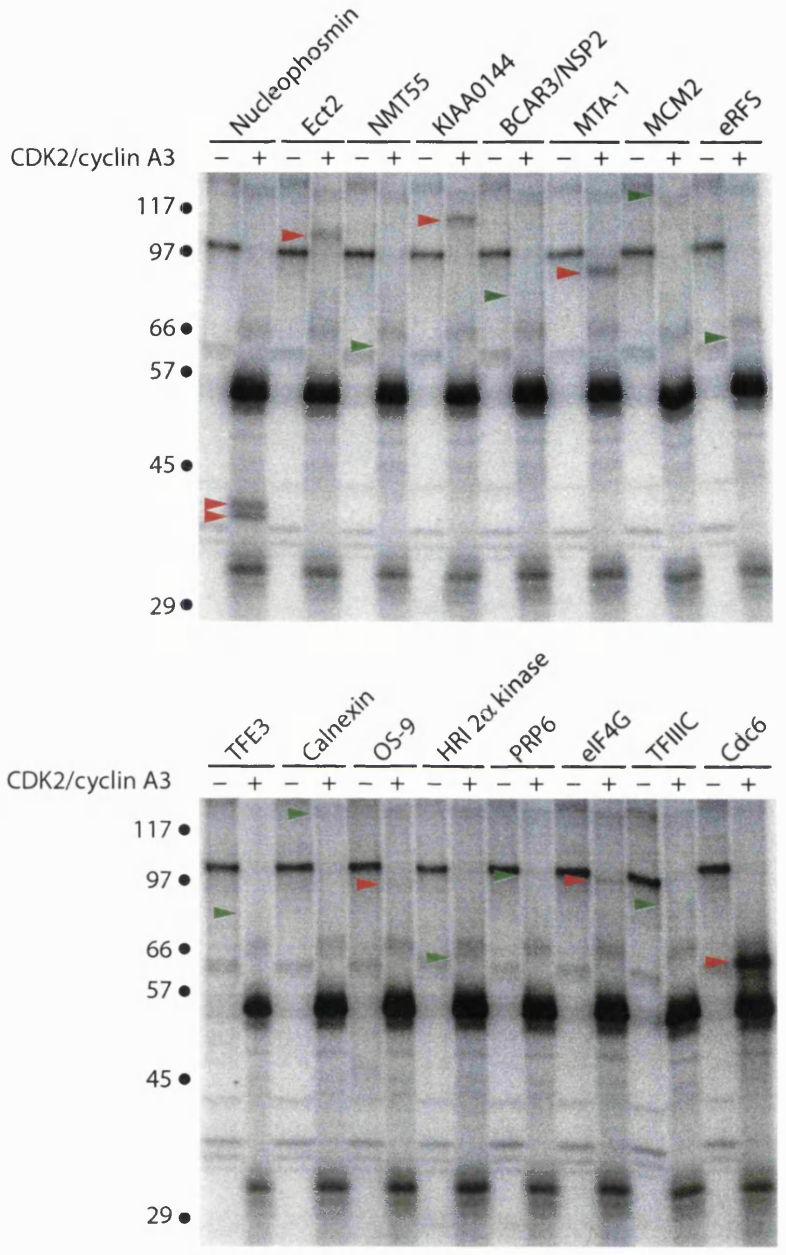


Figure 6-9. Direct phosphorylation of clones using CDK2:cyclin A3. Clones identified as substrates of CDK2:cyclin A3 were transcribed and translated in the presence of biotinyl-lysyl tRNA. Translation reactions were split in half and diluted to 200 µl with a buffer containing 0.1% NP-40 and 250 mM NaCl. Biotinylated proteins were harvested by the addition of 5 µl Softlink^{TM} monomeric streptavidin agarose and incubation on ice for 30 minutes. The resin and captured proteins were harvested by brief centrifugation and washed 3 times in the dilution buffer. The resin pellet was resuspended in 20 µl of kinase buffer containing 400 µM ATP, 1 µCi $\gamma[^{32}P]$ ATP. To 10 µl of this mix 110 nM GST CDK2:cyclin A3 was added while the other 10 µl lacked kinase. Reactions were incubated at 30°C for 30 minutes. Reaction products were separated by SDS-PAGE and the phosphorylated proteins were visualised using autoradiography. Arrowheads show where the indicated protein should migrate. Red shows the incorporation of $[^{32}P]$; green indicates a lack of incorporation of $[^{32}P]$.

The arrowheads indicate the predicted electrophoretic mobility of the proteins tested. Red arrowheads show where there is an incorporation of [^{32}P] and green arrowheads show where there is not. It has been documented that some proteins are non-functional when they contain biotin-lysine (Promega Corporation Publication) and it is possible that the proteins which are not phosphorylated in this assay are inappropriately folded due to disruption of internal structures, or that surface lysine residues important for protein-protein interactions are blocked by the presence of biotin. These data showed that nucleophosmin, hEct2, KIAA0144, MTA1, OS-9, eIF4G-1 and Cdc6 were phosphorylated *in vitro* directly by CDK2:cyclin A3. Surprisingly MCM2 was not phosphorylated in this assay, although this could be due to incorrect folding or the modification of surface lysines normally important in making protein-protein interactions.

6.7. Comparing substrate specificity between cyclins A and B (CDK2 vs CDK1).

An interesting question in cell cycle research is the substrate specificity of different CDK:cyclin complexes, that is, does CDK2:cyclin A phosphorylate a different set of proteins from CDK1:cyclin B? I had identified a number of *in vitro* substrates of CDK2:cyclin A3 and was thus in a strong position to be able to address whether the substrates I had identified could also be phosphorylated by other CDK:cyclin complexes. The main reason for using CDK2:cyclin A3 complexes to perform the initial screening was its ease of purification in an active recombinant form. Other CDK:cyclin complexes are not so easy to purify in a similar form (Wilhelm *et al.*, 1997). An alternative is to use an extract prepared from a source that is high in the kinase activity of interest.

6.7.1. Using frog egg extracts as a source of kinase activity to investigate substrate specificity of CDKs.

Frog eggs are arrested at metaphase of meiosis II (CSF arrest) with high levels of active CDK1:cyclin B. Extracts prepared from these cells in the presence of EGTA remain arrested at this point, and these extracts can be used as a source of CDK1:cyclin B activity (Shamu and Murray, 1992). Alternatively, addition of Ca^{2+} to egg extracts arrested at metaphase of meiosis II will degrade CSF and cause them to enter interphase, with low levels of active CDK1:cyclin B. Given the difficulties associated with preparing large

amounts of active CDK1:cyclin B I decided to use CSF extract as a source of CDK1:cyclin B and to compare the phosphorylation of the CDK2:cyclin A3 substrates in CSF and interphase extracts.

I transcribed and translated the substrates of CDK2:cyclin A3 in rabbit reticulocyte lysate in the presence of [^{35}S]-methionine. 1 μl of each of the translation reaction products were incubated in the presence or absence of either 100 ng CDK2:cyclin A3, or 1 μl CSF or interphase egg extracts as sources of kinase. The egg extracts were prepared by my colleague Dr. Jonathan Moore. Jon tested that the CSF extract was arrested by observing that added demembranated sperm formed individually condensed chromosomes on a spindle like structure. He also tested the interphase extracts by observing that added demembranated sperm were formed into nuclei (data not shown).

Figure 6-10A shows the results of the incubations using purified CDK2:cyclin A3 and crude CSF extract, interphase extract or interphase extract supplemented with p21. The kinase reactions were separated by SDS-PAGE and the [^{35}S] labelled proteins were visualised using autoradiography. As in the original screen, phosphorylation was visualised by the altered electrophoretic mobility of a protein through SDS-PAGE and the translation products are detected due to the incorporation of [^{35}S]. There was no [^{32}P] used in these assays. Some proteins were phosphorylated equally well by CDK2:cyclin A3 and by kinases present in CSF (hMCM2, hMTA1, hEct2, GTFIIICα, TFE3, calnexin and OS-9), but only TFE3 was phosphorylated equally well by CDK2:cyclin A3 and a protein kinase present in both CSF and interphase extracts. The kinase responsible for phosphorylating TFE3 in interphase extracts could not be inhibited by the addition of the CKI p21. These data suggest that as well as being an *in vitro* substrate of CDK2:cyclin A3 TFE3 can also be phosphorylated by a kinase present in both interphase and CSF extracts which cannot be inhibited by p21.

There were a number of proteins which were phosphorylated more efficiently by CDK2:cyclin A3 than by kinases present in CSF or interphase extracts (nucleophosmin, Heme-regulated initiation factor 2α kinase and PRP6, eIF4G-1 and BCAR3/NSP2). This suggested that these substrates are relatively specific for CDK2:cyclin A3 and are not phosphorylated (nucleophosmin, heme-regulated initiation factor 2α kinase and PRP6), or

are phosphorylated less efficiently (eIF4G-1 and BCAR3/NSP2) by the kinases present in CSF or interphase extracts. These kinases include CDK1 complexed with cyclins B1, B2, B4 and B5, MAP kinase, Polo-kinase and Aurora kinase in CSF extracts and low amounts of CDK2:cyclin E in interphase extracts. A final classification of specificity can be drawn in which there was greater phosphorylation of a substrate by kinases present in CSF extract relative to the phosphorylation observed using CDK2:cyclin A3 (KIAA0144 and HMG14/TSP-1). It was impossible to classify eRFS under these schemes as the largest translation product generated was degraded in the presence of CSF, interphase or interphase extracts supplemented with p21. A synopsis of these data is shown in figure 6-10B.

6.7.2. Some substrates of CDK1:cyclin B can also be phosphorylated by CDK2:cyclin A3.

To further extend this analysis of CDK specificity for substrates I obtained a number of clones which had been isolated in a screen for substrates of kinases present in CSF but not in interphase extracts (Lustig *et al.*, 1997; Stukenberg *et al.*, 1997). Although there are a number of kinases present in CSF and interphase extracts, these clones had also been further characterised to show that they were efficiently phosphorylated by CDK1:cyclin B *in vitro*. The constructs which the Kirschner Laboratory donated to me could be expressed *in vitro* using SP6 RNA polymerase and rabbit reticulocyte lysate (Promega Corporation). I expressed the proteins in the presence of [^{35}S]-methionine and used the translated products in kinase reactions with GST-CDK2:cyclin A3. Disappointingly, only 12 of the 20 clones translated sufficiently well to be detected using autoradiography. Eleven of these 12 clones exhibited altered mobilities in the presence of CDK2:cyclin A3 and therefore appeared to be substrates of CDK2:cyclin A3 (not shown). Interestingly, two of these proteins were XeCdc27 and XeWee1 both of which are involved in the regulation of CDK activity.

6.7.3. Are CDKs promiscuous enzymes?

These experiments showed that some proteins identified during a screen for substrates of CDK2:cyclin A3 could also be phosphorylated by protein kinases present in CSF which includes high levels of CDK1:cyclin B activity. It also showed that most proteins isolated during a screen for substrates of CDK1:cyclin B (Lustig *et al.*, 1997; Stukenberg *et al.*, 1997), could also be phosphorylated by purified, recombinant CDK2:cyclin A3. These data

raise the question of how selective the small-pool expression screening is when CDK:cyclin complexes are used. One interpretation is that there is little, if any selection for specific substrates of the kinase being used and that the proteins identified in a screen for substrates of CDK2:cyclin A3 could equally well be substrates for CDK2:cyclin E, CDK1:cyclin B or CDK1:cyclin A.

In support of this interpretation is the fact that I cloned nucleophosmin which was characterised as a physiological substrate of CDK2:cyclin E (Okuda *et al.*, 2000). Tokuyama *et al* (2001) have recently reported that nucleophosmin can be phosphorylated on T199 by CDK2 complexed with both cyclins E and A and also that it can be phosphorylated on T234 and T237 by CDK1:cyclin B (Okuda *et al.*, 2000; Tokuyama *et al.*, 2001). In cases like this, in the absence of extensive characterisation, it is difficult to distinguish which kinase(s) are exerting physiologically relevant effects on the activity of a target protein. Some reports have described a hydrophobic patch on the surface of cyclin A which could impart substrate selectivity to CDKs. However, this surface patch is conserved between many cyclins types which makes it difficult to understand how this patch could confer substrate selectivity on different cyclins (Schulman *et al.*, 1998). I decided to test whether the hydrophobic patch of cyclin A3 had any affect on the phosphorylation of the panel of substrates I had identified in my screen.

6.8. Mutating the hydrophobic patch of cyclin A3.

My colleague Jon Moore had constructed a mutant *Xenopus laevis* cyclin E in which Ile246 and Gln253 were mutated to alanines. Jon showed that these mutations impaired the ability of cyclin E to direct the kinase activity of associated CDK2 towards pRb (and Cdc6), but that they had no effect on the kinase activity of the complexes towards histone H1. Given the success Jon had with these mutations in cyclin E, I made the homologous mutations in cyclin A3. Figure 6-11 shows a representation of the crystal structure of the hydrophobic patch of cyclin A3 to highlight the residues which I mutated.

Figure 6-11. Representation of the crystal structure of the hydrophobic patch of cyclin A.

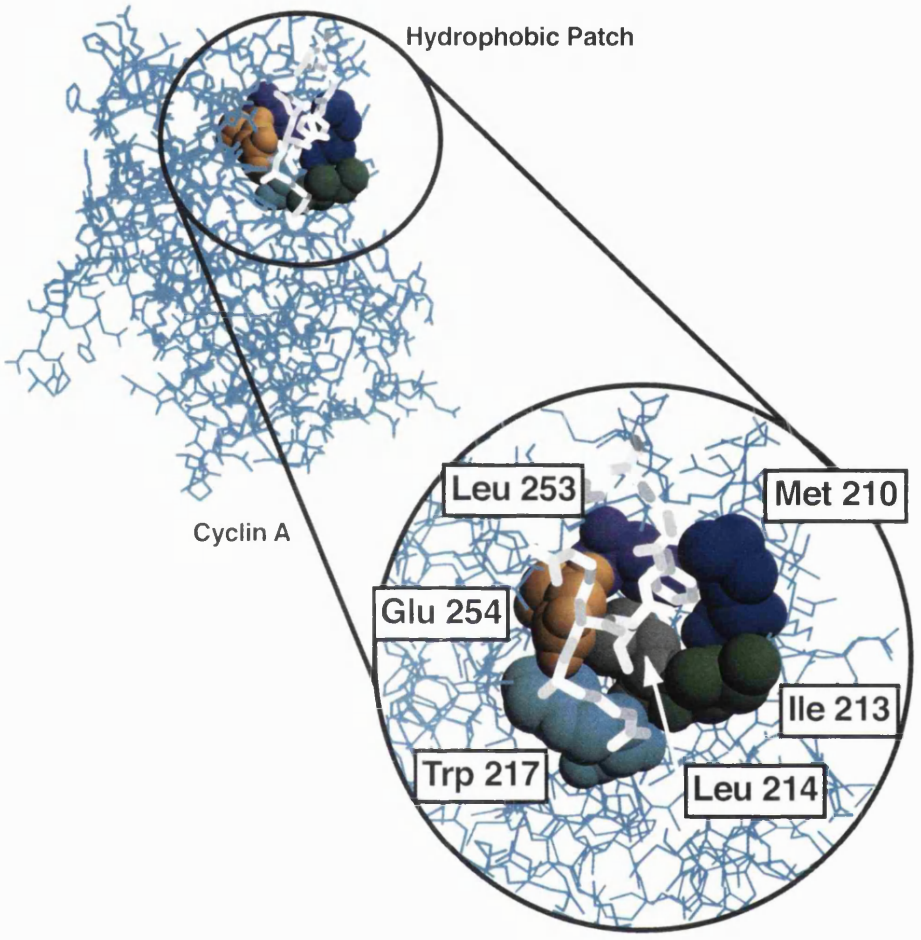
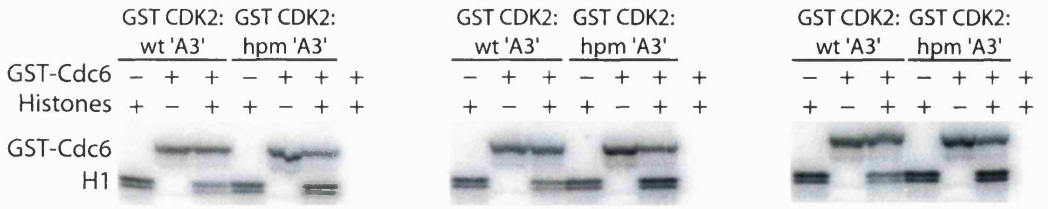


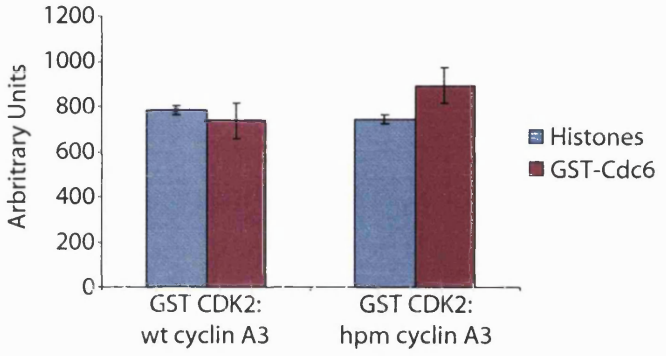
Figure 6-11. The hydrophobic patch of cyclin A is composed of Met210, Ile213, Leu214, Trp217, Glu254 and Leu 253. Analysis of the crystal structure showed that Glu254 and Ile213 both make contacts with Arg30 and Leu32 of the RXLFG motif of p27. Residues in cyclin A3 corresponding to Glu254 and Ile 213 were mutated to alanines by PCR which was expected to disrupt the interaction of cyclin A3 with RXL-containing substrates such as Cdc6.

Figure 6-12. Hydrophobic patch mutant cyclin A3 does not affect CDK2s kinase activity towards Cdc6.

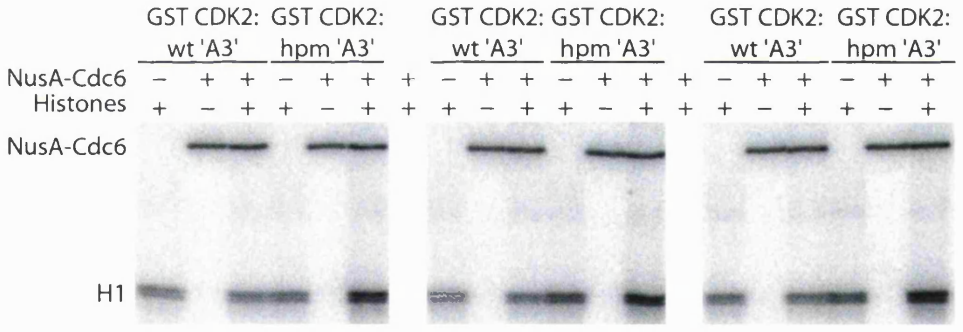
A. CDK2:wt 'A3' and CDK2:hpm 'A3' have similar kinase activity towards GST-Cdc6 and histones.



B. Quantitation of data shown in A.



C. CDK2:wt 'A3' and CDK2:hpm 'A3' have similar kinase activity toward NusA-Cdc6 and histones.



D. Quantitation of data shown in C.

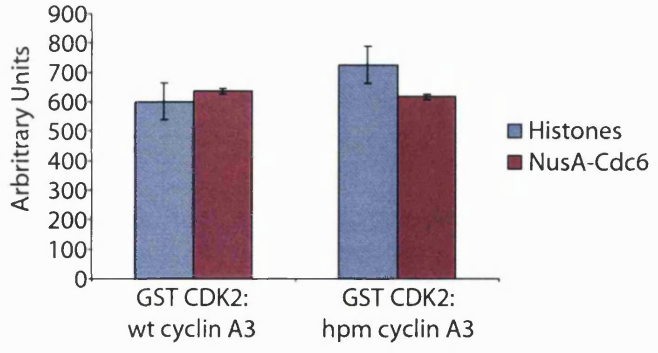


Figure 6-12. Comparison of wt cyclin A3 and hpm cyclin A3s CDK2 associated kinase activity towards histones and GST-, or NusA-Cdc6. **(A)** Kinase reactions were performed for 30 minutes at 30˚C using 5 ng of GST CDK2 complexed to either wt or hpm cyclin A3 in the presence of 0.1 μCi γ[^{32}P]-ATP and either 1 μg of GST-Cdc6 or histones or both combined. Reaction products were seperated by SDS-PAGE and visualised by autoradiography. Three duplicate experiments are shown. **(B)** Data in A were quantified using NIH image. **(C)** Kinase reactions were performed as in A except that GST-Cdc6 was substituted with NusA-Cdc6. Three duplicate experiments are shown. **(D)** Data in C were quantified using NIH image.

kinase reactions contained 5 ng GST-CDK2 complexed to either wt or hpm cyclin A3 and 1 μg of histones or NusA-Cdc6, or a mix of both in the presence of 0.1 μCi $\gamma[^{32}P]$-ATP. The kinase reaction products were separated using SDS-PAGE and visualised by autoradiography.

Again there was no difference in the ability of either GST-CDK2:wt or hpm cyclin A3 to phosphorylate histone H1 (figure 6-12C). There was also no difference in the abilities of GST-CDK2:wt or hpm cyclin A to phosphorylate NusA-Cdc6. The data in figure 6-12C were quantified using NIH image and are shown in figure 6-12D. These data showed that the I41->A41 and E48->A48 changes had no effect on the ability of cyclin A3 to recognise a substrate containing RXL motifs. This suggested that the contacts made between Ile41 and Glu48 of cyclin A3 and the arginine and the leucine of RXL motifs are not important in recruiting RXL containing substrates to CDK2:cyclin A complexes.

It was possible that kinase reactions using GST-CDK2: wt or hpm cyclin A3 were proceeding to completion under the conditions that I used. I tested whether GST-CDK2:wt cyclin A3 was more efficient at phosphorylating NusA-Cdc6 than GST-CDK2:hpm cyclin A3 I under limiting reaction conditions. Figure 6-13 shows that NusA-Cdc6 was phosphorylated with similar efficiencies by both GST-CDK2:wt and hpm cyclin A3 in kinase reactions performed across a time course of 1-6 minutes. These data were quantified using NIH image and are shown in figure 6-13B. Taken together these data suggested that the effect of the mutations I41A and E48A in cyclin A3 had little effect on phosphorylation of the RXL containing substrate Cdc6 by GST-CDK2:cyclin A3. It is possible that there are other sequence determinants in or around the hydrophobic patch which have more significant effects on the recruitment of RXL containing substrates to cyclin A. These results indicated that there is a significant difference in the way in which cyclins E and A recruit substrates to CDK. Mutation of the hydrophobic patch in cyclin E had a significant effect on the phosphorylation of an RXL containing substrate, while homologous mutations in cyclin A had little if any effect on the phosphorylation of the RXL containing substrate Cdc6.

6.8.3. Phosphorylation of *in vitro* translated Cdc6 by GST-CDK2:cyclin A3 can be inhibited by RXL containing peptides.

The hydrophobic patch of cyclin A is not defined solely by the residues Glu254 and Ile213, and it is possible that other residues in or around the patch are involved in the specific recruitment of substrate to cyclin A. I tested the ability of an RXL containing peptide derived from p27 to inhibit the phosphorylation of *in vitro* transcribed and translated Cdc6. The peptide I used spanned the region of p27 which interacts with the hydrophobic patch of cyclin A in the crystal structure of CDK2:cyclin A:p27 (Russo *et al.*, 1996).

Cdc6 was transcribed and translated *in vitro* in the presence of [^{35}S]-methionine. The labelled Cdc6 was used in kinase reactions containing either 220 or 55 nM CDK2:cyclin A3. As a control, I titrated increasing amounts of a control peptide, which happened to be derived from Cdc27, into the reactions. The reaction products were separated by SDS-PAGE and visualised using autoradiography. The effects of the Cdc27 peptide were evaluated by changes in the electrophoretic mobility of Cdc6. In the presence of the Cdc27 peptide, there was no detectable inhibition of the phosphorylation of Cdc6 by CDK2:cyclin A3 (figure 6-14A, upper and lower panels). When a peptide containing the RXL motif from the CKI p27 was titrated into the reactions there was an inhibition of the phosphorylation of Cdc6 by CDK2:cyclin A3. This inhibition of phosphorylation was dependent on the concentration of the peptide and the IC_{50} of the p27 peptide was in the range of 1-3 mM (figure 6-14B, upper and lower panels).

When the p27 peptide was titrated into kinase reactions containing CDK2:cyclin A3 and the *in vitro* translated substrate MCM2 there was no detectable inhibition of the phosphorylation. These data suggest that the p27 peptide-dependent inhibition of phosphorylation of Cdc6 by CDK2:cyclin A3 is dependent on an interaction between the hydrophobic patch of cyclin A3 and Cdc6. Although the experiments using hpm cyclin A3 showed that there was no dependence on I41 or E48 for a functional interaction between cyclin A3 and Cdc6, these experiments using the p27 peptide suggested that there are other elements in the hydrophobic patch which are important for the CDK2:cyclin A3-substrate interaction. When these elements in the hydrophobic patch were blocked with a competing peptide the phosphorylation of Cdc6 by CDK2:cyclin A3 was inhibited.

6.9. General Discussion.

In this chapter I have described the localisation of a number of the clones isolated during the screen for substrates of CDK2:cyclin A3. I have shown that the localisation of YFP-fusion proteins can be a useful way to determine if a protein is localised in the correct cellular compartment to be a candidate substrate of a CDK:cyclin complex (such as KIAA0144 was not). Also I have shown that YFP-fusion proteins in combination with microinjection and time-lapse video microscopy can be used to show that a protein is regulated in a cell cycle dependent manner in terms of its localisation (such as hEct2). It has also been shown that these methods can be used to identify new protein localisations which may be relevant to the cellular function of a protein, such as the plasma membrane localisation of hEct2 which was not detected in the immunofluorescence analysis of the localisation of hEct2.

I have used a novel streptavidin-biotin capture approach to show that [^{32}P] can be incorporated from γ[^{32}P]-ATP by a kinase into putative substrates which have proven refractory to expression in heterologous expression systems such as *E. coli*, or when substrates are present in such numbers as to make these systems an inviable option. This approach has shown that the electrophoretic mobility shifts seen in the primary screen were caused by phosphorylation and not by another protein modification. It is possible that this capture technique could be used to increase both the throughput and sensitivity of the small-pool expression screening technique that I used to identify substrates of CDK2:cyclin A3. Although I don't believe that it is possible to efficiently express more than 50-100 cDNAs using coupled transcription-translation reactions *in vitro*, it would be possible to pool several reactions containing this many clones. If each of these reactions included biotinyl-lysyl tRNA then the proteins synthesised would be biotinylated and could be captured using streptavidin agarose. If 10 translation reactions, each containing 70 primary library clones were pooled, and the biotinylated proteins were captured on streptavidin agarose, then this means that up to 10 times as many proteins could be screened. Further to this, if the kinase reactions were performed in the presence of γ[^{32}P]-ATP then this would change the selection criteria from altered electrophoretic mobility through SDS-PAGE to incorporation of [^{32}P]. This means that proteins which are phosphorylated but which do no

display altered electrophoretic mobility could also be identified. These are significant improvements to the screening technique as it currently exists.

Before I had found that I could use the incorporation of biotinylated lysine to tag *in vitro* expressed proteins, I had investigate the possibility of constructing plasmid libraries in which the cloned cDNAs could be expressed as fusions to GST using the *in vitro* transcription-translation system. Not wanting to invest more time in cDNA library construction however I constructed a plasmid which contained the T7 promoter upstream of the human influenza virus leader sequence, two copies of the c-Myc epitope, the coding sequence for GST and a *lox* recombination site.

The Cre-recombinase catalyses recombinational events between *lox* sites and so two sequence of DNA containing *lox* sites will be recombined together in the presence of Cre (Liu *et al.*, 1998a). This technique can be used to fuse plasmids together, thus placing a cDNA sequence under the control of a new promoter, or fusing it to other DNA sequences coding, for example, GST or hexahistidine tags. This approach can therefore be used to transfer the cDNA inserts from a whole library to another plasmid containing novel regulatory and tag sequences. Steve Elledge (Baylor College of Medicine, Houston, TX, USA) fused a HeLa cell cDNA library with my construct containing the GST-tag under the control of the T7 promoter. I had hoped to use this fusion library to generate pools of GST-tagged proteins which could be mixed together, purified on GSH-Sepharose and used as substrate in kinase reactions with $\gamma[^{32}P]$-ATP. Unfortunately when test constructs were expressed *in vitro*, very little translated product could be harvested using GSH-Sepharose.

I investigated the localisation of MTA1 using both immunofluorescence and a YFP-MTA1 fusion protein overexpressed in HeLa cells. It appeared that the localisation of YFP-MTA1 was dependent on expression level, and from that the analysis of the endogenous levels showed that this was a diffusely nuclear protein. In mitosis MTA1 was excluded from chromatin and localised partially to a spindle-like structure. These experiments showed the importance of being able to verify the results obtained with one technique using another. Although the antibodies I raised and purified against MTA1 were not of suitable quality to be useful, I was fortunate that a commercial antibody became available which I used to show that MTA1 was a mitotic phosphoprotein. It has become

apparent that MTA1 is a component of the nucleosome-remodelling and deacetylase complex (Iguchi *et al.*, 2000; Toh *et al.*, 2000).

Recently it has been shown that expression of the MTA1 mRNA is upregulated by human-mesenchymal growth factor (heregulin, HRG) which in turn leads to elevation of MTA1 protein levels. The increase in MTA1 protein levels leads to suppression of histone acetylation and an increase in deacetylase activity, which in turn leads to suppression of gene expression (Mazumdar *et al.*, 2001). In particular it has been shown that MTA1 is an important co-repressor of oestrogen receptor element (ERE) transcription because it blocks the ability of oestradiol to stimulate oestrogen-receptor (ER) mediated transcription. Interestingly when MTA1 protein levels were elevated in tissue culture cells by treatment with heregulin (or by overexpression from an endogenous promoter), it was seen that MTA1 localisation changed from being diffusely nuclear and excluded from nucleoli, to being localised in large nuclear foci (Mazumdar *et al.*, 2001). These are identical localisation patterns to those I observed when expression of MTA1 was low and high respectively.

I have shown that mutation of the cyclin A3 hydrophobic patch residues I48 and E48 had little effect on the phosphorylation of Cdc6 which is a substrate believed to be recruited to CDK2:cyclin A/E by its RXL motif. Dr. Jon Moore has shown that homologous mutations in cyclin E abolished the ability of bound CDK2 to phosphorylate Rb, another protein believed to be recruited in an RXL-motif dependent manner. He has also shown recently that the hydrophobic patch mutations in cyclin E greatly impaired the phosphorylation of Cdc6 by CDK2. Phosphorylation of *in vitro* transcribed and translated Cdc6 was inhibited by a peptide spanning the RXL motif of p27 however. These data suggest that the contacts made between both I41 and E48 in the hydrophobic patch of cyclin A3 and both R and L of the RXL motif in substrates such as Cdc6 are unimportant for regulating phosphorylation by CDK2:cyclin A3. In contrast however the homologous residues in cyclin E are important for the recruitment of RXL-containing substrates. This highlights an important fundamental difference in the way that cyclin A and cyclin E recruit substrates. The data from the peptide inhibition experiments do indicate however that there are additional elements in or around the hydrophobic patch which are important for substrate selection by cyclin A.

Concluding comments

I have described a systematic *in vitro* small-pool expression cloning screen based on the work of Lustig *et al* (1997). I invested a considerable amount of time constructing cDNA libraries which could be used to express pools of proteins using *in vitro* coupled transcription and translation systems. Initially I had hoped to use cyclin A3 to screen the pools of translated proteins for proteins which interacted with cyclin A. I decided to use complexes of CDK2:cyclin A3 to perform the screen however, and discovered serendipitously that I could use the kinase activity of the complexes to directly identify substrates which underwent an electrophoretic mobility shift in the presence of the kinase.

I screened 144 pools of proteins by affinity chromatography and only identified a single protein which appeared to interact with CDK2:cyclin A. This interaction eventually turned out to be artefactual, however, and I decided to stop the affinity chromatography based screen in favour of the kinase based screen which was proving more successful in identifying substrates. I screened through a total of 488 pools of plasmid using the kinase based screen and isolated 15 cDNAs encoding proteins which were *in vitro* substrates of CDK2. The proteins identified were involved in processes as diverse as transcriptional regulation, translation control, cell signalling, cell cycle regulation and RNA metabolism. In an attempt to identify *in vitro* substrates which may have been relevant *in vivo* substrates I investigated the cellular localisation by transfecting or microinjecting plasmid encoding the protein of interest fused to YFP.

I began to use the panel of substrates I had identified to investigate the substrate specificity of different CDK:cyclin complexes in an attempt to answer the question: do different CDK:cyclin complexes have different substrate specificity?. My colleague Yoshimi Tanaka was performing a similar screen to the one described in this thesis using CDK1:cyclin B complexes and the *Xenopus laevis* library I had constructed. This would have provided an excellent opportunity to investigate whether the proteins I had identified could be phosphorylated by CDK1:cyclin B, and moreover, whether the proteins Yoshimi had identified could be phosphorylated by CDK2:cyclin A. We have also recently developed techniques for the expression of CDK2:cyclin E in our laboratory and this would provide another opportunity for exploring the question of substrate specificity. From the experiments I have performed using clones encoding proteins which were identified as

substrates for CDK1:cyclin B, and which were donated to me by Marc Kirschner and Yoshimi Tanaka, I have seen that there appears to be little specificity of particular CDK:cyclin complexes for particular substrates. It appeared that CDK2:cyclin A could phosphorylate the proteins identified as substrates of CDK1:cyclin B and vice versa.

These results suggest that the small-pool expression screen employed here is not especially well suited to the identification of substrates specific to particular CDK:cyclin complexes. This means that this screen has the potential to identify proteins which are not exclusively substrates of CDK2:cyclin A but may be substrates of CDK1:cyclin B or CDK2:cyclin E. This is supported by the fact that my screen identified nucleophosmin as a substrate of CDK2:cyclin A which has since been independently identified as a substrate of CDK2:cyclin E, and that it also identified Ect2 which is a good candidate to be a substrate of CDK1:cyclin B.

This type of *in vitro* screen requires that proteins be removed from their normal cellular localisation and be exposed to a kinase which in their proper physiological environment they may never encounter. As has been described above, it is likely that subcellular localisation plays an important role in determining substrate specificity. Removing proteins from their normal physiological environment evades this important specificity determinant, and it could be argued that this reduces the possibility of identifying substrates specific for a particular kinase. Moreover, the *in vitro* expression of proteins means that they are not expressed in their normal temporal manner. Again, this is likely to be an important mediator of specificity, and the inappropriate temporal expression of proteins reduces the possibility of identifying specific substrates.

This type of screen is also limited by the number of proteins that can be tested. It is also limited by the nature of the selection imposed. I used a selection for substrates that was based on a change in the electrophoretic mobility of a protein through SDS-PAGE in the presence of CDK2:cyclin A. Not all proteins undergo a change in mobility when they are phosphorylated and so would not be identified using this selection criteria. There will have inevitably been proteins present in the pools I screened which were substrates of CDK2:cyclin A but which do not undergo a change in mobility upon phosphorylation. It would be much more useful if a selection could be performed based on the incorporation of [^{32}P]. In this scenario it becomes difficult to imagine how the translated proteins could be

separated away from the other rabbit reticulocyte lysate components. I found that addition of $\gamma[^{32}P]$-ATP to a translated pool of proteins in the presence of CDK2:cyclin A3 generated too much background to allow the identification of the translated proteins. I found that the incorporation of biotinylated lysine into translated proteins meant that they could be sufficiently purified from other reticulocyte lysate components using monomeric streptavidin agarose beads. The captured proteins could then be used in kinase assays containing $\gamma[^{32}P]$-ATP. Substrates of added kinase could then be detected based on the incorporation of $[^{32}P]$ and not electrophoretic mobility.

I mutated I41 and E48 in the hydrophobic patch of cyclin A3 to alanines and tested whether this affected the kinase activity of CDK2:hpm cyclin A3 towards Cdc6 or histone H1. Surprisingly, these mutations had no detectable affect on the phosphorylation of Cdc6 relative to the wild-type enzyme. Analogous mutations in cyclin E, however, impaired the ability of bound CDK2 to phosphorylate Cdc6. These mutations do not affect the interaction between cyclin and CDK2, and so it appears that there is some difference in the way that the hydrophobic patch of cyclin A and cyclin E is used to recruit proteins to CDK:cyclin complexes. This difference could confer different abilities on cyclins A and E to select substrates.

To summarise, the screen I have used here has identified known substrates of CDK2:cyclin A, and therefore it is likely that at least some of the other proteins identified are likely to be meaningful substrates of CDK:cyclin complexes. The utility of a secondary screen became apparent once I had identified a number of proteins which appeared to be *in vitro* substrates. One possible secondary screen would be to use the generation of MPM2 reactivity to assess whether phosphorylated CDK consensus sites were being generated when the candidate proteins were phosphorylated. The screen, like all screens, is in some ways flawed, but also avoids many of the problems associated with other *in vitro* and *in vivo* screening methods which have been employed to date. The screening method is amenable to improvement which would increase both the throughput and sensitivity.

References.

Abremski, K. and Hoess, R. (1984) Bacteriophage P1 site-specific recombination. Purification and properties of the Cre recombinase protein. *J Biol Chem*, **259**, 1509-1514.

Adamczewski, J.P., Rossignol, M., Tassan, J.P., Nigg, E.A., Moncollin, V. and Egly, J.M. (1996) MAT1, cdk7 and cyclin H form a kinase complex which is UV light-sensitive upon association with TFIIH. *EMBO J*, **15**, 1877-1884.

Adams, P.D., Sellers, W.R., Sharma, S.K., Wu, A.D., Nalin, C.M. and Kaelin, W.G., Jr. (1996) Identification of a cyclin-cdk2 recognition motif present in substrates and p21-like cyclin-dependent kinase inhibitors. *Mol Cell Biol*, **16**, 6623-6633.

Alber, T. (1992) Structure of the leucine zipper. *Curr Opin Genet Dev*, **2**, 205-210.

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

Anderson, C.W., Baum, P.R. and Gesteland, R.F. (1973) Processing of adenovirus 2-induced proteins. *J Virol*, **12**, 241-522.

Arroyo, M., Bagchi, S. and Raychaudhuri, P. (1993) Association of the human papillomavirus type 16 E7 protein with the S-phase-specific E2F-cyclin A complex. *Mol Cell Biol*, **13**, 6537-6546.

Baldino, F. Jr., Chesselet, MF. and Lewis, ME. (1989) High-resolution in situ hybridization histochemistry. *Methods Enzymol*, **168**, 761-777

Basco, R.D., Segal, M.D. and Reed, S.I. (1995) Negative regulation of G1 and G2 by S-phase cyclins of Saccharomyces cerevisiae. *Mol Cell Biol*, **15**, 5030-5042.

Bertolaet, B.L., Clarke, D.J., Wolff, M., Watson, M.H., Henze, M., Divita, G. and Reed, S.I. (2001) UBA domains of DNA damage-inducible proteins interact with ubiquitin. *Nat Struct Biol*, **8**, 417-422.

Borer, R.A., Lehner, C.F., Eppenberger, H.M. and Nigg, E.A. (1989) Major nucleolar proteins shuttle between nucleus and cytoplasm. *Cell*, **56**, 379-390.

Brown, N.R., Noble, M.E., Endicott, J.A., Garman, E.F., Wakatsuki, S., Mitchell, E., Rasmussen, B., Hunt, T. and Johnson, L.N. (1995) The crystal structure of cyclin A. *Structure*, **3**, 1235-1247.

Brown, N.R., Noble, M.E., Endicott, J.A. and Johnson, L.N. (1999) The structural basis for specificity of substrate and recruitment peptides for cyclin-dependent kinases. *Nat Cell Biol*, **1**, 438-443.

Callebaut, I., Courvalin, J.C. and Mornon, J.P. (1999) The BAH (bromo-adjacent homology) domain: a link between DNA methylation, replication and transcriptional regulation. *FEBS Lett*, **446**, 189-193.

Cappelli, E., Carrozzino, F., Abbondandolo, A. and Frosina, G. (1999) The DNA helicases acting in nucleotide excision repair, XPD, CSB and XPB, are not required for PCNA-dependent repair of abasic sites. *Eur J Biochem*, **259**, 325-330.

Chapman-Smith, A. and Cronan, J.E., Jr. (1999) *In vivo* enzymatic protein biotinylation. *Biomol Eng*, **16**, 119-125.

Chevray, P.M. and Nathans, D. (1992) Protein interaction cloning in yeast: identification of mammalian proteins that react with the leucine zipper of Jun. *Proc Natl Acad Sci U S A*, **89**, 5789-5793.

Clackson, T. and Wells, J.A. (1995) A hot spot of binding energy in a hormone-receptor interface. *Science*, **267**, 383-386.

Clezardin, P., Frappart, L., Clerget, M., Pechoux, C. and Delmas, P.D. (1993) Expression of thrombospondin (TSP1) and its receptors (CD36 and CD51) in normal, hyperplastic, and neoplastic human breast. *Cancer Res*, **53**, 1421-1430.

Cote, J.F., Charest, A., Wagner, J. and Tremblay, M.L. (1998) Combination of gene targeting and substrate trapping to identify substrates of protein tyrosine phosphatases using PTP-PEST as a model. *Biochemistry*, **37**, 13128-13137.

Cowell, I.G. (1997) Yeast two-hybrid library screening. *Methods Mol Biol*, **69**, 185-202.

Cross, F.R. and Blake, C.M. (1993) The yeast Cln3 protein is an unstable activator of Cdc28. *Mol Cell Biol*, **13**, 3266-3271.

Cross, F.R. and Jacobson, M.D. (2000) Conservation and function of a potential substrate-binding domain in the yeast Clb5 B-type cyclin. *Mol Cell Biol*, **20**, 4782-4790.

Cross, F.R., Yuste-Rojas, M., Gray, S. and Jacobson, M.D. (1999) Specialization and targeting of B-type cyclins. *Mol Cell*, **4**, 11-19.

D'Souza, S.E., Ginsberg, M.H. and Plow, E.F. (1991) Arginyl-glycyl-aspartic acid (RGD): a cell adhesion motif. *Trends Biochem Sci*, **16**, 246-250.

Dahmann, C., Diffley, J.F. and Nasmyth, K.A. (1995) S-phase-promoting cyclin-dependent kinases prevent re-replication by inhibiting the transition of replication origins to a pre-replicative state. *Curr Biol*, **5**, 1257-1269.

Darbon, J.M., Devault, A., Taviaux, S., Fesquet, D., Martinez, A.M., Galas, S., Cavadore, J.C., Doree, M. and Blanchard, J.M. (1994) Cloning, expression and subcellular localization of the human homolog of p40MO15 catalytic subunit of cdk-activating kinase. *Oncogene*, **9**, 3127-3138.

Desai, D., Gu, Y. and Morgan, D.O. (1992) Activation of human cyclin-dependent kinases *in vitro*. *Mol Biol Cell*, **3**, 571-82 Issn: 1059-1524.

Dimri, G.P., Nakanishi, M., Desprez, P.Y., Smith, J.R. and Campisi, J. (1996) Inhibition of E2F activity by the cyclin-dependent protein kinase inhibitor p21 in cells expressing or lacking a functional retinoblastoma protein. *Mol Cell Biol*, **16**, 2987-2997.

Dirick, L. and Nasmyth, K. (1991) Positive feedback in the activation of G1 cyclins in yeast. *Nature*, **351**, 754-757.

Draviam, V.M., Orrechia, S., Lowe, M., Pardi, R. and Pines, J. (2001) The localization of human cyclins B1 and B2 determines CDK1 substrate specificity and neither enzyme requires MEK to disassemble the Golgi apparatus. *J Cell Biol*, **152**, 945-958.

Drury, L.S., Perkins, G. and Diffley, J.F. (1997) The Cdc4/34/53 pathway targets Cdc6p for proteolysis in budding yeast. *EMBO J*, **16**, 5966-976.

Dulic, A., Bates, P.A., Zhang, X., Martin, S.R., Freemont, P.S., Lindahl, T. and Barnes, D.E. (2001) BRCT domain interactions in the heterodimeric DNA repair protein XRCC1- DNA ligase III. *Biochemistry*, **40**, 5906-5913.

Dumbar, T.S., Gentry, G.A. and Olson, M.O. (1989) Interaction of nucleolar phosphoprotein B23 with nucleic acids. *Biochemistry*, **28**, 9495-94501.

Dynlacht, B.D., Flores, O., Lees, J.A. and Harlow, E. (1994a) Differential regulation of E2F *trans*-activation by cyclin/cdk2 complexes. *Genes Dev.*, **8**, 1772-1786.

Elledge, S.J., Mulligan, J.T., Ramer, S.W., Spottswood, M. and Davis, R.W. (1991) Lambda YES: A multifunctional cDNA expression vector for the isolation of genes by complementation of yeast and E.coli mutations. *Proceedings of the National Academy of Sciences of the USA*, **88**, 1731-1735.

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

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

Ewen, M.E., Faha, B., Harlow, E. and Livingston, D.M. (1992) Interaction of p107 with cyclin A independent of complex formation with viral oncoproteins. *Science*, **255**, 85-87.

Feuerstein, N., Chan, P.K. and Mond, J.J. (1988) Identification of numatrin, the nuclear matrix protein associated with induction of mitogenesis, as the nucleolar protein

B23. Implication for the role of the nucleolus in early transduction of mitogenic signals. *J Biol Chem*, **263**, 10608-10612.

Fisher, D.L. and Nurse, P. (1996) A single fission yeast mitotic cyclin B p34cdc2 kinase promotes both S- phase and mitosis in the absence of G1 cyclins. *EMBO J*, **15**, 850-860.

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

Flint, A.J., Tiganis, T., Barford, D. and Tonks, N.K. (1997) Development of "substrate-trapping" mutants to identify physiological substrates of protein tyrosine phosphatases. *Proc Natl Acad Sci U S A*, **94**, 1680-1685.

Fukunaga, R. and Hunter, T. (1997a) MNK1, a new MAP kinase-activated protein kinase, isolated by a novel expression screening method for identifying protein kinase substrates. *EMBO Journal*, **16**, 1921-1933.

Gautier, J., Norbury, C., Lohka, M., Nurse, P. and Maller, J. (1988) Purified maturation-promoting factor contains the product of a Xenopus homolog of the fission yeast cell cycle control gene cdc2^{+}. *Cell*, **54**, 433-439.

Geley, S., Kramer, E., Gieffers, C., Gannon, J., Peters, J.M. and Hunt, T. (2001) Anaphase-promoting complex/cyclosome-dependent proteolysis of human cyclin A starts at the beginning of mitosis and is not subject to the spindle assembly checkpoint. *J Cell Biol*, **153**, 137-148.

Geng, Y., Whoriskey, W., Park, M.Y., Bronson, R.T., Medema, R.H., Li, T., Weinberg, R.A. and Sicinski, P. (1999) Rescue of cyclin D1 deficiency by knockin cyclin E. *Cell*, **97**, 767-777.

Goodwin, G.H. and Nicolas, R.H. (2001) The BAH domain, polybromo and the RSC chromatin remodelling complex. *Gene*, **268**, 1-7.

Graziano, V. and Ramakrishnan, V. (1990) Interaction of HMG14 with chromatin. *J Mol Biol*, **214**, 897-910.

Gunther, E.J., Murray, N.E. and Glazer, P.M. (1993) High efficiency, restriction-deficient *in vitro* packaging extracts for bacteriophage lambda DNA using a new E.coli lysogen. *Nucleic Acids Res*, **21**, 3903-3904.

Gurdon, J.B. (1968a) Changes in somatic cell nuclei inserted into growing and maturing amphibian oocytes. *J Embryol Exp Morphol*, **20**, 401-414.

Gurdon, J.B. (1968b) Transplanted nuclei and cell differentiation. *Sci Am*, **219**, 24-35.

Gurley, L.R., Valdez, J.G. and Buchanan, J.S. (1995) Characterization of the mitotic specific phosphorylation site of histone H1. Absence of a consensus sequence for the p34cdc2/cyclin B kinase. *J Biol Chem*, **270**, 27653-27660.

Haaf, T., Golub, E.I., Reddy, G., Radding, C.M. and Ward, D.C. (1995) Nuclear foci of mammalian Rad51 recombination protein in somatic cells after DNA damage and its localization in synaptonemal complexes. *Proc Natl Acad Sci U S A*, **92**, 2298-2302.

Haase, S.B. and Reed, S.I. (1999) Evidence that a free-running oscillator drives G1 events in the budding yeast cell cycle. *Nature*, **401**, 394-397.

Habelhah, H., Shah, K., Huang, L., Burlingame, A.L., Shokat, K.M. and Ronai, Z. (2001) Identification of new JNK substrate using ATP pocket mutant JNK and a corresponding ATP analogue. *J Biol Chem*, **276**, 18090-18095.

Hadwiger, J.A., Wittenberg, C., Richardson, H.E., de Barros Lopes, M. and Reed, S.I. (1989) A family of cyclin homologs that control the G_1 phase in yeast. *Proc. Natl. Acad. Sci. USA*, **86**, 6255-6259.

Hall, C., Nelson, D.M., Ye, X., Baker, K., DeCaprio, J.A., Seeholzer, S., Lipinski, M. and Adams, P.D. (2001) HIRA, the human homologue of yeast Hir1p and Hir2p, is a novel cyclin- cdk2 substrate whose expression blocks S-phase progression. *Mol Cell Biol*, **21**, 1854-1865.

Hood, J.K., Hwang, W.W. and Silver, P.A. (2001) The Saccharomyces cerevisiae cyclin Clb2p is targeted to multiple subcellular locations by cis- and trans-acting determinants. *J Cell Sci*, **114**, 589-597.

Iguchi, H., Imura, G., Toh, Y. and Ogata, Y. (2000) Expression of MTA1, a metastasis-associated gene with histone deacetylase activity in pancreatic cancer. *Int J Oncol*, **16**, 1211-1214.

Jackman, M., Firth, M. and Pines, J. (1995) Human cyclins B1 and B2 are localized to strikingly different structures: B1 to microtubules, B2 primarily to the Golgi apparatus. *EMBO J*, **14**, 1646-1654.

Jackson, P.K., Chevalier, S., Philippe, M. and Kirschner, M.W. (1995) Early events in DNA replication require cyclin E and are blocked by p21CIP1. *J Cell Biol*, **130**, 755-769.

Jeffrey, P.D., Russo, A.A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J. and Pavletich, N.P. (1995a) Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. *Nature*, **376**, 313-320.

Jeoung, D.I., Oehlen, L.J. and Cross, F.R. (1998) Cln3-associated kinase activity in Saccharomyces cerevisiae is regulated by the mating factor pathway. *Mol Cell Biol*, **18**, 433-441.

Jiang, P.S., Chang, J.H. and Yung, B.Y. (2000) Different kinases phosphorylate nucleophosmin/B23 at different sites during G(2) and M phases of the cell cycle. *Cancer Lett*, **153**, 151-160.

Johnson, R.T. and Rao, P.N. (1970) Mammalian cell fusion: induction of premature chromosome condensation in interphase nuclei. *Nature*, **226**, 717-722.

Johnson, R.T. and Rao, P.N. (1971) Nucleo-cytoplasmic interactions in the achievement of nuclear synchrony in DNA synthesis and mitosis in multinucleate cells. *Biological Reviews*, **46**, 97-155.

Jones, C.E., Busch, H. and Olson, M.O. (1981) Sequence of a phosphorylation site in nucleolar protein B23. *Biochim Biophys Acta*, **667**, 209-212.

Kellogg, D.R., Kikuchi, A., Fujii-Nakata, T., Turck, C.W. and Murray, A.W. (1995a) Members of the NAP/SET family of proteins interact specifically with B- type cyclins. *J Cell Biol*, **130**, 661-673.

Kim, T.Y. and Kaelin, W.G., Jr. (2001) Differential control of transcription by dna-bound cyclins. *Mol Biol Cell*, **12**, 2207-22017.

Kimura, K., Hirano, M., Kobayashi, R. and Hirano, T. (1998) Phosphorylation and activation of 13S condensin by Cdc2 *in vitro*. *Science*, **282**, 487-490.

Kimura, K., Tsuji, T., Takada, Y., Miki, T. and Narumiya, S. (2000) Accumulation of GTP-bound RhoA during cytokinesis and a critical role of ECT2 in this accumulation. *J Biol Chem*, **275**, 17233-17236.

King, R.W., Lustig, K.D., Stukenberg, P.T., McGarry, T.J. and Kirschner, M.W. (1997) Expression cloning in the test tube. *Science*, **277**, 973-974.

Kobayashi, H., Stewart, E., Poon, R., Adamczewski, J.P., Gannon, J. and Hunt, T. (1992) Identification of the domains in cyclin A required for binding to, and activation of, p34cdc2 and p32cdk2 protein kinase subunits. *Mol Biol Cell*, **3**, 1279-1294.

Koch, C., Schleiffer, A., Ammerer, G. and Nasmyth, K. (1996) Switching transcription on and off during the yeast cell cycle: Cln/Cdc28 kinases activate bound transcription factor SBF (Swi4/Swi6) at start, whereas Clb/Cdc28 kinases displace it from the promoter in G2. *Genes Dev*, **10**, 129-141.

Krek, W., Ewen, M.E., Shirodkar, S., Arany, Z., Kaelin Jr., W.G. and Livingstone, D.M. (1994) Negative regulation of the growth-promoting transcription factor E2F-1 by a stably bound cyclin A-dependent protein kinase. *Cell*, **78**, 161-172.

Labib, K. and Diffley, J.F. (2001) Is the MCM2-7 complex the eukaryotic DNA replication fork helicase? *Curr Opin Genet Dev*, **11**, 64-70.

Liu, Y. and Maizels, N. (2000) Coordinated response of mammalian Rad51 and Rad52 to DNA damage. *EMBO Rep*, **1**, 85-90.

Liu, Y., Shah, K., Yang, F., Witucki, L. and Shokat, K.M. (1998b) Engineering Src family protein kinases with unnatural nucleotide specificity. *Chem Biol*, **5**, 91-101.

Lowe, M., Rabouille, C., Nakamura, N., Watson, R., Jackman, M., Jamsa, E., Rahman, D., Pappin, D.J. and Warren, G. (1998) Cdc2 kinase directly phosphorylates the cis-Golgi matrix protein GM130 and is required for Golgi fragmentation in mitosis. *Cell*, **94**, 783-793.

Lustig, K.D., Stukenberg, P.T., McGarry, T.J., King, R.W., Cryns, V.L., Mead, P.E., Zon, L.I., Yuan, J. and Kirschner, M.W. (1997) Small pool expression screening: identification of genes involved in cell cycle control, apoptosis, and early development. *Methods Enzymol*, **283**, 83-99.

Ma, T., Van Tine, B.A., Wei, Y., Garrett, M.D., Nelson, D., Adams, P.D., Wang, J., Qin, J., Chow, L.T. and Harper, J.W. (2000) Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. *Genes Dev*, **14**, 2298-3313.

Makiniemi, M., Hillukkala, T., Tuusa, J., Reini, K., Vaara, M., Huang, D., Pospiech, H., Majuri, I., Westerling, T., Makela, T.P. and Syvaoja, J.E. (2001) BRCT domain-containing protein TopBP1 functions in DNA replication and damage response. *J Biol Chem*, **276**, 30399-30406.

Marenholz, I., Zirra, M., Fischer, D.F., Backendorf, C., Ziegler, A. and Mischke, D. (2001) Identification of human epidermal differentiation complex (EDC)-encoded genes by subtractive hybridization of entire YACs to a gridded keratinocyte cDNA library. *Genome Res*, **11**, 341-355.

Masai, H., Matsui, E., You, Z., Ishimi, Y., Tamai, K. and Arai, K. (2000) Human Cdc7-related kinase complex. *In vitro* phosphorylation of MCM by concerted actions of Cdks and Cdc7 and that of a criticial threonine residue of Cdc7 bY Cdks. *J Biol Chem*, **275**, 29042-29052.

Mazumdar, A., Wang, R.A., Mishra, S.K., Adam, L., Bagheri-Yarmand, R., Mandal, M., Vadlamudi, R.K. and Kumar, R. (2001) Transcriptional repression of oestrogen receptor by metastasis- associated protein 1 corepressor. *Nat Cell Biol*, **3**, 30-37.

Mendenhall, M.D., al-Jumaily, W. and Nugroho, T.T. (1995) The Cdc28 inhibitor p40SIC1. *Prog Cell Cycle Res*, **1**, 173-185.

Mendenhall, M.D. and Hodge, A.E. (1998) Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast Saccharomyces cerevisiae. *Microbiol Mol Biol Rev*, **62**, 1191-1243.

Miki, T., Smith, C.L., Long, J.E., Eva, A. and Fleming, T.P. (1993) Oncogene ect2 is related to regulators of small GTP-binding proteins. *Nature*, **362**, 462-465.

Miller, M.E. and Cross, F.R. (2000) Distinct subcellular localization patterns contribute to functional specificity of the Cln2 and Cln3 cyclins of Saccharomyces cerevisiae. *Mol Cell Biol*, **20**, 542-555.

Miller, M.E. and Cross, F.R. (2001) Mechanisms controlling subcellular localization of the g(1) cyclins cln2p and cln3p in budding yeast. *Mol Cell Biol*, **21**, 6292-6311.

Mittnacht, S. (1998) Control of pRB phosphorylation. *Curr Opin Genet Dev*, **8**, 21-27.

Mondesert, O., McGowan, C.H. and Russell, P. (1996) Cig2, a B-type cyclin, promotes the onset of S in Schizosaccharomyces pombe. *Mol Cell Biol*, **16**, 1527-1533.

Moore, D.J., Taylor, R.M., Clements, P. and Caldecott, K.W. (2000) Mutation of a BRCT domain selectively disrupts DNA single-strand break repair in noncycling Chinese hamster ovary cells. *Proc Natl Acad Sci U S A*, **97**, 13649-13654.

Moreno, S., Hayles, J. and Nurse, P. (1989) Regulation of $p34^{cdc2}$ protein kinase during mitosis. *Cell*, **58**, 361-372.

Mudryj, M., Devoto, S.H., Hiebert, S.W., Hunter, T., Pines, J. and Nevins, J.R. (1991) Cell cycle regulation of the E2F transcription factor involves an interaction with cyclin A. *Cell*, **65**, 1243-1253.

Nagase, T., Seki, N., Tanaka, A., Ishikawa, K. and Nomura, N. (1995) Prediction of the coding sequences of unidentified human genes. IV. The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by analysis of cDNA clones from human cell line KG-1. *DNA Res*, **2**, 167-174.

Nasmyth, K. (1995) Evolution of the cell cycle. *Philos Trans R Soc Lond B Biol Sci*, **349**, 271-281.

Nasmyth, K. (1996) At the heart of the budding yeast cell cycle. *Trends Genet*, **12**, 405-412.

Nguyen, G.P., Bomsel, M., Labrousse, J.P. and Gallien, C.L. (1986) Partial purification of the maturation-promoting factor MPF from unfertilized eggs of *Xenopus laevis*. *Eur J Biochem*, **161**, 771-777.

Nielsen, H., Engelbrecht, J., Brunak, S. and von Heijne, G. (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng*, **10**, 1-6.

Nissen, M.S., Langan, T.A. and Reeves, R. (1991) Phosphorylation by cdc2 kinase modulates DNA binding activity of high mobility group I nonhistone chromatin protein. *J Biol Chem*, **266**, 19945-19952.

Ochs, R., Lischwe, M., O'Leary, P. and Busch, H. (1983) Localization of nucleolar phosphoproteins B23 and C23 during mitosis. *Exp Cell Res*, **146**, 139-149.

Okuda, M., Horn, H.F., Tarapore, P., Tokuyama, Y., Smulian, A.G., Chan, P.K., Knudsen, E.S., Hofmann, I.A., Snyder, J.D., Bove, K.E. and Fukasawa, K. (2000) Nucleophosmin/B23 is a target of CDK2/cyclin E in centrosome duplication. *Cell*, **103**, 127-140.

Ookata, K., Hisanaga, S., Bulinski, J.C., Murofushi, H., Aizawa, H., Itoh, T.J., Hotani, H., Okumura, E., Tachibana, K. and Kishimoto, T. (1995) Cyclin B interaction with microtubule-associated protein 4 (MAP4) targets p34cdc2 kinase to microtubules and is a potential regulator of M-phase microtubule dynamics. *J Cell Biol*, **128**, 849-862.

Parks, T.D., Howard, E.D., Wolpert, T.J., Arp, D.J. and Dougherty, W.G. (1995) Expression and purification of a recombinant tobacco etch virus NIa proteinase: biochemical analyses of the full-length and a naturally occurring truncated proteinase form. *Virology*, **210**, 194-201.

Peeper, D.S., Parker, L.L., Ewen, M.E., Toebes, M., Hall, F.L., Xu, M., Zantema, A., van der Eb, A.J. and Piwnica Worms, H. (1993) A- and B-type cyclins differentially modulate substrate specificity of cyclin-cdk complexes. *EMBO J*, **12**, 1947-1954.

Perkins, N.D., Felzien, L.K., Betts, J.C., Leung, K., Beach, D.H. and Nabel, G.J. (1997) Regulation of NF-kappaB by cyclin-dependent kinases associated with the p300 coactivator. *Science*, **275**, 523-527.

Peter, M., Nakagawa, J., Doree, M., Labbe, J.C. and Nigg, E.A. (1990) *In vitro* disassembly of the nuclear lamina and M phase-specific phosphorylation of lamins by cdc2 kinase. *Cell*, **61**, 591-602.

Petersen, B.O., Wagener, C., Marinoni, F., Kramer, E.R., Melixetian, M., Denchi, E.L., Gieffers, C., Matteucci, C., Peters, J.M. and Helin, K. (2000) Cell cycle- and cell growth-regulated proteolysis of mammalian CDC6 is dependent on APC-CDH1. *Genes Dev*, **14**, 2330-2343.

Pines, J. and Hunter, T. (1991) Human cyclins A and B1 are differentially located in the cell and undergo cell cycle-dependent nuclear transport. *J Cell Biol*, **115**, 1-17.

Poon, R.Y., Yamashita, K., Adamczewski, J.P., Hunt, T. and Shuttleworth, J. (1993) The cdc2-related protein p40MO15 is the catalytic subunit of a protein kinase that can activate p33cdk2 and p34cdc2. *EMBO J*, **12**, 3123-3132.

Prymakowska-Bosak, M., Misteli, T., Herrera, J.E., Shirakawa, H., Birger, Y., Garfield, S. and Bustin, M. (2001) Mitotic phosphorylation prevents the binding of HMGN proteins to chromatin. *Mol Cell Biol*, **21**, 5169-5178.

Rao, P.N. and Johnson, R.T. (1970) Mammalian cell fusion: Studies on the regulation of DNA synthesis and mitosis. *Nature*, **225**, 159-164.

Shaw, G. (1993) Identification of novel pleckstrin homology (PH) domains provides a hypothesis for PH domain function. *Biochem Biophys Res Commun*, **195**, 1145-1151.

Shenoy, B.C., Magner, W.J., Kumar, G.K., Phillips, N.F., Haase, F.C. and Samols, D. (1993) The nonbiotinylated form of the 1.3 s subunit of transcarboxylase binds to avidin (monomeric)-agarose: purification and separation from the biotinylated 1.3 S subunit. *Protein Expr Purif*, **4**, 85-94.

Shiyanov, P., Bagchi, S., Adami, G., Kokontis, J., Hay, N., Arroyo, M., Morozov, A. and Raychaudhuri, P. (1996) p21 Disrupts the interaction between cdk2 and the E2F-p130 complex. *Mol Cell Biol*, **16**, 737-744.

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. *Curr Biol*, **4**, 973-982.

Stern, B. and Nurse, P. (1996) A quantitative model for the cdc2 control of S phase and mitosis in fission yeast. *Trends Genet*, **12**, 345-350.

Stolz, J., Ludwig, A. and Sauer, N. (1998) Bacteriophage lambda surface display of a bacterial biotin acceptor domain reveals the minimal peptide size required for biotinylation. *FEBS Lett*, **440**, 213-217.

Stuart, D. and Wittenberg, C. (1994) Cell cycle-dependent transcription of *CLN2* is conferred by multiple distinct *cis*-acting regulatory elements. *Mol. Cell. Biol.*, **14**, 4788-4801.

Stukenberg, P.T., Lustig, K.D., McGarry, T.J., King, R.W., Kuang, J. and Kirschner, M.W. (1997) Systematic identification of mitotic phosphoproteins. *Curr Biol*, **7**, 338-348.

Swanton, C., Mann, D.J., Fleckenstein, B., Neipel, F., Peters, G. and Jones, N. (1997) Herpes viral cyclin/Cdk6 complexes evade inhibition by CDK inhibitor proteins. *Nature*, **390**, 184-187.

Szebeni, A. and Olson, M.O. (1999) Nucleolar protein B23 has molecular chaperone activities. *Protein Sci*, **8**, 905-912.

Tashiro, S., Walter, J., Shinohara, A., Kamada, N. and Cremer, T. (2000) Rad51 accumulation at sites of DNA damage and in postreplicative chromatin. *J Cell Biol*, **150**, 283-291.

Tassan, J.P., Jaquenoud, M., Fry, A.M., Frutiger, S., Hughes, G.J. and Nigg, E.A. (1995) *In vitro* assembly of a functional human CDK7-cyclin H complex requires MAT1, a novel 36 kDa RING finger protein. *EMBO J*, **14**, 5608-5617.

Tassan, J.P., Schultz, S.J., Bartek, J. and Nigg, E.A. (1994) Cell cycle analysis of the activity, subcellular localization, and subunit composition of human CAK (CDK-activating kinase). *J Cell Biol*, **127**, 467-478.

Tatsumoto, T., Xie, X., Blumenthal, R., Okamoto, I. and Miki, T. (1999) Human ECT2 is an exchange factor for Rho GTPases, phosphorylated in G2/M phases, and involved in cytokinesis. *J Cell Biol*, **147**, 921-928.

Taylor, R.M., Wickstead, B., Cronin, S. and Caldecott, K.W. (1998) Role of a BRCT domain in the interaction of DNA ligase III-alpha with the DNA repair protein XRCC1. *Curr Biol*, **8**, 877-880.

Thain, A., Gaston, K., Jenkins, O. and Clarke, A.R. (1996) A method for the separation of GST fusion proteins from co-purifying GroEL. *Trends Genet*, **12**, 209-210.

Thornton, K.H., Krishnan, V.V., West, M.G., Popham, J., Ramirez, M., Thelen, M.P. and Cosman, M. (2001) Expression, purification, and biophysical characterization of the BRCT domain of human DNA ligase IIIalpha. *Protein Expr Purif*, **21**, 401-411.

Thuret, J.Y., Valay, J.G., Faye, G. and Mann, C. (1996) Civ1 (CAK *in vivo*), a novel Cdk-activating kinase. *Cell*, **86**, 565-576.

Toh, Y., Kuninaka, S., Endo, K., Oshiro, T., Ikeda, Y., Nakashima, H., Baba, H., Kohnoe, S., Okamura, T., Nicolson, G.L. and Sugimachi, K. (2000) Molecular analysis of a

candidate metastasis-associated gene, MTA1: possible interaction with histone deacetylase 1. *J Exp Clin Cancer Res*, **19**, 105-111.

Toh, Y., Pencil, S.D. and Nicolson, G.L. (1994) A novel candidate metastasis-associated gene, mta1, differentially expressed in highly metastatic mammary adenocarcinoma cell lines. cDNA cloning, expression, and protein analyses. *J Biol Chem*, **269**, 22958-22963.

Tokuyama, Y., Horn, H.F., Kawamura, K., Tarapore, P. and Fukasawa, K. (2001) Specific phosphorylation of nucleophosmin on Thr(199) by cyclin- dependent kinase 2-cyclin E and its role in centrosome duplication. *J Biol Chem*, **276**, 21529-21537.

Tyers, M., Tokiwa, G. and Futcher, B. (1993) Comparison of the Saccharomyces cerevisiae G1 cyclins: Cln3 may be an upstream activator of Cln1, Cln2 and other cyclins. *EMBO J*, **12**, 1955-1968.

Visintin, R., Hwang, E.S. and Amon, A. (1999) Cfi1 prevents premature exit from mitosis by anchoring Cdc14 phosphatase in the nucleolus. *Nature*, **398**, 818-823.

Wasserman, W.J. and Smith, L.D. (1978) The cyclic behavior of a cytoplasmic factor controlling nuclear membrane breakdown. *J Cell Biol*, **78**, R15-22.

Wilhelm, H., Andersen, S.S. and Karsenti, E. (1997) Purification of recombinant cyclin B1/cdc2 kinase from Xenopus egg extracts. *Methods Enzymol*, **283**, 12-28.

Wittenberg, C., Sugimoto, K. and Reed, S.I. (1990) G1-specific cyclins of *S. cerevisiae*: cell cycle periodicity, regulation by mating pheromone, and association with the p34CDC28 protein kinase. *Cell*, **62**, 225-237.

Wu, L., Yee, A., Liu, L., Carbonaro-Hall, D., Venkatesan, N., Tolo, V.T. and Hall, F.L. (1994) Molecular cloning of the human CAK1 gene encoding a cyclin-dependent kinase-activating kinase. *Oncogene*, **9**, 2089-2096.

Yung, B.Y., Busch, H. and Chan, P.K. (1985) Translocation of nucleolar phosphoprotein B23 (37 kDa/pI 5.1) induced by selective inhibitors of ribosome synthesis. *Biochim Biophys Acta*, **826**, 167-173.

Zatsepina, O.V., Rousselet, A., Chan, P.K., Olson, M.O., Jordan, E.G. and Bornens, M. (1999) The nucleolar phosphoprotein B23 redistributes in part to the spindle poles during mitosis. *J Cell Sci*, **112**, 455-466.

Zhao, J., Dynlacht, B., Imai, T., Hori, T. and Harlow, E. (1998) Expression of NPAT, a novel substrate of cyclin E-CDK2, promotes S- phase entry. *Genes Dev*, **12**, 456-461.

Zhao, J., Kennedy, B.K., Lawrence, B.D., Barbie, D.A., Matera, A.G., Fletcher, J.A. and Harlow, E. (2000) NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription. *Genes Dev*, **14**, 2283-2297.

Appendix 1

Complete codons for hEct2

```
        |  10     |  20     |  30     |  40     |  50     |  60     |  70     |  80     |  90     |  100
   1 tttttgaatc ggttgtggcg gccgcggcga ggaatggcgg tatttgtgag aggagtcggc gtttgaagag gtggaactcc tagggctttt ttgagagtga
 101 cggagtctac ctcttgttac ctagactgga gtgcagtggc acgatctcgg ctcactgcaa cctctgcctc ccgggttcaa gcgattctcc tgcctcagcc
 201 tcctgagtag ctgggattac aggtgcctgc caccaagccc agctaatttt tgtatttttа gtagagatgg ggtttcattg tgttggccag gctggtctcg
 301 aactcctgac ctcgtgatcc gcccgccttg gcctcccaaa gtgctaggat tacaagtgtg agccaccgcg tccggccttt caaatggtat ttttgatttt
 401 cctcttccag tccttaaagc agctgattta gaagaataca aatcATGGCT GAAAATAGTG TATTAACATC CACTACTGGG AGGACTAGCT TGGCAGACTC
 501 TTCCATTTTT GATTCTAAAG TTACTGAGAT TTCCAAGGAA AACTTACTTA TTGGATCTAC TTCATATGTA GAAGAGATGC CTCAGATTGA AACAAGAGTG
 601 ATATTGGTTC AAGAAGCTGG AAAACAAGAA GAACTTATAA AAGCCTTAAA GGACATTAAA GTGGGCTTTG TAAAGATGGA GTCAGTGGAA GAATTTGAAG
 701 GTTTGGATTC TCCGGAATTT GAAAATGTAT TTGTAGTCAC GGACTTTCAG GATTCTGTCT TTAATGACCT CTACAAGGCT GATTGTAGAG TTATTGGACC
 801 ACCAGTTGTA TTAAATTGTT CACAAAAAGG AGAGCCTTTG CCATTTTCAT GTCGCCCGTT GTATTGTACA AGTATGATGA ATCTAGTACT ATGCTTTACT
 901 GGATTTAGGA AAAAAGAAGA ACTAGTCAGG TTGGTGACAT TGGTCCATCA CATGGGTGGA GTTATTCGAA AAGACTTTAA TTCAAAAGTT ACACATTTGG
1001 TGGCAAATTG TACACAAGGA GAAAAATTCA GGGTTGCTGT GAGTCTAGGT ACTCCAATTA TGAAGCCAGA ATGGATTTAT AAAGCTTGGG AAAGGCGGAA
1101 TGAACAGGAT TTCTATGCAG CAGTTGATGA CTTTAGAAAT GAATTTAAAG TTCCTCCATT TCAAGATTGT ATTTTAAGTT TCCTGGGATT TTCAGATGAA
1201 GAGAAAACCA ATATGGAAGA AATGACTGAA ATGCAAGGAG GTAAATATTT ACCGCTTGGA GATGAAAGAT GCACTCACCT TGTAGTTGAA GAGAATATAG
1301 TAAAAGATCT TCCCTTTGAA CCTTCAAAGA AACTTTATGT TGTCAAGCAA GAGTGGTTCT GGGGAAGCAT TCAAATGGAT GCCCGAGCTG GAGAAACTAT
1401 GTATTTATAT GAAAAGGCAA ATACTCCTGA GCTCAAGAAA TCAGTGTCAA TGCTTTCTCT AAATACCCCT AACAGCAATC GCAAACGACG TCGTTTAAAA
1501 GAAACACTTG CTCAGCTTTC AAGAGAGACA GACGTGTCAC CATTTCCACC CCGTAAGCGC CCATCAGCTG AGCATTCCCT TTCCATAGGG TCACTCCTAG
1601 ATATCTCCAA CACACCAGAG TCTAGCATTA ACTATGGAGA CACCCCAAAG TCTTGTACTA AGTCTTCTAA AAGCTCCACT CCAGTTCCTT CAAAGCAGTC
1701 AGCAAGGTGG CAAGTTGCAA AAGAGCTTTA TCAAACTGAA AGTAATTATG TTAATATATT GGCAACAATT ATTCAGTTAT TTCAAGTACC ATTGGAAGAG
1801 GAAGGACAAC GTGGTGGACC TATCCTTGCA CCAGAGGAGA TTAAGACTAT TTTTGGTAGC ATCCCAGATA TCTTTGATGT ACACACTAAG ATAAAGGATG
1901 ATCTTGAAGA CCTTATAGTT AATTGGGATG AGAGCAAAAG CATTGGTGAC ATTTTTCTGA AATATTCAAA AGATTTGGTA AAAACCTACC CTCCCTTTGT
2001 AAACTTCTTT GAAATGAGCA AGGAAACAAT TATTAAATGT GAAAAACAGA AACCAAGATT TCATGCTTTT CTCAAGATAA ACCAAGCAAA ACCAGAATGT
2101 GGACGGCAGA GCCTTGTTGA ACTTCTTATC CGACCAGTAC AGAGGTTACC CAGTGTTGCA TTACTTTTAA ATGATCTTAA GAAGCATACA GCTGATGAAA
2201 ATCCAGACAA AAGCACTTTA GAAAAAGCTA TTGGATCACT GAAGGAAGTA ATGACGCATA TTAATGAGGA TAAGAGAAAA ACAGAAGCTC AAAAGCAAAT
2301 TTTTGATGTT GTTTATGAAG TAGATGGATG CCCAGCTAAT CTTTTATCTT CTCACCGAAG CTTAGTACAG CGGGTTGAAA CAATTTCTCT AGGTGAGCAC
2401 CCCTGTGACA GAGGAGAACA AGTAACTCTC TTCCTCTTCA ATGATTGCCT AGAGATAGCA AGAAAACGGC ACAAGGTTAT TGGCACTTTT AGGAGTCCTC
2501 ATGGCCAAAC CCGACCCCCA GCTTCTCTTA AGCATATTCA CCTAATGCCT CTTTCTCAGA TTAAGAAGGT ATTGGACATA AGAGAGACAG AAGATTGCCA
2601 TAATGCTTTT GCCTTGCTTG TGAGGCCACC AACAGAGCAG GCAAATGTGC TACTCAGTTT CCAGATGACA TCAGATGAAC TTCCAAAAGA AAACTGGCTA
2701 AAGATGCTGT GTCGACATGT AGCTAACACC ATTTGTAAAG CAGATGCTGA GAATCTTATT TATACTGCTG ATCCAGAATC CTTTGAAGTA AATACAAAAG
2801 ATATGGACAG TACATTGAGT AGAGCATCAA GAGCAATAAA AAAGACTTCA AAAAAGGTTA CAAGAGCATT CTCTTTCTCC AAAACTCCAA AAAGAGCTCT
2901 TCGAAGGGCT CTTATGACAT CCCACGGCTC AGTGGAGGGA AGAAGTCCTT CCAGCAATGA TAAGCATGTA ATGAGTCGTC TTTCTAGCAC ATCATCATTA
3001 GCAGGTATCC CTTCTCCCTC CCTTGTCAGC CTTCCTTCCT TCTTTGAAAG GAGAAGTCAT ACGTTAAGTA GATCTACAAC TCATTTGATA TGAagcgtta
3101 ccaaaatctt aaattataga aatgtataga cacctcatac tcaaataaga aactgactta aatggtactt gtaattagca cgttggtgaa agctggaagg
3201 aagataaata acactaaact atgctatttg atttttcttc ttgaaagagt aaggtttacc tgttacattt tcaagttaat tcatgtaaaa aatgatagtg
3301 attttgatgt aatttatctc ttgtttgaat ctgtcattca aaggccaata atttaagttg ctatcagctg atattagtag ctttgcaacc ctgatagagt
3401 aaataaattt tatgggtggg tgccaaatac tgctgtgaat ctatttgtat agtatccatg aatgaattta tggaaataga tatttgtgca gctcaattta
3501 tgcagagatt aaatgacatc ataatactgg atgaaaactt gcatagaatt ctgattaaat agtgggtctg tttcacatgt gcagtttgaa gtatttaaat
3601 aaccactcct ttcacagttt attttcttct caagcgtttt caagatctag catgtggatt ttaaaagatt tgccctcatt aacaagaata acatttaaag
3701 gagattgttt caaaatattt ttgcaaattg agataaggac agaaagattg agaaacattg tatattttgc aaaaacaaga tgtttgtagc tgtttcagag
3801 agagtacggt atatttatgg taattttatc cactagcaaa tcttgattta gtttgatagt cgtcgtcgga attttatttt gaaggataag accatgggaa
3901 aattgtggta aagactgttt gtacccttca tgaaataatt ctgaagttgc catcagtttt actaatcttc tgtgaaatgc atagatatgc gcatgttcaa
4001 ctttttattg tggtcttata attaaatgta aaattgaaaa ttcatttgct gtttcaaagt gtgatatctt tcacaatagc ctttttatag tcagtaattc
4101 agaataatca agttcatatg gataaatgca tttttatttc ctatttcttt agggagtgct acaaatgttt gtcacttaaa tttcaagttt ctgttttaat
4201 agttaactga ctatagattg ttttctatgc catgtatgtg ccacttctga gagtagtaaa tgactctttg ctacatttta aaagcaattg tattagtaag
4301 aactttgtaa ataaatacct aaaacccaag tgtaaaaaaa aaaaaaaaaa aaaaaaaaaa aa
        |  10     |  20     |  30     |  40     |  50     |  60     |  70     |  80     |  90     |  100
```

Conceptual translation of hEct2 open reading frame.

```
        |  10     |  20     |  30     |  40     |  50     |  60     |  70     |  80     |  90     |  100
   1 MAENSVLTST TGRTSLADSS IFDSKVTEIS KENLLIGSTS YVEEMPQIET RVILVQEAGK QEELIKALKD IKVGFVKMES VEEFEGLDSP EFENVFVVTD
 101 FQDSVFNDLY KADCRVIGPP VVLNCSQKGE PLPFSCRPLY CTSMMNLVLC FTGFRKKEEL VRLVTLVHHM GGVIRKDFNS KVTHLVANCT QGEKFRVAVS
 201 LGTPIMKPEW IYKAWERRNE QDFYAAVDDF RNEFKVPPFQ DCILSFLGFS DEEKTNMEEM TEMQGGKYLP LGDERCTHLV VEENIVKDLP FEPSKKLYVV
 301 KQEWFWGSIQ MDARAGETMY LYEKANTPEL KKSVSMLSLN TPNSNRKRRR LKETLAQLSR ETDVSPFPPR KRPSAEHSLS IGSLLDISNT PESSINYGDT
 401 PKSCTKSSKS STPVPSKQSA RWQVAKELYQ TESNYVNILA TIIQLFQVPL EEEGQRGGPI LAPEEIKTIF GSIPDIFDVH TKIKDDLEDL IVNWDESKSI
 501 GDIFLKYSKD LVKTYPPFVN FFEMSKETII KCEKQKPRFH AFLKINQAKP ECGRQSLVEL LIRPVQRLPS VALLLNDLKK HTADENPDKS TLEKAIGSLK
 601 EVMTHINEDK RKTEAQKQIF DVVYEVDGCP ANLLSSHRSL VQRVETISLG EHPCDRGEQV TLFLFNDCLE IARKRHKVIG TFRSPHGQTR PPASLKHIHL
 701 MPLSQIKKVL DIRETEDCHN AFALLVRPPT EQANVLLSFQ MTSDELPKEN WLKMLCRHVA NTICKADAEN LIYTADPESF EVNTKDMDST LSRASRAIKK
 801 TSKKVTRAFS FSKTPKRALR RALMTSHGSV EGRSPSSNDK HVMSRLSSTS SLAGIPSPSL VSLPSFFERR SHTLSRSTTH LI*
        |  10     |  20     |  30     |  40     |  50     |  60     |  70     |  80     |  90     |  100
```