Cyclin-dependent kinases and nuclear functions in *Arabidopsis thaliana*

Konstantinos G. Alexiou
John Innes Centre

A thesis submitted for the degree of Doctor of Philosophy to
The University of East Anglia

To the memory of my father

ACKNOWLEDGEMENTS

I want to thank my supervisors, Professor John Doonan and Dr. Mike Naldrett, for their guidance and encouragement throughout the course of my PhD. I would also like to thank my adviser, Professor Peter Shaw, for his recommendations and help. I want also to acknowledge the valuable help given to me, during the proteomics work of the project, by Dr. Gerhard Saalbach from the JIC proteomics facility and Dr. Alex Jones from The Sainsbury Laboratory.

Acknowledgements also go to past and present members of the Doonan group, especially to Dr Max Bush and Dr. Veronika Mikitova, as well as Grant Calder for his help in microscopy.

I would also like to thank BBSRC for paying my tution fees.

Finally, I am grateful to my wife Rachil for psychological support and to my son Andreas, who came into my and Rachil's life and redefined its meaning.

ABSTRACT

The eukaryotic nucleus contains most of the cell's genetic material and is the primary site of many fundamental cellular processes such as transcription, replication and pre-mRNA processing. A liquid chromatography/mass spectrometry-based (LC/MS) analysis on nuclear extracts from *Arabidopsis thaliana* cell cultures identified more than 600 putative nuclear proteins, providing the first comprehensive catalogue of a plant nuclear proteome. The nuclear proteome of proliferating and stationary phase cells were compared using the iTRAQ technology and identified three broad classes of proteins based on whether their levels increased, decreased or remained constant.

The majority of proteins remain at approximately constant levels suggesting post-translational regulation of many functions. Cyclin-dependent kinase C (CDKC), a nuclear localised protein kinase, was found to co-localise with splicing factors and a component of the exon-junction complex (EJC). Expression of a mutant form of CDKC induced re-arrangement of splicing factors into large speckles, particularly in response to anoxia. CDKC2 interacts with and phosphorylates CyclinT1:3 and the C-terminal domain of RNA polymerase in vitro. However, no evidence for direct interaction with other co-localising proteins could be obtained in vitro, raising the possibility the CDKC indirectly regulates splicing factor distribution.

Abbreviations

2H	Yeast two hybrid
ACN	acetonitrile
APS	ammonium persulphate
ATP	adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
bp	base pair
CAK	CDK-activating kinase
CaMV	cauliflower mosaic virus
CB	cajal body
CCD	charge coupled device
CDK	cyclin-dependent kinase
CTD	c terminal domain
DFC	dense fibrillar component
DSIF	DRB sensitivity-inducing factor
DRB	5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole
EJC	exon-junction complex
EIF	elongation initiation factor
FITC	fluorescein isothiocyanate
EM	electron microscopy
ESI	electrospray ionisation
FC	fibrillar component
G1	gap 1 phase
G2	gap 2 phase
GFP	green fluorescent protein
GST	glutashione-S-transferase
GO	gene ontology
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	high performance liquid chromatography
IGC	interchromatin granule cluster
ITPG	Isopropyl β-D-1-thiogalactopyranoside

iTRAQ	isobaric tag for relative and absolute quantification
Kb	kilobase
kDa	kilo Dalton
LB	Luria Bertani
MAPK	mitogen-activated protein kinase
min	minutes
ml	mili-litre
m/z	mass-to-charge ratio
M	molarity
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MudPIT	multi-dimensional protein identification technology
NIB	nuclei isolation buffer
NMD	nonsense-mediated decay
no	nucleolus
NoLS	nucleolar localization signal
NOR	nucleolar organizing centre
Nu	nucleus
rDNA	ribosomal DNA
RFP	red fluorescent protein
PIC	pre initiation complex
PolII	Polyerase II
pTEFb	positive transcription elongation factor b
RNA	ribonucleic acid
RNAi	RNA interference
rRNA	ribosomal RNA
RP	reverse phase
SCX	strong cation exchange
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
sec	second
siRNA	short interfering RNA

S-phase	DNA synthesis phase
snRNA	small nuclear RNA
snoRNA	small nucleolar RNA
SR	Serine/arginine
TF	transcription factor
µl	micro-litre
UTR	untranslated region

CHAPTER 1 – INTRODUCTION

CHAPTER 2 - MATERIAL AND METHODS

CHAPTER 3 – PROTEOMICS ANALYSIS OF THE *ARABIDOPSIS* NUCLEUS

CHAPTER 4 –
QUANTITATIVE PROTEOMICS ANALYSIS OF THE *ARABIDOPSIS* NUCLEUS

Chapter 1 - Introduction

The present chapter is divided into three major sections. In the first section, I introduce the nucleus in more detail, some theories regarding its origin and also describe sub-nuclear bodies that participate in important functions, particularly mRNA processing. In the second section, I introduce the theory and technology of mass spectrometry (MS)-based analysis of proteins and its application in the study of the nucleus. In the final section I introduce cyclin-dependent kinases (CDKs) as regulatory proteins that control nuclear structure during the cell cycle and discuss the current knowledge of CDK-dependent regulation of fundamental processes, such as mRNA transcription, processing and translation in mammals and plants.

1.1 The nucleus

The cellular organisation of eukaryotes appears to be much more complex than that of prokaryotes. One aspect of increased complexity is the compartmentalisation of the eukaryotic cell into multiple types of membrane-bound organelles that have distinct functions. One of the most prominent of these organelles is the nucleus, which harbours most of the genetic information in the form of chromatin (Harris, 1999). Many important cellular functions take place within the nucleus, including pre-mRNA processing and ribosome biogenesis (Moore and Proudfoot, 2009; Bentley, 2005; Henras *et al*., 2008; Fromont-Racine *et al*., 2003). The nucleus forms a continuum with the rest of the cell through nuclear pores that traverse the nuclear membrane and provide a route for two-way communication between nucleus and cytoplasm (Xu and Massague, 2004).

Another important nuclear function is gene expression. As soon as transcription is initiated, parallel multi-stage processing of the nascent pre-mRNA prepares the mature transcript for export to the cytoplasm. Pre-mRNA processing consists of 5'UTR–capping, pre-mRNA splicing and polyadenylation (Kornblihtt *et al*., 2004; Proudfoot *et al*., 2002). Mature mRNA is exported to the cytoplasm via association with a multi-protein exon-junction complex (EJC), which binds 21-24 nucleotides upstream of an exon-exon junction in mRNA (Le Hir *et al*., 2000). Components of the ribosome, the translation factory of the cell, are assembled in the nucleolus, a distinct subnuclear

compartment. Ribosomal subunits are exported to the cytoplasm where they assemble into functional ribosomes.

These nuclear processes are regulated at a number of levels, such as reversible phosphorylation, carried out by protein kinases (Vermeulen *et al*., 2009; Stamm, 2008; Calvert *et al*., 2008; de la Fuente van Bentem *et al*., 2006). Some protein kinases have a solely nuclear localisation profile. Thus, nuclear-localised protein kinases are good candidates for dissecting nuclear processes and their dynamics.

1.2 Theories of nuclear origin

There are at least three major theories regarding the origin and evolution of the nucleus.

One hypothesises that the nucleus evolved from an archaeal endosymbiont taken up by a bacterial host. In this scenario, the archaeum's membrane evolved to be the nuclear membrane whereas its cytoplasm became the nucleolus (Lopez-Garcia and Moreira, 2006).

A second theory postulates that eukaryote-like organisms were present before the appearance of bacteria and archaea (Fuerst, 2005). This theory is supported by the observation that members of the planctomycetes bacteria contain membrane-bound compartments; a prominent characteristic of a eukaryotic cell. The planctomycetes, *Gemmata obscuriglobus* and *Pirella marina,* contain two membrane-bound compartments; one large compartment that contains proteins, ribosomes and RNA and a smaller one, inside the large compartment, that holds DNA, RNA and nucleic acid-processing proteins (Fuerst, 2005). The membrane around the DNA is double, but not continuous and therefore resembles the structure of the nuclear envelope in eukaryotic cells. Thus, planctomycetes might be related to eukaryotic progenitor cells but alternatively could represent an example of parallel evolution.

A third theory, "viral eukaryogenesis", states that the nucleus is the result of viral invasion (Bell, 2001). The organisation of some viruses resembles that of the eukaryotic nucleus and based on such functional similarities, this theory postulates that a virus living in an archaeum provided the necessary material for nucleus establishment (Bell, 2001).

Regardless of which theory (if any) is correct, the separation of chromatin and cytoplasm creates complexity and provides opportunities for regulating gene expression. In the next part I will introduce the interplay between nuclear structure and function.

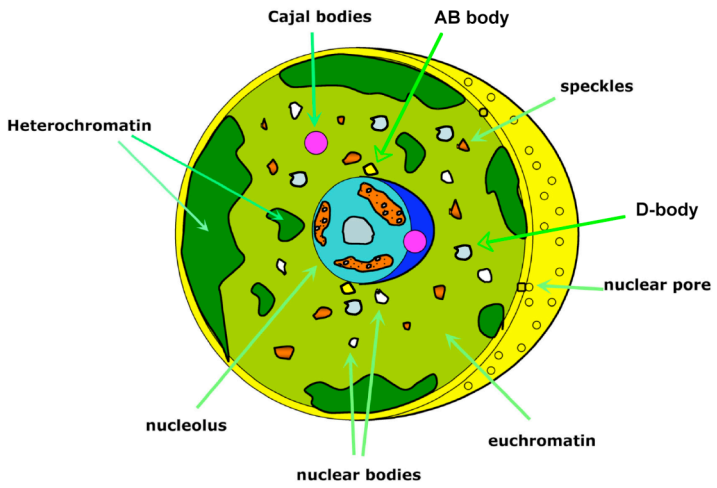


Figure 1.1 The plant nucleus. Chromatin regions, nucleolus and other nuclear bodies are shown (arrows) (Adapted by Shaw and Brown, 2004 and updated based on the most recent discoveries concerning nuclear bodies)

1.2.1 Chromatin architecture

The genetic material in the nucleus is organised into linear units known as chromosomes. Chromosomes define the highest order organisation of DNA and represent a complex packaging of the DNA and associated proteins. At other end of the scale, DNA threads are formed by DNA molecules wrapping around proteins called histones to produce nucleosomes. Gene-rich chromosome pools associate with interchromatin granule clusters (IGCs), non-chromatin regions consisting of proteins

involved in mRNA metabolism (Shopland *et al*., 2003; Saitoh *et al*., 2004), and transcription factories (TFs) where RNA PolII machineries reside (Puvion-Dutilleul *et al*., 1997; Osborne *et al*., 2004). These gene pools are highly dynamic and extensive intranuclear repositioning is observed. Upon transcriptional activation, gene pools from the nuclear periphery move into the nuclear interior towards the IGCs of TFs (Moen *et al*., 2004; Ragoczy *et al*., 2006) whereas the opposite movement is observed when transcriptional repression occurs (Brown *et al*., 1997). Thus, the nucleus is structured yet highly dynamic. Proteins complexes move within the nucleus interacting with their respective targets, while whole chromosome territories rearrange their position to facilitate interaction of gene foci with TFs. In the next section, some of the known functional domains of the nucleus will be considered.

1.2.2 Nuclear bodies in animal and plant cells

Nucleolus

The nucleolus is the largest single domain of the nucleus. It is not a sub-nuclear organelle *per se*, as it is not membrane-bound. Its major role is to assemble ribosomal units, which will form functional ribosomes (reviewed by Leung and Lamond, 2003; Cmarko *et al*., 2008). The nucleolus is organised around rDNA in tandem repeats known as nucleolar organizer regions (NORs). Precursor rRNAs (pre-rRNAs) are produced by RNA Polymerase I-driven transcription of rDNA, followed by cleavage and modification of pre-rRNAs. Mature rRNAs are assembled into mature ribosomal subunits through a multi-complex cascade of interactions involving small nucleolar ribonucleoproteins (snoRNPs) as well as non-ribosomal proteins (Tschochner and Hurt, 2003; Fatica and Tollervey, 2002). The 90S pre-ribosomal unit produced is subsequently cleaved into 40S and 60S pre-ribosomal subunits by the SSU processome (Schafer *et al*., 2003) in the nucleolus followed by their export to the nucleoplasm. While in the nucleoplasm, the majority of 40S- and 60S-associated factors dissociate from their respective subunits and are replaced by export factors that facilitate the exit of pre-ribosomal subunits to the nucleoplasm through the nuclear pore. Upon entry to the cytoplasm, export factors dissociate from their partners, leading to mature 40S and

60S ribosomal subunits (Trotta *et al*., 2003; Ferreira-Cerca *et al*., 2005; Hedges *et al*., 2005) that assemble to produce functional ribosomes.

When viewed by transmission electron microscopy (TEM), the nucleolus contains three distinctive components; the fibrillar centre (FC), the dense fibrillar component (DFC) and the granular component (GC). Even though both mammalian and plant nucleoli possess the aforementioned components, their arrangement is different (Figure 1.2). In mammalian cells, individual FCs are surrounded by DFCs whereas the rest of the nucleolus is filled with GC. In plant cells, multiple FCs are embedded into a large DFC and the rest of the nucleolus is filled with FC, whereas in the centre of the nucleolus, usually there is a central clear region called the "nucleolar cavity" or "vacuole" that is absent from the mammalian nucleolus (Shaw and Brown, 2004). Immunogold EM has shown that, in mammalian cells, FCs contain RNA polymerase I (Sheer and Rose, 1984) whereas transcription sites have been found to correspond to the DFC regions. Chromosome spreads show that active transcription sites appear as "Christmas trees", which correspond to nascent pre-rRNAs (Gonzalez-Melendi *et al*., 2001). Further processing of pre-rRNAs and the assembly of pre-ribosomal subunits takes place in the GCs.

Apart from its instrumental role in ribosome biogenesis, other functions of the nucleolus include nonsense-mediated decay (NMD) and involvement in virus trafficking (Pendle *et al*., 2005; Andersen *et al*., 2005; Hori and Watanabe, 2005; Emmott *et al*., 2008; Canetta *et al*., 2008). Involvement of the nucleolus in NMD has been inferred by the identification of exon-junction complex (EJC) proteins in plant and human nucleolar proteomes. The EJC recruits NMD-related factors both in animals and plants (Conti and Izaurralde, 2005; Arciga-Reyes *et al*., 2006). Simultaneous identification of mRNA populations in whole cells, in the nucleus and in the nucleolus of *Arabidopsis*, has shown that the percentage of aberrantly spliced transcripts is dramatically increased in the nucleolus, with 90% being putative NMD substrates (Brown and Shaw, 2008). These data suggest either that aberrant mRNAs are stored in the nucleolus to be degraded later via NMD in the nucleoplasm or that part of the NMD process itself takes place in the nucleolus.

Some nucleolar proteins, including viral proteins, exhibited characteristic nucleolar localisation signals (NoLS) (Emmott *et al*., 2008). Moreover, viral proteins in plants

have been shown to interact with the nucleolar protein fibrillarin and this interaction is important for the transport of viral RNAs through phloem (Canetta *et al*., 2008).

Thus, nucleolus emerges as an important subnuclear compartment that, apart from being the site of ribosome biogenesis, is involved in important processes that regulate cellular homeostasis.

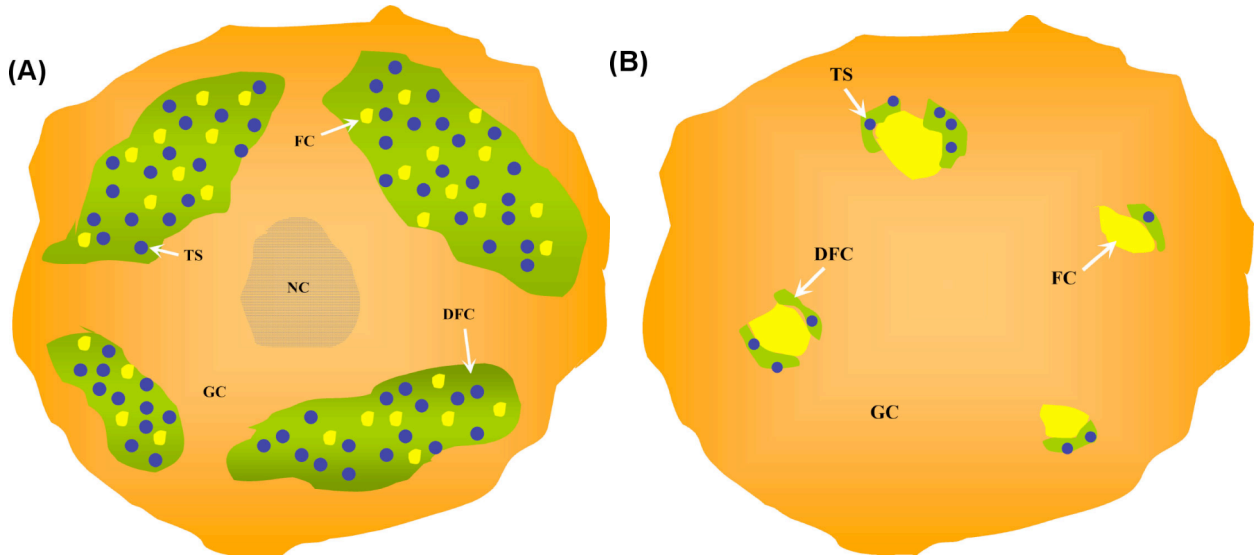


Figure 1.2. Structural organisation of the nucleolus.

(A) Organisation of the plant nucleolus. Transcription sites (TS) and fibrillar centres (FC) are surrounded by the dense fibrillar component (DFC). At the centre of the nucleolus there is the nucleolar cavity and the rest of the body is filled by the granular component (GC). **(B)** Organisation of the mammalian nucleolus. Like in plants, the GC occupies the majority of the nucleolar space. However, the sites of rDNA transcription and processing differ dramatically and an organisation into groups can be seen. As shown, TS are surrounded by DFC, which in turn are attached to FC. In this way multiple tri-partite groups are formed that are spread around the nucleolar space (Adapted by Shaw and Brown, 2004).

Nuclear speckles

The majority of protein-coding genes in eukaryotic cells are transcribed as immature pre-mRNAs that contain both exons and intervening introns. Before these pre-mRNAs are exported from the nucleus, they undergo processing that leads to the production of mature mRNAs ready to be translated. The first step of pre-mRNA processing is the removal of introns, which is carried out by a multi-protein complex called the spliceosome. The spliceosome consists of five snRNPs (U1, U2, U2/U6, U5) and a large number of non-snRNP splicing factors (Rappsilber *et al*., 2002; Zhou *et al*., 2002; Jurica and Moore, 2003). Both in plant and mammalian cells, splicing factors localise to nuclear spots, known as speckles, which have an irregular distribution around the nucleoplasm (Lorkovic *et al*., 2004; Lamond and Spector, 2003). Speckles are believed to be pools for the storage or assembly of splicing factors that can then be recruited to sites of high transcriptional activity and, therefore, may be important for the coordination of transcription and pre-mRNA splicing (Misteli *et al*., 1997; Misteli, 2000; Sacco-Bubulya and Spector, 2002). Speckle distribution is very dynamic and its pattern depends on the transcriptional or phosphorylation status of the cells as well as cell type and developmental stage (Misteli, 2000; Docquier *et al*., 2004; Fang *et* al., 2004; Tillemans *et al*., 2005, 2006).

Different serine-arginine (SR) splicing factor proteins localise into distinct populations of nuclear speckles with no, partial or complete co-localisation depending on the proteins under examination (Lorkovic *et al*., 2008). This diverse localisation profile of SR proteins might reflect recruitment of splicing factors to transcription and mRNA-processing sites of specific genes, depending on the cell type and developmental stage of a tissue (Lorkovic *et al*., 2008). Nevertheless, in some cases co-localisation or not of SR proteins could be misleading when considering functional relationships among proteins (Rino *et al*., 2007). Specific SR proteins that did not co-localize *in vivo* were found to interact in yeast two-hybrid and co-immunoprecipitation assays, suggesting that splicing factors move freely around the nucleoplasm and that "speckles" are just stochastic localised concentrations of these freely moving splicing factors (Rino *et al*., 2007). More studies will be needed to decipher the real nature of nuclear speckles.

Other nuclear bodies

Cajal bodies (CB) are structures of about 0.5–1.0 μm in size, not surrounded by a membrane and found within the nuclei of most plant and animal cells. They are dynamic structures that move, split, rejoin and exchange their contents with the surrounding nucleoplasm. The size and number of CBs depend on cell type, cell cycle stage, and metabolic activity (Andrade *et al*., 1993; Boudonck *et al*., 1998, 1999; Platani *et al*., 2000; Shaw and Brown, 2004). In plants, a distant homologue of the vertebrate coilin was necessary for the formation of CBs and its expression levels affect the size of CBs (Collier *et al*., 2006). It is believed that CBs function in the metabolism of different classes of snRNP particles (Nesic *et al*., 2004; Schaffert *et al*., 2004) and also act as the site of generation of siRNA/protein complexes mediating RNA-directed DNA methylation (Li *et al*., 2006).

Cyclophilin-containing bodies were identified when the localization pattern of proteins containing RS/SP (serine/proline) and cyclophilin domains was determined. The cyclophilins, CypRS64 and CypRS92, have the ability to interact with snRNPs and SR proteins, suggesting a role in spliceosome assembly (Lorkovic *et al*., 2004). Localisation profile of CypRS64 protein was quite distinct from that of CBs, whereas, when co-expressed with SR proteins, CypRS64 tranlocated to SR-containing speckles (Lorkovic *et al*., 2004). This suggested that cyclophilin-containing bodies are storage pools for cyclophilins, which translocate to speckles in order to affect SR protein structure, and probably facilitate or respond to SR protein phosphorylation (Lorkovic *et al*., 2004; Schiene *et al*., 2000).

In plants, processing of microRNAs (miRNAs) is performed in the nucleus, as opposed to animal cells where processing occurs in both nucleus and cytoplasm. miRNA processing in plants is mediated by the enzyme Dicer-like1 (DCL1), together with other proteins such as the zinc finger protein SERRATE (SE) and dsRNA-binding protein HYPONASTIC LEAF1 (HYL1) (Chen, 2005; Han *et al*., 2004). DHL1 and HYL1 co-localise in D-bodies, which are distinct from both CBs and speckles (Fang and Spector, 2007). Bi-molecular fluorescence complementation (BiFC) showed that DCL1, HYL1 and SE interact in the D-bodies and a pre-miRNA was targeted to D-bodies (Fang and Spector, 2007) suggesting that they could be the site of pre-miRNA processing and biogenesis. Also, since DCL1 is involved in processing of pathogen-induced small-

interfering RNAs (Navarro *et al*., 2006), D-bodies might play a role in plant defence (Fang and Spector, 2007).

Li *et al*. (2008) identified a new class of nuclear bodies in the *Arabidopsis* nucleus using ARGONAUTE4 (AGO4) protein, a regulator of siRNA-mediated gene silencing (Zilberman *et al*., 2003). AGO4 localised into two distinct nuclear body populations; one that co-localised with CBs and the other that co-localised with a subunit of the plant-specific RNA PolIV, NRPD1b (Li *et al.*, 2006; 2008). This second population was called "AB-bodies". NRPD2 and DNA methyltransferase DRM2 are additional proteins that localised in AB-bodies (Li *et al*., 2008). These nuclear bodies may mediate silencing with a possible target being the 45S rDNA loci, since AB-bodies are adjacent or overlap with the 45S loci, whereas CBs were never found to be close (Li *et al*., 2008).

1.3. The nucleus during proliferation and quiescence

Nuclear structure and function also changes over longer time periods. Thus, proliferating nuclei are profoundly different from differentiated or quiescent nuclei. In the next section I consider the definition of the state of quiescence in cell cultures, how it is regulated at the genetic level and describe what is known about quiescence in *Arabidopsis*.

1.3.1 Quiescence in cell cultures

When nutrients or growth factors become depleted, eukaryotic cells growing in culture will exit from the cell cycle and enter a resting phase called "quiescence" or "G0" (Su *et al*., 1996). These terms are also used to describe a cell culture grown to saturation. In yeast, "stationary phase" refers to the state of a saturated culture and the term "quiescence" refers to the state of the individual cells within the stationary phase culture (Gray *et al*., 2004). The stage of the culture is determined by measuring, for instance, its optical density or dry weight at different times, whereas "quiescence" is determined by examining cellular processes, like transcription and translational rate or metabolite concentrations (Choder, 1991). Re-supplying a stationary phase culture with essential

components such as nutrients or growth factors, can induce re-entry into the proliferative phase. So, by analogy to the cell division cycle, the cells in such cultures can be considered to cycle between proliferation and quiescence (Gray *et al*., 2004). The budding (*Saccharomyces cerevisiae*) and fission (*Schizosaccharomyces pombe*) yeast have been used as models to study the growth cycle and how cells enter and exit from a quiescent state. One major difference between the two species is that, contrary to *S. pombe* cells, *S. cerevisiae* cells go through a "post-diauxic shift" phase before entering the stationary phase, where cells utilize ethanol as an energy source (Herman, 2002). This basic difference in the culture growth profiles of the two yeast species suggests that they respond differently to nutrient limitations and, most probably, employ different genetic programs for entering into and exiting from quiescence (Yanagida, 2009). Analysis of cell division parameters in *Arabidopsis* suspension cultures (Hutchins, 2004) showed that the growth profile of the culture may resemble that of *S. pombe*, since cultures enter directly into stationary phase at days 5-6 (see Figure 1.3), where cell division is minimal and culture grows only due to cell biomass accumulation (Richard, *et al*., 2001).

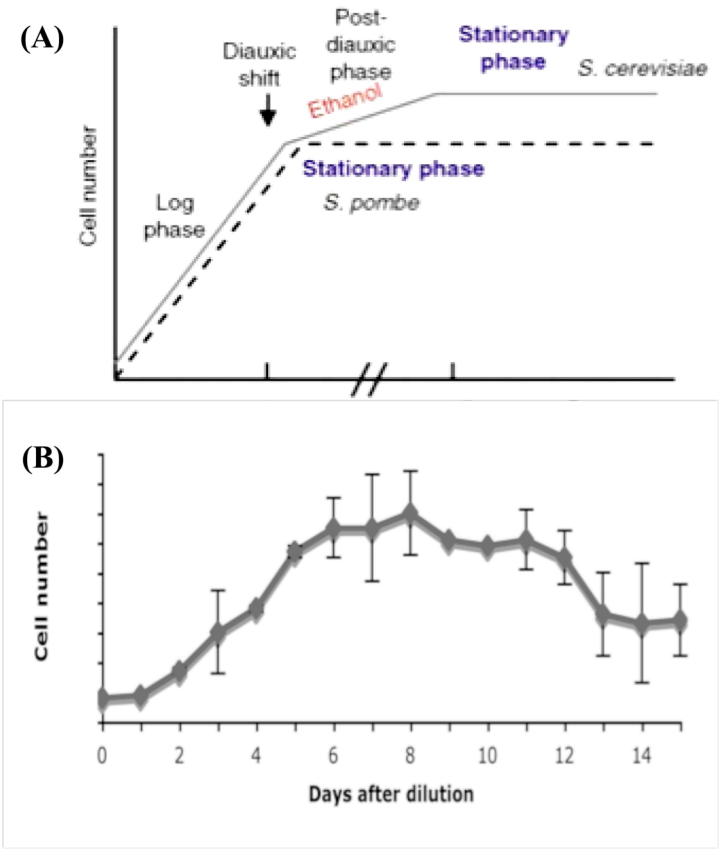


Figure 1.3 Cell numbers of yeast and *Arabidopsis* cell cultures as they progress from proliferation to stationary phase

(A) The cell number profile of *S. cerevisiae* (-) and *S. pombe* (--) are shown. Note that *S. cerevisiae* cells go through the "post-diauxic shift" phase that is not observed in *S. pombe* (Yanagida, 2009). **(B)** The cell number profile of *Arabidopsis* cell cultures pass from the log phase of growth to the stationary phase, resembling the growth profile of *S. pombe* cell cultures (Hutchins, 2004).

1.3.2 Genetics of cell quiescence

Both nitrogen- and glucose-starvation of *S.pombe* lead to entry into stationary phase, but the cellular responses are distinctly different. In the first case, yeast cells undergo two rounds of cell division and acquire a shorter and round shape compared to

proliferating cells (Su *et al*., 1996). In the second case, cells arrest their growth without further divisions, do not change their shape, and the number of dead cells gradually increases. The latter response resembles cells in a 6-day-old *Arabidopsis* cell culture as they cease growth (Hutchins, 2004). Carbon source depletion also reduces cell division and cell cycle regulatory genes (i,e the G1 cyclin *CYCD2*) are induced very rapidly by sucrose addition to a stationary cell culture (Riou-Khamlichi *et al*., 2000), leading to activation of its Cyclin-dependent kinase (CDK) partner CDKA1 which allow cell cycle progression. Thus, sucrose may be a key limiting factor for entry into and exit from quiescence.

Target-of-rapamycin (TOR) protein kinases are conserved core regulators of cell proliferation in yeast and human cells (Loewith, *et al*., 2002; Ma and Blenis, 2009). Upon sucrose or nitrogen starvation, TOR signalling is inactivated through co-ordinated steps of activation and de-activation of target genes (Gasch *et al*., 2000; Powers and Walter, 1999), allowing the cell culture to enter stationary phase. The *Arabidopsis* homologue, *AtTOR*, is expressed only in meristems, embryo and endosperm of *Arabidopsis* plants but not in differentiated tissues (Menand *et al*., 2002) and RNAi-mediated knockout of the gene led to developmental defects such as early senescence, growth arrest and reduction of mRNA translation (Deprost *et al*., 2007). These data suggest that TOR kinase in *Arabidopsis* plays a role in plant development via control of mRNA translation.

Three additional signalling pathways that control entry into quiescence have been described in budding yeast; the Protein kinase A, the Protein kinase C and the Snf1p pathway. The first pathway may be a negative regulator of entry into quiescence, whereas the other two are considered as positive regulators (Gray *et al*., 2004 and references therein). Again, it is not clear if similar pathways exist in plants.

A genome-wide transcriptomics approach revealed seven distinct gene clusters with altered mRNA expression levels as yeast cells exit quiescence in response to re-addition of a nitrogen source (Shimanuki, 2007). Down-regulated genes were implicated in catabolic and oxidation processes whereas up-regulated genes were involved in ribosome biogenesis and protein synthesis. Surprisingly, mRNA levels of genes involved in cell cycle, splicing and chromatin modification did not show significant changes. Considering the fact that both cell cycle and increased transcriptional activity

are resumed after nitrogen addition, genes involved in these processes are most probably regulated at the post-translational level.

In plants there are two studies that analyze transcriptome response upon sucrose removal (Contento, *et* al., 2004) or sucrose addition (Menges *et al*., 2003) in *Arabidopsis* cell cultures. In both studies, genes involved in protein synthesis, ribosome biogenesis and transcription were inversely regulated indicating that *Arabidopsis* cells probably modify the activity of common cellular processes as they enter or exit quiescence, albeit via the regulation of different genes. Despite these studies, no data have been published yet about the transcriptomic response of subcellular organelles, like the nucleus, in sucrose addition or removal.

1.3.3 Nuclear size and morphology during the growth cycle

Nuclear size and morphology varies between proliferating and quiescence cells, and some differentiated cell types contain nuclei with extreme nuclear phenotypes. In this section I will first describe some of the variation in nuclear morphology associated with differentiation and then discuss which features are associated with quiescence.

1.3.3.1 Variation in nuclear size and morphology associate with the switch between proliferation and quiescence/differentiation

The majority of eukaryotic cells contain round or elliptic nuclei. However, specialized cells often contain nuclei of altered shapes. For example, spermatid cells contain highly elongated nuclei, neutrophil cells have extremely lobed nuclei that is associated with loss of lamin A/C and expression of lamin B receptor. The precursor stem cells have large and round nuclei that lack lamin A protein. These structural alterations may cause chromatin repositioning, which could increase the accessibility of the transcriptional machineries to tissue-specific genes or, conversely, prevent unnecessary genes from being transcribed (Dahl *et al*., 2008). Diseases collectively called laminopathies are due to defects in genes encoding lamins A and C and are associated with deformed nuclei and changes in chromatin organization as well as sensitivity to mechanical stress (Hoffmann *et al*., 2002; Olins *et al*., 2008; Rusinol and Sinensky, 2006). Lamins also

directly regulate the transcription of certain genes as well as DNA binding and activity of transcription factors (Lammerding *et al*., 2004; Ivorra *et al*., 2006).

Nuclear shape and structural nuclear proteins may be directly related to other cellular phenotypes For example, collagen synthesis and epithelial cell tissue morphogenesis are dictated by nucleus shape rather that cell shape. Shape of the nucleus, and not cell morphology, was found to be strongly correlated with the breast cancer cell phenotype as well as other cancer types (Thomas *et al*., 2002; Lelievre, *et al*., 1998; Zink *et al*., 2004).

Plant nuclei lack lamins (Moreno Diaz de la Espina *et al*., 1991; Meier, 2007), and presumably the cell wall protects nuclei from shear stresses. However, their shape and size vary in ways reminiscent of animal nuclei. Meristematic plant cells tend to contain round nuclei, whereas differentiated tissues such as leaves and trichomes contain elongated nuclei. In *Arabidopsis*, nuclear size positively correlates with cell size, even though the causal link between the two phenotypes has not been fully explained (Breuer *et al*., 2009; Sugimoto-Shirasu *et al*., 2005: Sugimoto-Shirasu and Roberts, 2003). The "karyoplasmic ratio" theory states that cell size is coupled to the size of the nucleus and, consequently, to ploidy levels (Sugimoto-Shirasu and Roberts, 2003). DNA content is increased by the process of endoreduplication; successive rounds of DNA replication without intervening mitoses. However, it is not clear whether the size of the plant cell controls the final size of the nucleus or vice versa. Experimental evidence in mammalian cells supports the first case, as early experiments on transplantation of small nuclei into the cytoplasm of nuclei-depleted HeLa cells resulted in their enlargement (Harris, 1967). Since cytoplasmic factors in mammalian cells are responsible for retaining a constant ratio of cell to nuclear volume (Gregory, 2005; Kiseleva *et al*., 2007), it is implied that nuclei size can be altered without change in the DNA content. It is not clear if this is also the case in plants.

The ratio between cell volume and nuclear volume may be important for the correct timing of cell cycle entry (Futcher, 1996) and alterations in this ratio are associated with certain types of cancer (Zink *et al*., 2004). Also, there is strong correlation between the nucleus size, transcriptional activity and cell size, suggesting that larger cells need larger nuclei in order to accommodate the requirement for high rate of transcription (Schmidt and Schibler, 1995). So far, no alterations to the karyoplasmic ratio is reported

for plants, so no conclusions can be drawn on the importance of nucleus size to cellular functions.

1.3.3.2 Variation in nuclear compartments associated with transcriptional activity

Proliferating cells tend to be more active transcriptionally than quiescent cells (Menges *et al*., 2003; Contento, *et* al., 2004; Liu *et al*., 2007). Many aspects of nuclear structure are directly or indirectly related to transcription and other aspects of gene expression. For example, the non-random distribution of chromosomal DNA into chromosome territories (Cremer *et al*., 2006) and localised distribution of nuclear bodies and protein complexes involved in gene expression, suggest that gene positioning is functionally important for regulation of gene expression (reviewed in Fraser *et al*., 2007; Kumaran *et al*., 2008). In animal cells, nuclear lamina proteins and peri-nucleolar chromatin (PNC) are believed to form a transcriptionally repressive region in the nucleus that is maintained by recruitment of heterochromatin proteins, histone deacetylases and other transcriptional repressors or, in the case of PNC, via the action of non-coding RNAs (ncRNAs). On the other hand, nuclear pore complexes (NPCs), promyelocytic leukaemia (PML) nuclear bodies and nuclear speckles are believed to be sites of active gene transcription (Zhao *et al*., 2009 and references therein). Extensive chromosomal re-organisation, including recruitment of euchromatin-rich promoter regions to the NPCs, is caused by a global increase of acetylation, whereas PML bodies are believed to induce local changes in the chromatin architecture via interaction with other factors. Other nuclear bodies, such as nuclear speckles, have been consistently found to associate with transcriptionally active genes, even though there are reports suggesting that this association is stochastic and occurs only after gene activation (Brown *et al*., 2008). Conclusively, the nuclear neighbourhoods clustered around genes are structurally specialised to mediate either gene transcription or suppression (Zhao *et al*., 2009).

Changes of the transcriptional status in both plant and mammalian nuclei induce redistribution of nuclear speckles and the portion of RNA Polymerase II associated with speckles (Bregman *et al*., 1995). Upon drug-induced transcriptional inhibition, PolII and splicing factors aggregate into 5-8 “mega-speckles” in the nucleus (Phair and Misteli,

2000; Tillemans *et al*., 2006; Fang *et al*., 2004), imposing a dramatic change in the nuclear structure. Thus, under unfavourable conditions in the nucleus, transcriptional machineries are sequestered away from their transcription sites and stored in speckles, until the physiological conditions in the cell are in favour of transcription resumption. Distribution of splicing factors also depends on the developmental status of the cell. In meristematic cells splicing factors show a diffused localisation pattern whereas in differentiated cells they acquire a more speckled localisation. These different protein localisation profiles were attributed to different transcriptional status of the cells (Fang *et al*., 2004). In meristematic cells, transcription rate is high and thus larger amounts of splicing factors are recruited from speckles to transcriptionally active sites, compared to differentiated cells where the transcription rates are low. This increased recruitment results to the observed diffused localisation profile. Moreover, topological reasons also preclude the formation of large number of speckles in the meristematic cells; high transcription rates result to highly decondensed chromatin and, thus, smaller interchromatin spaces that hinder the formation of smaller bodies, like speckles (Fang *et al*., 2004: Docquier *et al*., 2004).

The organisation of chromosome domains also varies between proliferating and quiescent mammalian cells (Galiova *et al*., 2008; Kuroda et *al*., 2004; Bridger *et al*., 2000; Meaburn *et al*., 2007; Mehta *et al*., 2010). Such chromosomal movements can be very rapid and take place within 15 minutes after serum removal, whereas when quiescent cells re-enter the proliferation phase by repeated addition of serum, is considerably slower and takes up 24 to 36 hours (Mehta *et al*., 2010). These chromosome movements are linked to active gene transcription, since they are sensitive to inhibition of RNA polymerase II activity (Heard and Bickmore, 2007). In *Arabidopsis*, chromosome positioning does not differ significantly among nuclei of similar shape that originated from different organs or different cell types and that non-specific interactions were sufficient to explain the position of heterochromatin and nucleolus in nuclei (Berr and Schubert, 2007). This stochastic assembly of chromosomes and the nucleolus in the *Arabidopsis* nucleus suggests that regulation of gene transcription is not mediated by the movement of chromosomal DNA to transcriptionally-permissive regions, as in mammalian cells (Zhao *et al*., 2004).

1.4 Methods for studying nuclear proteome dynamics

1.4.1 Live cell imaging

The nucleus is a highly dynamic organelle whose structure alters in ways that reflect its activity. Live cell imaging has proven very useful for monitoring the protein movements and protein-protein interactions that mediate nuclear activity. Fluorescence-imaging methods have been developed for detecting protein-protein interactions both in time and space with high resolution. These methods include fluorescence recovery after photobleaching (FRAP) and quantum dots (QDs) for detecting protein movement (Brown *et al*., 1999; Michalet *et al*., 2005) and fluorescence resonance energy transfer (FRET) and fluorescence correlation spectroscopy (FCS) for detecting interaction between two proteins fused to different fluorophores (Jares-Erijman and Jovin, 2003; Haustein and Schwille, 2007). Modifications of these methods along with others led to the development of a repertoire of live cell imaging fluorescent techniques with high temporal and spatial resolution (reviewed in Lidke and Wilson, 2009).

1.4.2 Biochemical methods for monitoring protein dynamics

As the largest organelle, physical isolation of the nucleus has allowed its composition to be investigated at the biochemical level. Subcellular fractionation, coupled to western blotting, can provide information on how nuclear protein composition changes through time along a developmental sequence. Protein-protein interactions can be followed directly in such nuclear preparations using immunoprecipitation assays or indirectly using the yeast two-hybrid system and complementation assays. However, such methodologies have limitations, including low sensitivity, and difficulties in terms of scaling. Recently, the development of mass spectrometry-based protein identification methods has provided the means to survey complex protein mixtures with high sensitivity.

1.4.3 Mass spectrometry-based analysis of nuclear protein dynamics

High quality genome annotation has tremendously increased the potential of using mass spectrometry (MS)-based technologies for assessing the proteome. Subcellular fractionation to isolate the organelle of interest further improves this approach by reducing the complexity of the protein sample and has revealed the nuclear proteome in mammalian and plant cells (Andersen *et al*., 2005; Saitoh *et al*., 2004; Schirmer *et al*., 2005; Pendle *et al*., 2005; Bae et al., 2003; Calikowski *et al*., 2003). Due to the robustness of mass spectrometry as a tool for identifying large numbers of proteins and its potential in uncovering novel roles of an organelle, like the nucleus, in the next sections I will discuss the methodology used in MS-based proteomics experiments and how MS data are used in downstream analysis.

1.4.3.1 An introduction to mass spectrometry

MS-based protein analysis is a multi-step process and a flowchart of the different steps involved is shown in Figure 1.4. In the following subsections I will describe mass spectrometry methods for peptide fragmentation and subsequent sequence determination, the emergence of quantitative proteomics as a powerful tool for studying proteome dynamics and a method for the automated analysis of MS spectral data.

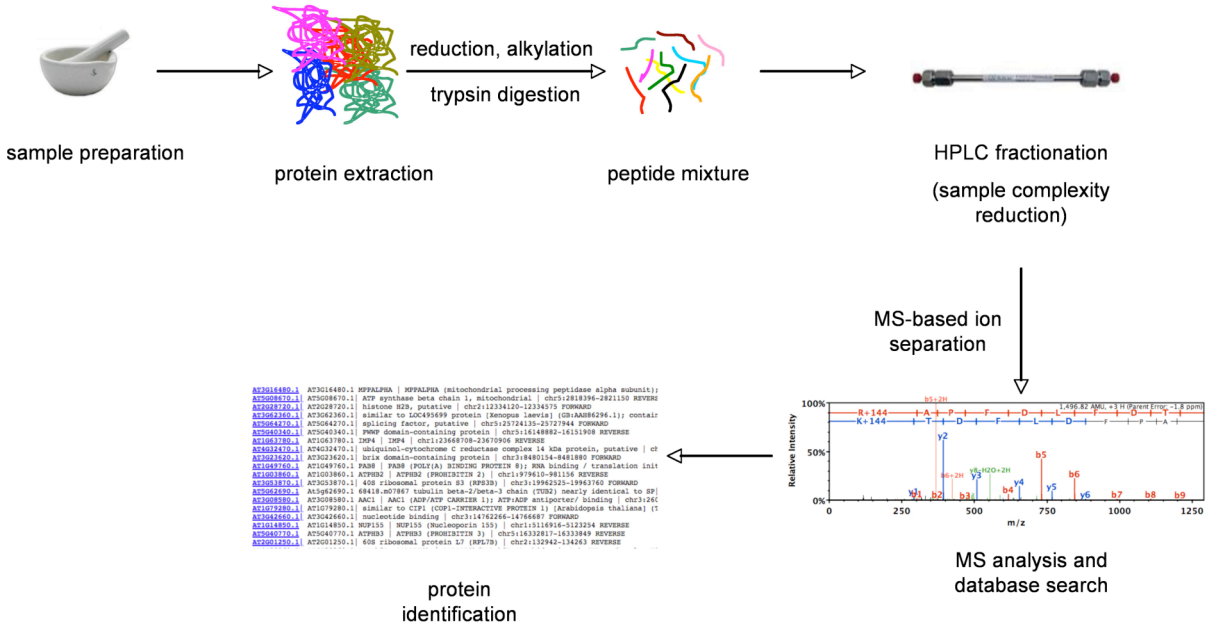


Figure 1.4 Flow-chart of MS-based proteomics analysis

MS-based protein identification starts with protein extraction from whole cells, subcellular organelles or complete organs, Protein extracts are then digested and samples for relative quantification can be labelled with different isobaric chemical tags and mixed. HPLC separation of these labelled peptides gives multiple less complex fractions, which can be separated more fully by an on-line HPLC on the front of an electrospray mass spectrometer. Separated peptide ions enter the mass spectrometer and eventually lead to the generation of mass spectra, which are fed into search engines to achieve protein identification by querying databases of theoretical mass spectra (see Section 1.4.3.2 for details).

1.4.3.2 MS-based methods for protein identification

Ionization methods

A major technological advance in the field of MS-based proteomics came about with the development of the "soft" ionization methods, electrospray ionization (ESI) (Fenn *et al*., 1989) and matrix-assisted laser desorption/ionization (MALDI) (Karas and Hillenkamp, 1988). In the case of the MALDI method, the analyte is mixed with an UV-absorbing matrix and positioned on an inert surface as discrete spots. Each spot is then irradiated with a laser leading to instant vaporisation of the matrix that transfers the non-volatile analyte into the gas phase (Figure 1.5A). Each time the laser fires, a packet of ions is accelerated, by electric potentials, down the flight tube and detected (see *Ion Separation Methods* in this Section). The spectra from many hundreds of shots are summed together to give the final spectrum.

During ESI, liquid is sprayed from the tip of a needle, in the presence of a high electrical field, leading to the dispersion of ionized droplets (Figure 1.5B). Solvent in these airborne droplets gradually evaporates bringing the charged ions closer, up to the point where Coulomb forces are stronger than the surface tension of the solvent and the droplets explode into ever-smaller ones. Successive rounds of solvent evaporation and droplet explosion ideally result in solvent droplets containing only one ion. Due to the nature of ionization, ESI produces a range of multiply charged ions. Following ionization, individual ions are packaged into the ion separation device, where they are separated based on their mass-to-charge ratio, and then exit to the mass spectrometer detector. One major difference with ESI is that MALDI produces singly charged ions, which can affect the degree of ion resolution by the mass spectrometer.

Sample prefractionation methods

In the case of complex protein mixtures, there are advantages in reducing the complexity of the sample prior to peptide ionization, in order to obtain the highest number of protein identifications later on in the mass spectrometer. The most common method used is the pre-fractionation of the peptide mixture using a high performance liquid chromatography (HPLC) column packed with a negatively charged resin that binds positively charged ions [i.e. a strong cation exchanger (SCX)]. This is done off-line away from the mass spectrometer. Increasing salt concentration is used to

progressively displace bound cations (peptides) based on charge. Fractions are dried out to remove organic compounds (25% acetonitrile is usually added to denature peptides and make them bind more reproducibly), redissolved in aqueous buffer and loaded onto an HPLC system with a reverse phase (RP) column on-line to an ESI mass spectrometer. The RP column is packed with a hydrophobic material (e.g. C18 resin) that binds peptides with an affinity proportional to a peptide's hydrophobicity. A novel on-line approach called "Multi-dimensional Protein Identification Technology" (MudPIT; Link *et al*., 1999), where the SCX and RP column phases are packed into a single column, affords full on-line automation of the separation of peptides prior to the ESI mass spectrometer. This has been widely applied to the identification of complex protein mixtures in mammals (reviewed in Yates *et al*., 2009) and plants (Lohrig *et al*., 2009; Feng *et al*., 2009; Maor *et al*., 2007).

All of these methods have their advantages and disadvantages. In MALDI the discrete spots of sample/matrix on the target can be archived and the user can come back later and repeat the analysis on exactly the same sample. For ESI, the sample is gone and to repeat the same analysis requires a replicate injection that may not be identical. MALDI is often not compromised by salty samples whereas ESI requires samples to be desalted first. For the aims of this project (see Section 1.8) I used the nano-ESI method (Brugger *et al*., 1997) that allows for the use of minimal amounts of sample without a compromise in the resolution of the analysis.

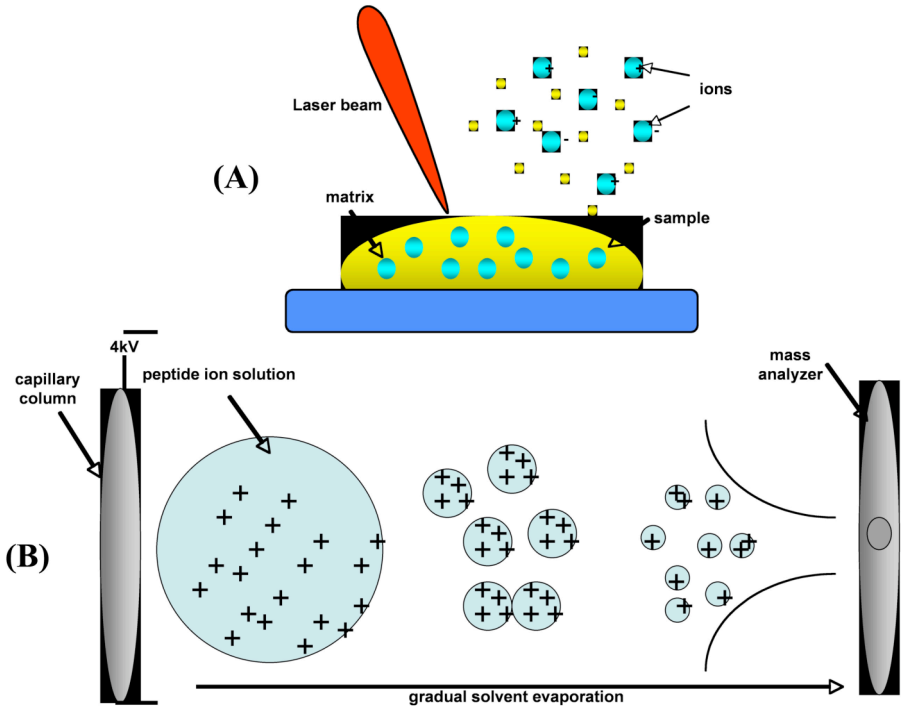


Figure 1.5 Ionization of a peptide mixture prior to mass spectrometry analysis

(A) The MALDI ionization method. A UV laser is fired at a mixture of co-crystallised peptides and matrix. UV ablation of the mixture results in the generation of mostly singly charged ions that are channelled to the mass analyzer. **(B)** The ESI ionization method. As the acidified peptide analyte exits the capillary,positively charged droplets are produced in the presence of high electric field. Heated gasevaporates the solvent and eventually singly- or multiply-charged ions are left which are guided into the mass analyzer.

Ion separation methods

Following ionisation in the source of the mass spectrometer, ions are moved into the analyser of mass spectrometer. There are three main types of analyser used for peptide work; Time-of-Flight (ToF), quadrupole and Orbitrap. In the first case, ions of equal energies enter a cylindrical tube under vacuum and their time of flight (ToF) along the length of the tube is measured. Since the ions are of equal energy, the ToF will depend on the mass of each ion (Figure 1.6A). Quadrupole mass analyzer consists of four cylindrical rods arranged in parallel to each other (Figure 1.6B). Opposite rods are connected electrically and a fixed radio frequency is applied to both pairs, superimposed by a direct current (DC) voltage (de Hoffmann and Stroobant, 2003). Depending on the mass-to-charge ratio, and for a given DC value, an ion can either reach the detector or collide with the rods. In this way, specific ions can be selected. A more advanced version is the quadrupole ion trap (Douglas *et al*., 2005) where ions are confined within the four rods and those of a specific mass-to-charge ratio are selectively released. The third type of ion separation system, the Orbitrap, consists of one cylindrical electrode and an inner spindle-like electrode (Figure 1.6C; Makarov, 2000; Hu *et al*., 2005). Injected ions orbit around the inner electrode in the presence of an electric field and the attraction of the ions to the inner electrode is balanced out by the centrifugal forces. In addition to cycling, ions oscillate along the axis of the inner electrode at a frequency inversely proportional to the square root of the mass-to-charge ratio (Makarov, 2000). The system detects the frequency of each oscillation and Fourier transformation is used to deconvolute the individual masses. Of the three ion-separation methods described, the Orbitrap generally has the highest mass accuracy (1-2ppm or better) and the highest resolving power (up to 200,000) (Makarov *et al*., 2006). Modern tandem mass spectrometers often use a combination of the above analyser types e.g. Quadrupole-Time-of-Flight machines (Q-ToFs) or Time-of-Flight / Time-of-Flight (ToF/ToF) combinations.

Protein identification

Mass spectra of peptides can be output as text files and act as input to specialist MS-based search engines. These can match and therefore identify mass spectral profiles by comparison with *in silico* digests of a proteome. Depending on the tool, mass spectral

matches are assigned a statistical value that indicates the significance of that match. The identified peptide is assigned to a protein or proteins and, based on the number of assigned peptides and scores given to each peptide, each protein is assigned a different significance score that is included in the program result output. Programs used for protein identification include Mascot™ (Perkins *et al*., 1999), SEQUEST (Gentzel *et al*., 2003), X!Tandem (Fenyo and Beavis, 2003) and each one of these uses a different algorithm for scoring peptide and protein identifications.

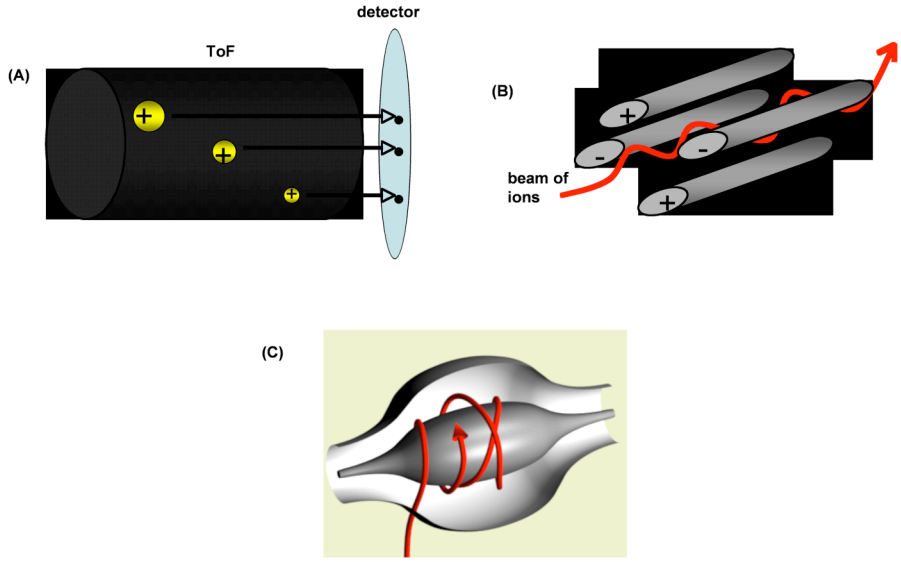


Figure 1.6 Commonly used ion separation methods (analysers) for proteomics analysis

(A) Time-of-Flight (ToF) ion separation method. Three ions are shown that have the same charge but different masses. The lighter ion reaches the detector first, followed by the middleweight ion and the heaviest ion reaches the detector last. **(B)** Quadrupole-based ion separation method. The beam of ions passes through the four rods and, based on the DC applied some ions will reach the detector and other will collide with the rods. **(C)** Orbitrap ion separation method. Cycling and oscillating ions around the spindle-like electrode process unique frequencies that relate to their mass-to-charge ratio, and these frequencies are used as a measure for mass analysis. Source of image in (C) is: http://upload.wikimedia.org/wikipedia/commons/8/8a/Orbitrappe.png

Because each of the search engines uses different scoring values comparison of results is difficult. PeptideProphet (Keller *et al*., 2002) overcomes this problem by converting the score from each of the programs into values on the same scale. PeptideProphet is now part of the Scaffold™ software (www.proteomesoftware.com) and was used for the proteomics analysis in Chapters 5 and 6. Since this algorithm provides important statistical strength to my results, it will be useful to briefly introduce the reader with the methodology followed by the PeptideProphet for assigning scores to peptides. Peptide score values from SEQUEST, X!Tandem or Mascot are converted by PeptideProphet to a "discriminant score", through the use of conversion equations that are specific for each search engine (see also Scaffold software manual). Plotting of the discriminant scores and curve fitting allows the algorithm to assign matches as "incorrect" and "correct", assuming that both types follow a normal distribution (Figure 1.7). After curve fitting, the algorithm uses Bayesian statistics to identify the discriminant score that matches the minimum probability (in my case this was 99%) for assigning a peptide as "correct" (this is denoted with the black vertical line in Figure 1.7).

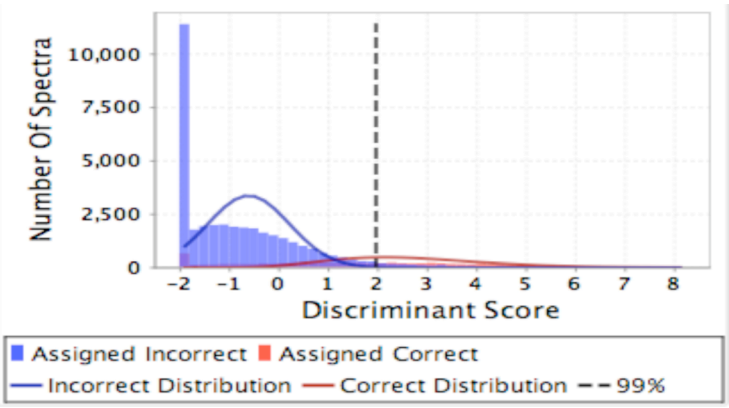


Figure 1.7 Assignment of discriminant score to mass spectra

Search results using the SEQUEST search engine were loaded into Scaffold and, using the PeptideProphet algorithm, a discriminant score was assigned to each of the spectra. This is followed by the determination of "correct" and "incorrect" identification of peptides. The black line matches the discriminant score with the lowest peptide probability (set by the statistics and in my case it was 99%) at which the peptide is assigned as "correct".

1.4.4 Quantitative proteome dynamics

Determination of organelle proteomes using MS-based technology has the potential to provide new insights into protein dynamics within those organelles. To study proteome dynamics, accurate quantitative analysis is required. Relative quantitative methods can be broadly grouped into 4 groups: gel-based methods, label-free methods, metabolic labelling methods and chemical labelling methods. Since I used a chemical labelling method (iTRAQ), I will refer briefly to the first 3 groups of quantitative methods and focus in more detail on the last one.

1.4.4.1 Gel-based approaches for protein quantification

Gel-based methods were the first approach used in protein quantification (Unlu *et al*., 1997). In the classic 2D gel approach, samples are separated on different gels, spot intensities compared, spots excised, digested with trypsin and identified by MS. In the Differential in Gel Electrophoresis (DIGE) approach, two samples are labelled with different fluorescent dyes [Cysteine-3 (Cys-3) and Cys-5] and then mixed before separation on a single 2D gel, overcoming the reproducibility problem inherent to previous gel-based approaches. Relative protein quantification depends on the ratio between fluorescence intensities of the two dyes within a given spot, which is then excised for MS analysis. Plant biologists have used these approaches extensively to quantify proteome responses to stress and characterise organelle proteome dynamics during growth (Nilo *et al*., 2010; Ingle *et al*., 2007; Bae *et al*., 2003; Kleffmann *et al*., 2007; Majeran *et al*., 2008). Nevertheless, protein quantification using DIGE analysis has some disadvantages. First, protein identification and protein quantification are not coupled; that is, differentially expressed spots should be first excised and then go though MS analysis for protein identification, increasing labour time. Second, loss of quantitative information may occur, due to the fact that one spot could contain more than one proteins. In this case, if one of the proteins is up-regulated under a treatment and the other one is down-regulated then the colour of the spot will not change. This will lead to missing of quantitative information for potentially important proteins for the processes under study. These inherent disadvantages, led me to decide on using chemical labelling (Section 1.3.4.4) of peptides that could be coupled to LC/MS and

achieve a high resolution quantitative analysis of the *Arabidopsis* nuclear proteome, as cells move from proliferation to quiescence.

1.4.4.2 Label-free approaches for protein quantification

In label-free approaches, quantification is made at the level of either the protein or the peptide. In the first case, the parameter used is the average ion count of total peptide spectra identified for a given protein in each sample, whereas in the second case the ion intensity of each peptide is measured by calculating the peak volume corresponding to each identified peptide at a given m/z ratio. In this way peak volumes for each peptide can be calculated across the LC-MS/MS runs within an experiment, and then be quantitatively compared with the same dataset from a different sample, outputting an average quantitative value for the corresponding protein. Despite limitations (Old *et al*., 2005), the label-free approach has been employed different tissues in *Arabidopsis*, revealing dominant gene models and tissue-level proteome dynamics (Baerenfaller *et al*., 2008).

1.4.4.3 Metabolic labelling methods in quantitative proteomics

Metabolic labelling refers to the labelling of the whole cell proteome during active cellular metabolism. The label can be either a heavy isotope, such as ^{15}N in the form of ^{15}N-ammonium salts (Conrads *et al*., 2001), or an amino acid containing a stable isotope, such as lysine containing ^{13}C [SILAC (stable isotope labelling of amino acids in culture; Ong *et al*., 2002)], and they are supplied during the growth of a cell culture. Following label incorporation, labelled and unlabelled protein samples are mixed at the earliest stage before any other type of processing commences, thus minimising technical variance issues. SILAC has been used many times in mammalian cells for studying *e.g.* signalling, protein-protein interactions, protein dynamics and micro-RNA-dependent proteome changes (reviewed in Schulze *et al*., 2010). In plants, a single study used SILAC-mediated phospho-proteomic analysis of *Arabidopsis* cell cultures, but the efficiency of label incorporation was too low for reliable quantification (Gruhler *et al*., 2005). An alternative labelling method grows cells in either normal nitrogen (^{14}N) or

with the heavier isotope ^{15}N and compares the two samples. Almost 100% ^{15}N incorporation was achieved in *Arabidopsis* cell cultures (Lanquar et al., 2007) as well as in *Arabidopsis* and tomato plants (Hebeler *et al*., 2008). In the case of plants, the labelling method is called either SILIP (stable isotope labelling *in planta*; Schaff *et al*., 2008) or HILEP (hydroponic isotope labelling of entire plants; Bindschedler *et al*., 2008). Thus, heavy isotope labelling is a very robust method for studying proteome dynamics in plants during normal growth or alterations of physiological conditions.

1.4.4.4 Chemical-based labelling methods

A wide range of chemical reagents has been generated for chemical labelling of proteins but only a few have been used in biological research (reviewed in Ong and Mann, 2005). Two widely used methods are the "isotope-coded affinity tag" (ICAT; Gygi *et al*., 1999) and the "isobaric tag for relative and absolute quantification" (iTRAQ; Ross *et al*., 2004). The chemical reagent used in ICAT contains a reactive group that targets cysteines, a linker region and a biotin group that allows the recovery of labelled peptides. However, the lack of cysteines in some peptides compromises coverage and gives lower confidence in the quantification results. Also, the ICAT method is limited to only two samples per experiment, rendering it less flexible for experimental designs employing multiple treatments such as, time-course experiments.

iTRAQ can label up to 8 independent samples, using a unique reagent for each sample, thus allowing for the design of multi-sampled experiments. Each iTRAQ reagent labels primary amines of amino acids, generating a population of peptides that are labelled at their amino terminus as well as at any lysine residues. The general structure of the reagent consists of an amine-reactive group, a charged reporter group that is unique for each of the eight reagents and a balance group as a linker between the reporter and amine-reactive groups (Figure 1.8). The masses of the individual reporter ions range from 114 to 121Da whereas the linker group masses range from 31 to 24Da, so that the total mass of the isobaric tag in each of the eight iTRAQ reagents is 145Da. This special mass combination facilitates peptide quantification in the MS/MS mode of the mass spectrometry analysis (see below).

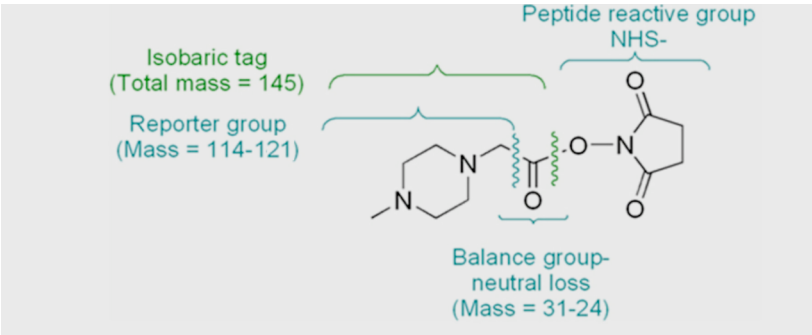


Figure 1.8 Structure of the iTRAQ™ reagent.

The structure of the reagent is divided into 3 parts. The first part is the peptide reactive group; this is attached to the balance group followed by the reporter species. In total, the balance and reporter groups have a mass of 145kDa. During MS analysis the balance group is cleaved from the amino acid reactive moiety. When parent ions enter the MS/MS mode, the vast majority of balance-reporter pairs break at the joining point releasing the reporter ion. (Source: iTRAQ manual, Applied Biosystems)

ICAT and iTRAQ also differ in that in the former method, as mentioned above, quantification is undertaken in the MS mode whereas for iTRAQ it occurs in the MS/MS mode. Upon peptide ionization, iTRAQ-labelled peptides that are common between the different samples are indistinguishable at the MS mode due to the addition of the isobaric tag to their sequence. Upon parent ion fragmentation by collision-induced dissociation (CID), the singly charged reporter group is cleaved from the isobaric tag and the neutral linker group is lost. The reporter ions from the isobaric tag are detected at the low end of the mass spectrum, (at an m/z between 114 and 121 depending on which reagent has been used) and at the same time peptide identification is taking place. The ion abundance of the reporter groups for the same peptide from different samples give a relative quantitative ratio and, consequently, quantitative information for the amount of protein present in each sample (Figure 6.2). At least two peptides, preferably many more, from each protein are needed for reliable protein quantification. Complete labelling of all peptides in a mixture ensures that all the

proteins in the sample can be potentially quantified, in contrast to ICAT where only cysteine-containing peptides are labelled. Also, complete peptide labelling with iTRAQ may facilitate simultaneous quantification of post-translational modifications (PTMs).
The iTRAQ approach has been used to determine protein distributions across different sub-cellular organelles and structures (Zhu *et al*., 2009; Dunkley *et al*., 2006), to determine quantitative changes in plant proteomes during biotic or abiotic stresses (Marsh *et al*., 2010; Lippert *et al*., 2009; Schneider *et al*., 2009; Nuhse *et al*., 2007; Patterson *et al*., 2007; Jones *et al*., 2006) and to follow proteome changes during physiological processes, such as fruit ripening (Lucker *et al*., 2009).

1.4.5 Plant nuclear proteomics

Systematic analysis of the nucleolar proteome, using LC/MS combined to *in vivo* validation with the use of fluorescence-tagged proteins, has allowed systematic analysis of the nucleolar proteome (Pendle *et al*., 2005) and the same is potentially possible for the nucleus. Up to date, a considerable amount of studies of nuclear proteome in plants has been published, and these include nuclei from leaves and cultured cells of *Arabidopsis* (Jones *et al*., 2009), from rice endosperm (Bae *et al*., 2003), from chickpea seedlings and from *Medicago* seeds (Calikowski *et al*., 2003; Khan and Komatsu, 2004; Pandey *et al*., 2006; Repetto *et al*., 2008). The number of proteins identified in these studies varied (Table 1.1) but they illustrate the potential to provide a deeper understanding of how the nuclear proteome is related to metabolic processes. For example, in *Medicago* seeds just prior to seed filling, almost 30% of the nuclear proteome consists of ribosomal and RNA processing proteins (Repetto *et al.*, 2008), an indication of the high translational activity associated with storage protein accumulation. Conversely, in nuclear extracts from *Arabidopsis* leaves and chickpea seedlings, these types of proteins were less abundant whilst proteins related to cell organization and carbohydrate metabolism, which are necessary for sustaining plant development, are relatively more abundant (Bae *et al.*, 2003; Pandey *et al*., 2006). However, with the exception of nucleolus (Pendle *et al*., 2005), these studies fail to capture the expected complexity of the nucleus, and this complexity is part of my work at this thesis aims to reveal (Section 1.8).

Plant species	Material	Proteome	Method	No. of proteins	Reference
Arabidopsis thaliana	cell culture	nuclear phosphoproteome	SCX;LC-MS/MS	317	Jones *et al*., 2009
Arabidopsis thaliana	leaf	nuclear	2D-SDS; MALDI-ToF-ToF	54	Bae *et al*., 2003
Arabidopsis thaliana	cell culture	nucleolus	LC-ESI-Q-ToF	217	Pendle *et al*., 2005
Arabidopsis thaliana	cell culture	nuclear matrix	1D-SDS; ESI-Q-ToF * 2D-SDS; MALDI-ToF-ToF	36	Calikowski *et al*., 2003
Medicago trancatula	seed	nuclear	1D-SDS; ESI-Q-ToF	143	Repetto et al., 2008
Cicer arietinum	seedling	nuclear	2D-SDS; ESI-Q-ToF	150	Pandey et al., 2006
Cicer arietinum	seedling	nuclear-dehydration stress	2D-SDS; ESI-Q-ToF	147	Pandey et al., 2008
Oryza sativa	cell culture	nuclear	2D-SDS; MALDI-ToF-ToF	190	Khan and Komatsu, 2004
Oryza sativa	cell culture	chromatin proteome and phosphoproteome	2D-SDS; MALDI-ToF-ToF	269	Tan *et al*., 2007
Oryza sativa	endosperm	nuclear	SCX; MALDI-ToF-ToF	468	Li *et al*., 2008

Table 1.1 List of proteomic studies of the plant nucleus

1.4.6 Proteomics of the nucleolus

The most prominent region inside most nuclei is the nucleolus, which is the site of transcription, processing and assembly of ribosomal RNAs (rRNAs) into ribosomal subunits before their export to the cytoplasm (Fatica and Tollervey, 2002) and these features provide a means for physical isolation. A comprehensive study of the *Arabidopsis* nucleolar proteome (aNP) identified 217 proteins (Pendle, *et al*., 2005). Comparison of the plant nucleolar proteome with the human nucleolar proteome (hNP; Andersen *et al*., 2002, 2005) identified both conserved and organism-specific proteins. Initial protein homology comparisons indicated that 69% of the identified plant nucleolar proteins have a counterpart in the hNP. Another 23 putative plant-specific proteins provided a possible basis for known structural differences between plant and animal nucleoli (Shaw and Jordan, 1995), such as differences in rRNA gene transcription/regulation and the presence of a nucleolar cavity in plants. The plant-specific proteins included putative DNA-interacting proteins, perhaps reflecting differences in rDNA gene transcription. Exon-junction complex (EJC) proteins were

also found in the plant nucleolar preparation and this was confirmed by expressing GFP fusion proteins in living cells (Pendle, *et al*., 2005). The association of some EJC components with the nucleolus in animal cells was also reported (Andersen *et al*., 2005). The above studies demonstrate the importance of high-throughput proteomics studies at identifying novel roles of a structure or an organelle, as well as identifying inter-connectedness between cellular compartments.

1.5 Regulatory kinases as components of nuclear bodies

Many nuclear-based processes, such as DNA replication and gene expression, are controlled by reversible phosphorylation (Vermeulen *et al*., 2009) and many protein kinases are known to reside in the nucleus. Some, such as CDKA from *Arabidopsis* are present both in the cytoplasm and in the nucleus (Weingartner *et al*., 2001), but others reside solely in the nucleus. For example, CDKC1, CDKC2, CDKD1 and CDKD3 in *Arabidopsis* are nuclear proteins that locate to speckle-like structures (Kitsios, 2007; Kitsios *et al.,* 2008) and do not have a direct role in cell cycle progression (Menges *et al*., 2003). Understanding the role of these nuclear kinases would help us to obtain a more holistic view of the functional importance of the CDK family in plants. In the following sections I will present a literature review about the role of CDKs in important nuclear processes of the eukaryotic cell.

1.5.1 Cyclin dependent protein kinases

Cyclin-dependent kinases (CDKs) are a large family of heterodimeric serine/threonine kinases that belong to the CMGC group of eukaryotic protein kinases. Apart from CDKs, this group includes mitogen-activated kinases (MAPKs), glycogen synthase kinases (GSKs) and CDK-like kinases (CLKs) (reviewed by Champion *et* al., 2004; Hanks, 2003; Manning *et* al., 2002;). By definition, CDKs require the binding of cyclins for their activity (although additional proteins bind and activate CDKs in the absence of cyclins (reviewed by Nebreda, 2006). Cyclins function as the regulatory subunit of the cyclin/CDK complex whereas CDKs act as the catalytic subunit.

1.5.1.1 Discovery of the first cyclin dependent protein kinase

The first comprehensive description of cell division was published by Strasburger in 1880, revealing that nuclear division immediately precedes division of the cell. DNA replication occurs much earlier in the cell cycle (Swift, 1950). The non-mitotic stages are collectively known as interphase (see below), a term that was coined by H. Lundergardh in 1913, whereas the process of nuclear division was called "mitosis". Mitosis is divided into five different stages: prophase, metaphase, anaphase, telophase and cytokinesis (Strasburger in 1884, "telophase" was introduced by Martin Heidenhain in 1894). Interphase also has several distinct stages including Gap1 (G1), DNA synthesis or replication (S) and Gap2 (G2). During G1 and G2 stages, the cell may monitor its size and physiological environment and can "decide" whether to progress to S phase and mitosis, respectively, or to delay progression or to exit the cell cycle and enter quiescent or differentiated states.

In 1968, Van't Hoff observed that when pea root tip were starved of sucrose the cells stopped dividing. When sucrose was re-supplied to the root tips the cells resumed cell division. He suggested that cells have an inherent mechanism that allows them to monitor nutrient levels and then "decide" whether these levels are adequate for progressing to division. This mechanism is known as a "checkpoint". During the eukaryotic cell cycle, there are at least 3 checkpoints; the G1/S checkpoint, the G2/M checkpoint and spindle assembly checkpoint. Amongst other things, cell size is checked at G1/S before the cell proceeds to DNA replication. At the G2/M checkpoint, the fidelity and completeness of DNA replication is assessed before the cell enters mitosis. The spindle assembly checkpoint is responsible for making sure that chromosomes are properly aligned to the equatorial plane before any attempt is made to separate them.

A revolution in the understanding of cell cycle control came with the discovery of the first core cell cycle regulator in fission yeast (*Schizosaccharomyces pombe*) (Nurse and Bissett, 1981). The gene was called cell division control-2 (CDC2), encoded for a protein kinase and it was found to control progression during both G1/S and G2/M checkpoints (Nurse and Bissett, 1981). Subsequent to the discovery of *CDC2*, Evans *et al*., (1983) described a protein in sea urchin that was destroyed every time the cell

divided. Due to its cyclic behaviour, the protein was called cyclin. A decade after its discovery, cyclin was shown to bind and regulate CDC2 activity. This discovery led to the adoption of a new nomenclature regarding CDC proteins, and subsequently discovered CDC2-related proteins were named cyclin-dependent kinases (CDKs).

1.5.1.2 Structure of cyclin-dependent kinases

Crystallographic studies of CDKs have provided insight into their protein structure and the conformational changes that occur upon binding to cyclin partners or other CDK-associated proteins (Day *et al*., 2009; Lu *et al*., 2006; Lolli *et al*., 2004; Brotherton *et al*., 1998; De Bondt *et al*., 1993). The core catalytic domain of all protein kinases consists of 250-300 residues and it has a structure similar to that of human cyclic-AMP (Knighton *et al*., 1991) and CDK2 (De Bondt *et al*., 1993).

Mammalian CDK2 was the first CDK whose crystal structure was resolved (De Bondt *et al*., 1993). As all CDKs are structurally related, I will use the structure of CDK2 for describing the basic structure of the CDK protein family. The catalytic core of the CDK2 structure consists of an N-terminal lobe and a C-terminal lobe (Figure 1.9), characteristic of all protein kinases. The two lobes are separated by a stretch of 40 amino acids that form the catalytic cleft of the kinase, where ATP and substrate binding sites reside. The N-terminal lobe consists of five anti-parallel β-strands and a single α-helix, whereas the C-terminal lobe has six α-helices and a small ribbon (De Bondt *et al*., 1993). The N-terminal lobe contains a glycine-rich motif and the PSTAIRE motif both of which are important for kinase activity. The glycine-rich sequence facilitates binding of the kinase to ATP phosphates and also contains two conserved residues, Thr14 and Tyr15, whose phosphorylation state is important for CDK activation (Gu *et al*., 1992; Morgan, 1995). The PSTAIRE motif is a CDK hallmark peptide sequence and defines part of the site where cyclin binds. Whereas the N-terminal lobe is the CDK activation domain, the C-terminal lobe is responsible for the catalytic activity of the kinase. The C-terminal lobe contains the catalytic loop, which is responsible for the transfer of a phosphate group from ATP to the substrate protein. The T-loop acts as an inhibitor of CDK activation until the conserved threonine residue (hence the name "T-loop") is

phosphorylated and allows access of the substrate to the catalytic cleft (Jeffrey *et al*., 1995).

1.6 Cyclin-dependent kinases have diverse cellular functions

Cyclin-dependent kinases are well-established regulators of the eukaryotic cell cycle (Nigg, 1995; Morgan, 1997; Nurse, 2002) and have profound effects on nuclear structure and function. Activation of the mitotic CDKA leads to disassembly of many nuclear structures, including the nucleolus and many other structures involved in gene expression. Other CDK's appear to have roles in specific aspects of nuclear function, including gene transcription, pre-mRNA processing, translation and DNA damage responses. In this section I will focus on these "non cell cycle" roles of CDKs. The roles of each CDK group and the respective homologues in plants, yeast and animals are collectively presented in Table 1.3 at the end of this section and can be used in reference to the text. Also a phylogenetic tree of all the CDKs involved in pre-mRNA processing is provided at Figure 1.12, in order to visualise sequence-based relationships between the different CDK families.

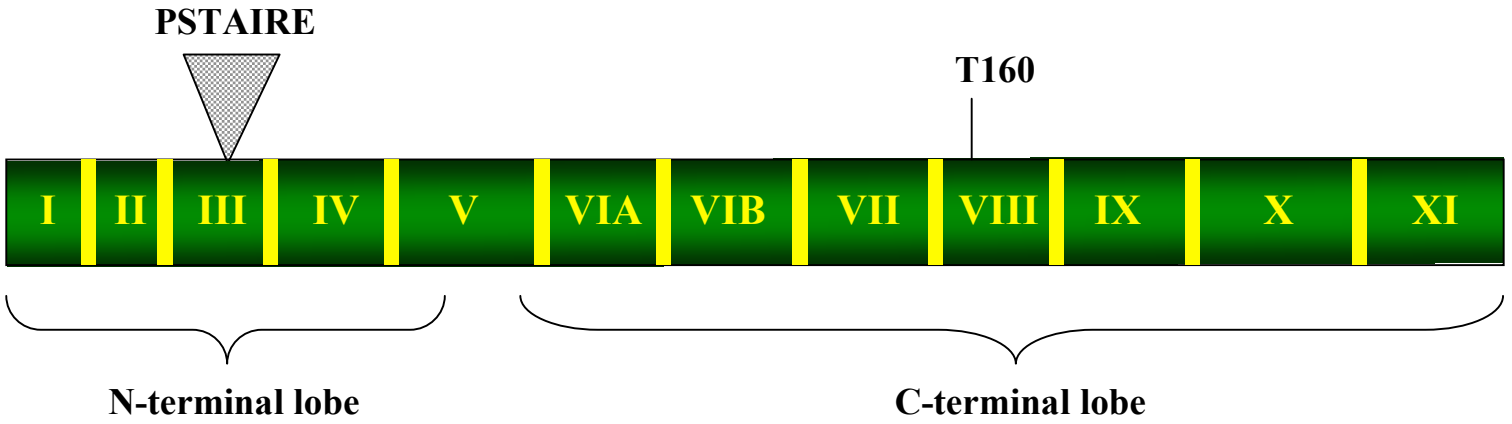


Figure 1.9 Graphical representation of domain organisation in a eukaryotic kinase.

The catalytic domain of the kinase consists of an N- and a C-terminal lobe. In total, there are 12 domains (numbered with latin numerals) separated by less conserved amino acid sequences (yellow bars). Conserved motifs and functionally important amino acids for CDK activity are shown above the sequence (see text). Amino acids are numbered based on the human CDK2 sequence.

1.6.1 CDKs and transcription initiation

Gene transcription is a highly complex and multi-step process involving chromatin modification, assembly of transcription initiation complexes on gene promoters, pre-mRNA transcript initiation, elongation and termination. The first step for transcription initiation is the formation of the pre-initiation complex (PIC) on the promoter DNA of the gene. Assembly of the PIC is a step-by-step process (Buratowski *et al*., 1989; Ranish and Hahn, 1996; see Figure 1.10). Assembly of the transcriptional apparatus for transcription initiation begins with the recruitment of general transcription factors (GTFs) to gene promoters. TFIID is first recruited and binds to the DNA promoter, followed by TFIIA and TFIIB and then RNA Polymerase II associated with TFIIF. PIC is finally established with the loading of TFIIE and TFIIH. Even though PIC is sufficient to drive transcription *in vitro*, it is unable to respond to activators. This observation suggested the existence of a bridge complex that responds to activator molecules and signals the PIC to initiate transcription. This complex was first isolated in yeast and termed "Mediator" (Kelleher *et al*., 1990; Flanagan *et al*., 1991; Kim *et al*., 1994). The Mediator complex is conserved from yeast to plants and mammals and is important in mediating the activation of RNA Polymerase II that leads to promoter clearance and transcript elongation (Boube *et al*., 2002; Max *et al*., 2007).

RNA Polymerase II (PolII) is the enzyme responsible for pre-mRNA transcription in eukaryotic cells. It is composed of two large subunits and a collection of smaller subunits. The unique characteristic of PolII is the presence of a domain at the C-terminus of its large subunit, comprised of tandem repeats of the consensus sequence Tyr-Ser-Pro-Thr-Ser-Pro-Ser (Corden, 1990) and termed the C-terminal domain (CTD). The number of this hepta-peptide repeat varies across eukaryotes, from 26-27 in yeast, 42 repeats in *Caenorhabditis elegans*, 44 in *Drosophila melanogaster* to 25 in *Arabidopsis thaliana*.

Reversible phosphorylation of the CTD is an important regulatory mechanism for transcription initiation, pre-mRNA processing and transcription termination. The phosphorylation status of the CTD dictates its binding specificity (Hirose and Manley, 2000). The CTD is phosphorylated by the TFIIH and positive transcription elongation b (P-TEFb) transcriptions factors. In mammals, the catalytic subunit of TFIIH is the CDK7/CyclinH whereas for P-TEFb is the CDK9/CycT (Serizawa *et al*., 1995;

Marshall *et al*., 1996). Further to the involvement of CDK7- and CDK9- containing complexes in the regulation of gene transcription, new CDK/Cyclin complexes have been found to contribute to the complex regulatory network of gene transcription, pre-mRNA processing and mRNA translation. In the next sections I will discuss the role of each CDK/Cyclin complex in this fundamental cellular process.

1.6.2 The CDK7/CyclinH group

CDK7/CyclinH complex exhibits a dual role in eukaryotic cells. One role is to act as a CDK-activating kinase (CAK) by phosphorylating Cdc2/CyclinA and Cdc2/CyclinB complexes in Drosophila embryos leading to their activation (Larochelle *et al*., 1998). The second role is to participate in the TFIIH transcription factor complex as the catalytic subunit that phosphorylates the Ser-5 of the CTD in PolII (Serizawa *et al*., 1995; Shiekhattar *et al*., 1995). Substrate specificity is dictated by its interaction with the assembly factor MAT1; in the absence of MAT1, CDK7/CylinH complex shows CAK activity but when in association with MAT1, it participates in the TFIIH factor and phosphorylates the CTD (Rossignol *et al*., 1997; Yankulov *et al*., 1997). In yeast, the Kin28 protein is the catalytic subunit of TFIIH and was found to phosphorylate the Ser-5 of the PolII CTD domain and to have a CAK activity (Feaver *et al*., 1994; Hengartner *et al*., 1998).

Binding of unphosphorylated CTD to components of the PIC, such as the TATA-binding protein (TBP) and the Mediator complex, recruits PolII to gene promoters (Usheva *et al*., 1992; Myers and Kornberg, 2000). Also, unphosphorylated CTD has high affinity for DNA and the CTD/DNA complex is a better substrate for CDK7/CyclinH than free CTD (Lolli, 2009). Upon phosphorylation at Ser-5, CTD dissociates from DNA and from components of the PIC, licensing PolII to initiate transcription (Kobor and Greenblatt, 2002; Oelgeschlager, 2002). Transcription initiation-competent PolII is not allowed to proceed to elongation due to its association with DRB (5,6-dichlorobenzimidazole 1-β-*D*-ribofuranoside)-sensitivity-inducing factor (DSIF) and the negative elongation factor (NELF), the two inhibitors of the elongation phase of transcription (Wada *et al*., 1998). Transcript elongation is established when DSIF and NELF dissociate from PolII after being phosphorylated by CDK9/CyclinT complex (see Section 1.5.5 and Figure 1.10).

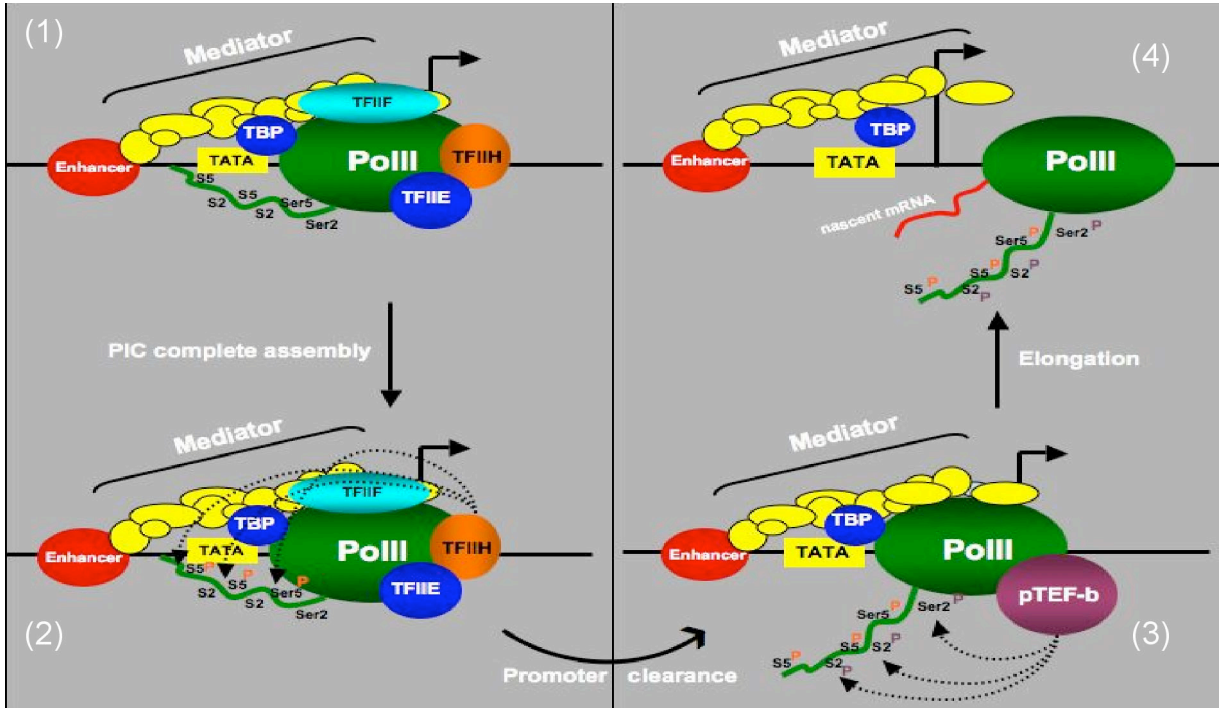


Figure 1.10 A simplified model of nascent transcript initiation in the mammalian nucleus.

Upon PIC assembly, PolII and TFIIH are recruited to the gene promoter where PolII binds to TATA-binding protein (TBP) in a Mediator-dependent manner (1). In the next step, the catalytic subunit of TFIIH transcription factor, CDK7/CycH, phosphorylates the Ser-5 of the CTD domain and licences PolII for promoter clearance (2). Phosphorylation on the Ser-2 of CTD by the CDK9/CyclinH complex (3), the catalytic subunit of the pTEF-b transcription factor, initiates transcript elongation (4).

In a phylogenetic analysis of eukaryotic CDK protein sequences, a group of plant CDKs was found to share high homology with the mammalian CDK7 (Guo and Stiller, 2004). Based on proposed nomenclature by Joubes *et al*., (2000) this group was designated as CDKD and consisted of three *Arabidopsis* CDK7-like proteins (*At*CDKD1-3) and a rice CDK7-like protein (*Os*CDKD1). *Os*CDKD1 was found to phosphorylate the rice homolog of CDK1, acting as a CAK, and the *Arabidopsis* PolII CTD domain (Hata, 1991; Yamaguchi *et al*., 1998). Thus, *Os*CDKD1 was the first identified ortholog of mammalian CDK7 in plants. The presence of three CDK7-like proteins in *Arabidopsis* indicated a more diverse role for this protein family in plants. Indeed, the three members were found to have considerable functional divergence. CDKD2 and CDKD3 were both found to have CAK and CTD-kinase activities that were enhanced *in vitro* when incubated with *At*CyclinH;1, the plant homolog of human CyclinH (Yamaguchi *et al*., 2000; Shimotohno *et al*., 2003, 2004, 2006). On the other hand, CDKD1 showed very low affinity for binding *At*CyclinH1 *in vitro* and exhibited very low kinase activity *in vivo*. These data suggest that the functional role of CDKD1 has diverged from the other two members of the family, with its role still remaining unresolved.

1.6.3 The CDK8/CyclinC group

The mammalian CDK8 protein binds to CyclinC and, together with Med12 and Med13 proteins, form active CDK8-CyclinC complexes that can directly affect the activity of major components of the PIC, such as TFIIH transcription factor and PNA PolII (Hengartner *et* al., 1998; Akoulitchev *et* al., 2000). In mammals, CDK8-CyclinC was found to suppress transcriptional re-initiation not by acting on the CTD, but by preventing Mediator-PolII complex formation (Elmlund *et al*., 2006; Knuesel *et al*., 2009a). Disruption of the Mediator-PolII complex resulted in improper assembly of the PIC on gene promoters and, thus, inhibition of transcription re-initiation. Thus, CDK8-CyclinC provides a fine-tuning, allowing the cell to readily activate/reactivate or suppress gene transcription in response to internal or external stimuli (Knuesel *et al*., 2009a). Contrary to its indirect inhibition of transcription in mammalian cells, the yeast CDK8 homologue, Srb10, was found to directly phosphorylate CTD-Ser of PolII prior

to the CTD binding to the DNA and, thus, preventing formation of PIC (Hengartner *et al*., 1998).

Recently, a surprising additional property of CDK8 ia histone kinase activity that led to transcriptional activation, contrasting with the well-known inhibitory effect of CDK8 in gene transcription (Knuessel *et al*., 2009a, b). CDK8 phosphorylated the Ser-10 of histone H3; a histone modification known to favour transcriptional activation (Nowak and Corces, 2000; Knuesel *et al*., 2009b). Thus, CDK8-CyclinC appears to regulate gene transcription either in association with Mediator, or in its own right functioning at the level of chromatin. Since most transcription factors target the Mediator complex, CDK8-CyclinC association with the Mediator implies regulation of gene transcription over multiple transcriptional sites (Knuessel *et al*., 2009b).

Apart from acting as a switch affecting global gene transcription, CDK8 also specifically regulates the Wnt/β-catenin pathway that is aberrantly activated during colorectal cancer development. Ip-regulation of the β-catenin pathway is CDK8 dependent and leads to proliferation of colon cancer cells (Firestein *et al*., 2008). Moreover, CDK8 suppresses expression of the transcription factor, E2F1, a transcriptional target of retinoblastoma (RB) tumour suppressor protein (Morris *et al*., 2008). E2F1 is a suppressor of β-catenin-dependent transcription and, thus, an inhibitor of colon cancer development. The suppression of E2F1 transcription suggests that CDK8 is an oncogene, involved in regulating both Wnt/β-catenin and pRB-signalling pathways in favour of proliferating colon cancer cells (Morris *et al*., 2008).

1.6.4 CDK8-like proteins in plants

In plants, the closest homolog of CDK8 is CDKE (Joubes *et al*., 2000). The *Arabidopsis* CDKE, encoded by the HUA ENHANCER 3 (HEN3) gene, regulates leaf cell expansion and cell fate specification in floral meristems (Wang and Chen, 2004). Similarly to yeast Srb10, CDKE exhibits a CTD kinase activity and *hen3* mutants showed transcriptional up-regulation of floral homeotic genes *AGAMOUS*, *APETALA1* and *APETALA2,* suggesting that CDKE inhibits gene transcription. An interesting difference between *Arabidopsis* CDKE and its animal counterpart is that, instead of

binding to a C-type cyclin, CDKE interacts with a D-type cyclin in yeast two-hybrid (Wang and Chen, 2004).

Gonzalez and colleagues (2007) found a Mediator/CDK8-subcomplex in *Arabidopsis* that mimics the function of CDK8 in suppressing gene transcription. Identification of its role was achieved by studying the mechanism of action of the transcriptional co-repressor LEUNING (LUG), a member of the Gro/TLE transcription co-repressor family conserved in all eukaryotes. Repressor activity was found to be impaired in a yeast strain depleted in the Srb10, the yeast counterpart of CDK8, suggesting that activity of LUG depends on the presence of CDK8 (Gonzalez *et al*., 2007). Identification of proteins homologous to mammalian Mediator and CDK8-subcomplexes in *Arabidopsis* suggested that LUG exerts its repression activity in a similar fashion to its mammalian counterpart. Moreover, evidence from yeast two-hybrid assays and *in vivo* interaction assays showed that CDKE interacted with LUG and SEUSS; a mediator of LUG interaction with transcription factors (Gonzalez *et al*., 2007; Sridhar *et al*., 2004, 2006). Thus, LUG may act as a transcriptional repressor by interacting with the plant Mediator via CDKE, mirroring the mechanism of widespread transcriptional repression in mammals.

1.6.5 The CDK9/CyclinT group

Mammalian CDK9 has been found to interact with two types of cyclins, T-type and K-type, with T-type being its predominant partner (Garriga *et al*., 1996; Peng *et al*., 1998). CDK9 is the catalytic subunit of the positive transcription elongation factor b complex (P-TEFb) that is responsible for Ser-2 phosphorylation at the CTD domain of PolII (Marshall *et al*., 1996; Price, 2000). Ser-2 phosphorylation is the decisive step in the process of producing an elongation-competent PolII holoenzyme, with the first step being the phosphorylation of CTD at the Ser-5 position by the CDK7/CyclinH complex (see Section 1.3.2). DNA binding of the unphosphorylated CTD, upon recruitment of PolII to the PIC by the Mediator, assures the ordered phosphorylation of the CTD by TFIIH and P-TEFb at the Ser-5 and Ser-2 sites, respectively. Binding of the CTD to DNA makes the Ser-2 residue inaccessible to the CDK9/CycT complex. Dissociation of the CTD/DNA complex is achieved via phosphorylation of Ser-5 by the CDK7/CyclinH

complex that shows high affinity for binding CTD/DNA complexes, compared to free CTD (Loli, 2009).

Interest in the mammalian P-TEFb complex has been intense due its potential as an anti-HIV drug target. The CDK9/CyclinT complex has been found to regulate the transcription of HIV genes via binding to the HIV transcription factor TAT (Garriga *et al*., 1998; Wei *et al*., 1998) whereas incubation of mammalian cells with the P-TEFb-specific inhibitor, flavopiridol, suppressed the HIV transactivation.

Budding yeast contains two P-TEFb homologs; the Ctk1/CtDk-I (kinase/cyclin respectively) and Bur1/Bur2. The two complexes act in sequence, with Bur1/Bur2 being the catalytic unit that acts first, followed by the second P-TEFb homolog. Bru1/Bru2 do not have any CTD-kinase activity but mediate chromatin modifications by catalyzing ubiquitination of histone H2B and trimethylation of H3K4 (Wood *et al*., 2005). H2B ubiquitination acts as a signal for SAGA-associated ubiquitin protease that removes ubiquitin from histone H2B and allows recruitment of Ctk1/Ctdk-I (Wyce *et al*., 2007). Upon its recruitment, Ctk1/Ctdk-I phosphorylates the Ser-2 of the CTD domain initiating transcript elongation (Wood *et al*., 2005).

Plant homologs of CDK9 were identified as members of the CDKC family (Joubes *et al*. 2000), whereas the plant homolog of mammalian CyclinT interacts with *Arabidopsis* CDKC1/CDKC2 proteins in yeast two-hybrid and immunoprecipitation assays (Barroco *et al*., 2003). CDKC/CyclinT complexes phosphorylate the CTD domain of PolII in *Arabidopsis* and Medicago (Cui *et al*., 2007; Fulop *et al*., 2005), suggesting a role in transcriptional regulation similar to that of the mammalian CDK9/CyclinT complex. In agreement with this hypothesis, *cdkc2*/CDKC1:RNAi or *CyclinT1;4/CyclinT1;5* double mutants are extremely resistant to the Cauliflower Mosaic Virus (CaMV) indicating that the CDKC2/CyclinT complex is indispensable for the transcriptional activation of CaMV viral genes (Cui *et al*., 2007). In addition to viral resistance, the double mutants exhibited developmental defects such as altered leaf growth and trichome architecture as well as delayed flowering, suggesting a role of the plant P-TEFb in other aspects of plant development.

1.7 Other CTD-kinases

1.7.1 CDK10-like proteins

The mammalian CDK10 and CDK11 proteins have been implicated in PolII-dependent transcription. Even though CDK10 contains a variant of the PSTAIRE motif, no cognate cyclins have been found. CDK10 associates with the N-terminal domain of the Ets2 transcriptional activator and inhibits its action (Kasten and Giordano, 2001). Ets2 inhibition was not dependent on the activity of the kinase, as both wild type and an inactive kinase mutant of CDK10 impose the same level of inhibition. CDK10 may affect the activity of Ets2 by interfering with the binding of Ets2 to members of the transcription apparatus (Kasten and Giordano, 2001). Ets family members regulate the expression of genes involved in tumour development and cancer, functioning either as trans-activators or suppressors (Graves, *et al.*, 1998; Graves and Petersen, 1998; Buggy *et al*., 2006). CDK10-dependent regulation of Ets2 trans-activation suggests that this kinase could affect signalling pathways involved in cancer development. Indeed, CDK10 is an important determinant of resistance to endocrine therapy for breast cancer patients (Iorns *et al*., 2008), since RNAi silencing of CDK10 resulted in elevated resistance of breast cancer cells to tamoxifane, an anti-cancer agent that targets estrogen signalling.

1.7.2 CDK11-like proteins

CDK11 proteins contain the PITSLRE variant motif and associate with L-type cyclins (Dickinson *et al*., 2002). In humans, three isoforms of CDK11 proteins have been described; CDK11^{p46}, CDK11^{p58} and CDK11^{p110}, where the numbers correspond to the size of respective polypeptides in kilodaltons (Bunnell *et al*., 1990; Xiang *et al*., 1994; Gururajan *et al*., 1998). The first two isoforms were found to be involved into apoptosis and mitosis, respectively, whereas the third one has been implicated in pre-mRNA transcription and splicing (Lahti *et al*., 1995; Bunnell *et al*., 1990; Trembley *et al*., 2003). In this section, I discuss only the CDK11^{p110} isoform and refer to it hereafter as CDK11. CDK11 co-immunoprecipitates with transcription elongation factors and PolII but casein kinase 2 (CK2), and not CDK11, is the kinase responsible for phosphorylating the PolII, whereas CDK11 is also a substrate of CK2 itself (Trembley

et al., 2003). These data suggest that CDK11 participates in complexes that include CK2 and that its activation might be the important step in co-ordinating pre-mRNA transcription and processing. CDK11 was found to be part of the Mediator complex in humans and that its kinase activity inhibited viral activator VP16-dependent transcriptional activity (Tsutsui *et al*., 2008). In the same report, CDK8 knockout had the opposite effect on transcriptional activation suggesting that, even though members of the same complex, CDK11 and CDK8 play distinct roles in transcriptional regulation.

Comparative genomic analysis of the evolution of CTD-kinases in animals, plants, yeast and protists, identified three plant proteins that were grouped into the CDK11/CDK10-like phylogenetic clade; two from *Arabidopsis* and one from rice (Guo and Stiller, 2004). The *Arabidopsis* genes correspond to CDK-like protein kinases CDKG;1 and CDKG;2 (Menges *et al*., 2005) whereas the rice gene corresponds to Os04g0488000 locus. Even though they were grouped into the CDK10/CDK11 clade, all three proteins were more closely related to CDK11. CDKG1 mRNA expression is up-regulated during cell cycle re-entry whereas CDKG2 mRNA levels remain stable all the way through the cell cycle (Menges *et al*., 2005). Cell cycle exit or re-entry, that is the transition from G1 to G0 stage and *vice versa*, is controlled by the suppression/transactivation of E2F-type transcriptional activators that recruit chromatin modification factors to gene promoters (Takahashi *et al*., 2000; Trimarchi and Lees, 2002). That CDKG1 is up-regulated at the G0/G1 transition (Menges *et al*., 2005) and is also related to the mammalian CDK11, suggests that CDKG1 probably has an important role in resumption of transcriptional activity during the cell cycle re-entry. Moreover, GFP protein localization data showed that CDKG1 fusion protein localizes into nuclear speckles suggesting involvement into transcription and splicing (Kitsios, 2007).

1.8 CDKs and pre-mRNA processing

Production of a mature mRNA that is ready to be exported to the cytoplasm and translated by the ribosome, involves three major processing steps: capping, splicing and polyadenylation. Transcription and pre-mRNA processing are concomitant events that take place as PolII progresses through the coding region of the gene. Most interestingly,

the link between pre-mRNA transcription and processing is mediated, at least partially, by CDK-dependent phosphorylation events in yeast and mammals (Guiguen *et al*., 2007; Chen *et al*., 2007; Loyer *et al*., 2005; Pei *et al*., 2003, 2006; Ko *et al*., 2001).

1.8.1 P-TEFb and pre-mRNA processing

As discussed in Section 1.3.5, PolII-dependent transcript elongation is regulated by the P-TEFb transcription factor in eukaryotes. Recruitment of yeast P-TEFb onto the PolII holoenzyme is dependent on the presence of mRNA capping factors cap methyltransferase Pcm1 and RNA triphosphate Pct1 (Guiguen *et al*., 2007; Pei *et al*., 2003, 2006), suggesting a link between pre-mRNA elongation and processing. Loss of Pcm1 from yeast cells causes dephosphorylation of CTD at Ser-2 position whereas chromatin immunoprecipitations from depleted cells show a dramatic decrease of chromatin-bound P-TEFb complex (Guiguen *et al*., 2007). Also, the CDK9/Pch1 complex in fission yeast, the homolog of mammalian P-TEFb, interacts with the RNA triphosphate Pct1 and phosphorylates the elongation factor Spt5 (Pei *et al*., 2003). Pct1 is necessary for the formation of the cap complex whereas the Spt5 factor belongs to the DSB-sensitivity inducing factor (DSIF) complex. Upon phosphorylation of CTD Ser-5 by TFIIH, DSIF complex binds Pcm1 and Pct1 in fission yeast and suppresses transcript elongation (Pei and Shuman, 2002).

In conclusion, transcript initiation and pre-mRNA processing are co-ordinated (see Figure 1.11). Upon phosphorylation of CTD Ser-5 by CDK7/CycH CTD tail is released from the DNA and PolII leaves the promoter, pausing soon after due to association with the DSIF and NELF inhibitor proteins (Wada *et al*., 1998). This temporary pause allows the loading of the capping complex to the 5'-end of the nascent pre-mRNA. The cap complex recruits the P-TEFb factor by interaction with Pcm1 and Pct1 capping proteins and then the catalytic subunit of P-TEFb phosphorylates the Spt5 protein component of the DSIF complex, the RD subunit of NELF (Fujinaga *et al*., 2004) and the Ser-2 of CTD. These phosphorylation events lead to relief of PolII by inhibitory factors and establishment of transcript elongation.

1.8.2 CDC2-like proteins and splicing

A family of mammalian cdc2-like proteins associate with spliceosomal components, suggesting a CDK role in pre-mRNA splicing. These CDKs share the PITA(I/V)RE cyclin binding motif and contain a serine/arginine-rich (RS) motif at their N-terminus, which is a typical feature of kinases that phosphorylate SR-splicing factors (Marques *et al*., 2000). In humans, the cdc2-related protein CrkRS/CRK7 might be a link between splicing and transcription as it co-localised with the splicing factor SC35 and phosphorylated *in vitro* the CTD domain of PolII (Ko *et al*., 2001). The rat homologue of CrkRS was renamed CDK12 and both its isoforms, CDK12(L) and CDK12(S), interact with CyclinL1 and CyclinL2 and are involved in alternative splicing (Chen *et al*., 2006). L-type cyclins also interact with the mammalian CDC2L5, renamed to CDK13, and regulate alternative splicing as CDK12 (Chen *et al*., 2007). Even though regulation of splicing by CDK12 and CDK13 involves the ASF/SF2 splicing factor, the mechanism of action seems to differ between the two complexes. The effect of CDK12/CyclinL complex on splicing was counteracted by over-expressing the splicing factor, and this alteration did not involve phosphorylation events. This suggests that CDK12/CyclinL complexes affect splicing by ASF/SF2 sequestration, leading to inhibition of its splicing activity (Chen *et al*., 2006).

Conversely, CDK13 directly phosphorylates the ASF/SF2 splicing factor and positively regulates alternative splicing of viral mRNAs (Berro *et al*., 2008). Inhibition of CDK13 activity was achieved by interaction of the kinase complex with the splicing inhibitor protein p32, leading to sequestration of CDK13 away from ASF/SF2 substrate (Berro *et al*., 2008; Even *et al*., 2006; Petersen-Mahrt *et al*., 1999).

Even though inhibition of ASF/SF2 phosphorylation eventually blocks splicing, ASF/SF2 dephosphorylation is a necessary step for the first trans-esterification event in the splicing process *in vitro* (Cao *et al*., 1997). Therefore, CDK13-mediated ASF/SF2 phosphorylation is the first step in initiating splicing, followed by a step of ASF/SF2 dephosphorylation. However, inhibition of CDK13 activity by binding to p32, does not allow the phosphorylation/dephosphorylation cycle to resume leading eventually to splicing inhibition.

Taken together, CDK12 and CDK13 complexes seem to be a part of an ASF/SF2-dependent network of splicing regulation but act in different ways. CDK12 inhibits

splicing by sequestering splicing factors away from their interacting proteins whereas CDK13 promotes it by directly phosphorylating components of the splicing machinery.

1.8.3 CDK11 protein family and pre-mRNA processing

CDK11 (PITSLRE proteins) may affect the activity of spliceosomal components, in addition to its role in transcription regulation (Trembley *et al*., 2004). From the three CDK11 isoforms produced, $CDK11^{p110}$, $CDK11^{p58}$ and $CDK11^{p46}$, only the p110 one has been related to splicing regulation. The other two isoforms, p46 and p58, are involved in the regulation of apoptosis and mitosis, respectively (Lahti *et al*., 1995; Bunnell *et al*., 1990).

The first indication of $CDK11^{p110}$ isoform (hereafter, called CDK11) involvement in splicing was the finding that the protein co-localized with the splicing component SC35 in nuclear speckles in human cells (Loyer *et al*., 1998). A yeast two-hybrid assay identified novel splicing interactors of CDK11 among which was the splicing factor 9G8, which is an *in vivo* substrate of CDK11, whereas splicing assays with CDK11-immunodepleted protein extracts confirmed the importance of CDK11 in the splicing process (Hu *et al*., 2003). In Section 1.4.2, I discussed the involvement of CDK11 in transcriptional activation. The above-presented CDK11-dependent regulation of pre-mRNA splicing suggests that CDK11, along with CrkRS, might be a link between transcription and splicing. The identification of the RNPS1, an RNA-binding protein with an RS-rich motif that is characteristic of splicing factors, as an interactor of CDK11 supports this idea (Mayeda *et al*., 1999). When RNPS1 was over-expressed in human cells it altered the normal distribution of nuclear speckles causing aggregation of most nuclear speckles into 5-6 “large speckles” (Loyer *et al*., 1998), mimicking the one generated when cells are treated with transcription inhibitors.

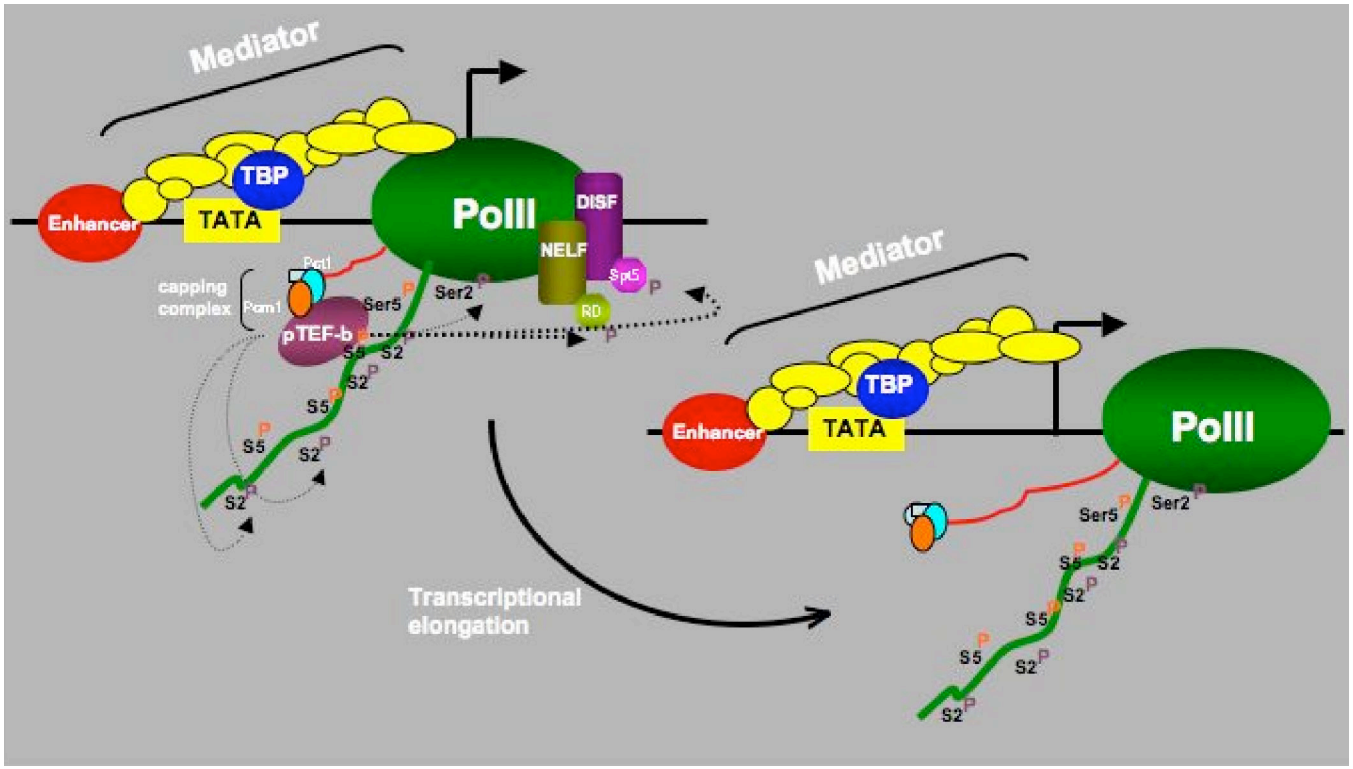


Figure 1.11 Coupling of transcription to pre-mRNA processing.

P-TEFb complex is recruited to PolII holoenzyme by the combinatory effect of the capping complex (PCT1/Pcm1 in yeast) and the phosphorylated CTD tail at Ser-5. Upon binding, the catalytic subunit of P-TEFb, CDK9/CycT, phosphorylates Ser-2 of the CTD and inhibitor protein complexes DSIF and NELF, releasing PolII and initiating transcript elongation.

Since RNPS1 is also a component of the EJC, responsible for mRNA export and nonsense-mediated decay of aberrant mRNAs (Le Hir *et al*., 2000), it can be suggested that CDK11 kinase constitutes a link between splicing, transcription and mRNA export in mammals.

1.8.4 Plant CDKs and mRNA processing

Phylogenetic analysis of CTD-kinases grouped plant CDKC proteins (*At*CDKC1 and *At*CDKC2, *Os*CDKC1) into a sub-group of the CDK9-like protein clade, containing the human CrkRS and CDK12/CDC2L5 kinases (Guo and Stiller, 2004). Positioning of the CDKC family within the CDK9-like group is in agreement with the suggestion that CDKC/CyclinT complexes in *Arabidopsis* are the mammalian homolog of P-TEFb (Barroco *et al*., 2003; Fulop *et al*., 2005). However, a more detailed phylogenetic comparison with known regulators of splicing activity in eukaryotes suggested that CDKC2 might be involved in splicing regulation as well; and therefore be functional homologs of the mammalian CrkRS as the link between transcription and splicing. Indeed, experimental data generated support this idea; *Arabidopsis* CDKC2 was found to co-localize with SR-splicing factors and the co-localization persisted when the transcriptional status of the cell was altered (Kitsios *et al*., 2008). Moreover, abolishment of CDKC kinase activity altered the speckled localization of both the kinase and the splicing factors to "mega" speckles, resembling localization profiles seen after drug-induced inhibition of transcription (Kitsios *et al*., 2008). This suggests that, at least in *Arabidopsis*, CDKC proteins control transcription and splicing.

1.8.5 Summary

Following their discovery as core regulators of the cell cycle in eukaryotes, CDKs have been found to affect many aspects of cellular homeostasis and especially gene expression. A summary of CDK families from human, yeast and *Arabidopsis* and their various functions is shown in Table 1.2. A phylogenetic tree of CDK proteins, based on amino acid sequence similarity across the three different organisms, is showed in Figure 1.12. In the following section, I present the rationale of my project and the aims of my research.

Table 1.2 Eukaryotic CDKs and their role in gene transcription and pre-mRNA processing.

The first column indicates kinase families, members of which affect transcription and/or pre-mRNA processing. The second column contains homologs of budding yeast (red), human (green) and Arabidopsis (blue) of kinase families, the third column mentions the mechanism of action and the forth column the effect that the respective action has on transcript processing, if known.

At: *Arabidopsis thaliana*, Hs: *Homo sapiens*, Sc: *Saccharomyces cerevisiae*, Sp: *Schizosaccharomyces pombe*

"-": no homologs identified, **"(?)":** no mechanism of action or effect determined for identified homologs**, "N/A":** verified functions that are not related to transcription and/or transcript processing

Kinase family	Kinase name	Mechanism of action	Effect
CDK7-like	ScKin28	CTD-Ser5 phosphorylation	transcription initiation
	HsCDK7/CyclinH	CTD-Ser5 phosphorylation	transcription initiation
	AtCDKD2, AtCDKD3	CTD phosphorylation	(?)
CDK8-like	SpSrb10	CTD-Ser5 phosphorylation	transcription inhibition
	HsCDK8	protein interaction	transcription inhibition
	AtCDKE	(?)	transcription inhibition?
CDK9-like	ScBur1, ScCtk1	CTD-Ser2 phosphorylation	transcript elongation, polysome association (Ctk1)
	HsCDK9	CTD-Ser2 phosphorylation	transcript elongation
CDK10-like	SpPpk23	(?)	(?)
	HsCDK10	protein interaction	transcription inhibition
	AtCDKG1, AtCDKG2	(?)	(?)
CDK11-like	-	-	-
	HsCDK11 (p110) HsCDK11 (p46)	protein interaction eIF3p47 phosphorylation	transcription inhibition, splicing translation inhibition
	AtCDKG1, AtCDKG2	(?)	(?)

Kinase family	Kinase name	Mechanism of action	Effect
CDK12-like	-	-	-
	HsCDK12 (CrkRS)	protein interaction	alternative splicing
	AtCDKC2	(?)	putative splicing regulator viral transcription inhibition
CDK13-like	-	-	-
	HsCDK13	ASF/SF2 phosphorylation	splicing activator
	-	-	-
CDK5-like	ScPho85	N/A	N/A
	HsCDK5	eIF4E phosphorylation	translation activator
	-		
CDC2-like	ScCDC28	N/A	N/A
	HsCDC2	S6K1 phosphorylation	translation activation (G1/S) or translation inhibition (G2/M)
		eIF4E-BP phosphorylation	translation activation
	AtCDKA1	eIF4A phosphorylation	(?)

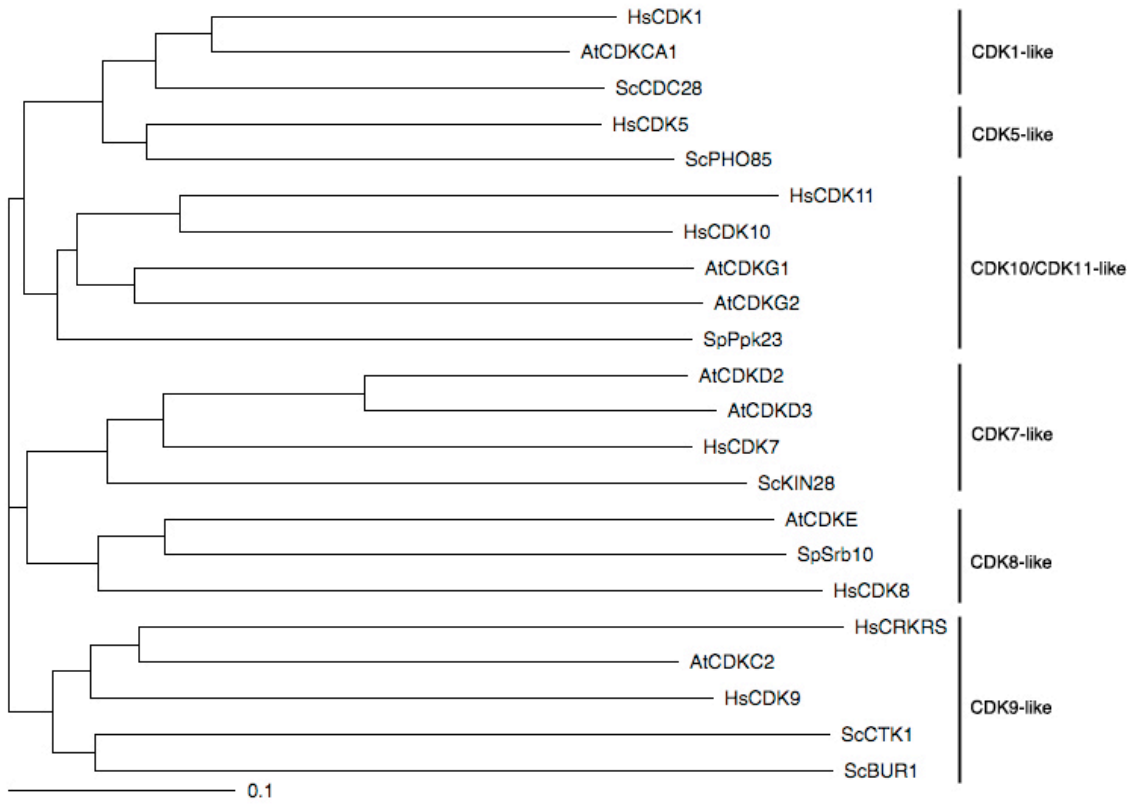


Figure 1.12 Phylogenetic tree of CDK proteins involved in different aspects of pre-mRNA processing

The multiple alignments of proteins were generated using the ClustalX program (www.clustal.org) and the phylogenetic tree was constructed using the TreeView X software (http://darwin.zoology.gla.ac.uk/~rpage/treeviewx/index.html).

1.9 Aims of the project

The mammalian nucleus is well known to be a highly dynamic organelle both structurally and in terms of protein composition, but few similar studies have been undertaken on plant nuclei. A major aim of my PhD research was to characterise the *Arabidopsis* nuclear proteome using subcellular fractionation, followed by high performance liquid chromatography (HPLC) coupled to mass spectrometry (MS)-based protein identification. Comprehensive identification of nuclear proteins would provide the necessary framework to study proteome dynamics. As a first application of the newly obtained Arabidopsis nuclear proteome for studying organellar dynamics, I used MS-based methods to follow changes of the nuclear proteome as *Arabidopsis* cell cultures progress from proliferation to quiescence.

One major mechanism of controlling nuclear processes at the molecular level is post-translational modifications of proteins, which do not necessarily change protein abundance. One type of modification is protein phosphorylation. Therefore, my second aim was to study the role of nuclear-localised protein kinases, specifically CDKC2, in different nuclear processes, using live cell imaging and biochemical techniques to follow protein dynamics and identify possible substrates, respectively.

Overall, my ambition is that my PhD research will provide a protein framework of the Arabidopsis nucleus that will allow future research on the specific roles of identified proteins in a wide range of nuclear processes and also obtain, for the first time, quantitative information on nuclear proteome behaviour between proliferation and quiescence. Moreover, shedding light on the specific roles of a nuclear-localized protein kinase will give us an insight on how this protein can affect certain nuclear processes in *Arabidopsis*.

Chapter 2 - Material and Methods

2.1 Biological material

2.1.1 Cell culture material

Arabidopsis cell lines used were: *Landsberg erecta* cell line MM2d (May and Leaver, 1993) and Columbia-0 (Col-0; Mathur and Koncz, 1998). The tobacco Bright Yellow-2 cell line (BY-2; Kato *et al*., 1972) was also used.

2.1.2 Plant material

The following T-DNA insertion and overexpression lines were used. One T-DNA line was SALK_029546 that contains an insertion in the 8th exon of the *CDKC2* gene (At5g64960) and thereafter called *cdkc2;2*. The second one was SAIL_269_C02 that contains a T-DNA insertion into the first intron of *Magoh* gene (At1g02140). Seeds from both lines were purchased from ABRC (Colombus, USA). I also used seeds from a line overexpressing a GFP fusion of Magoh protein in a Col-0 background, a gift from Ali Pendle (John Innes Centre, UK). The *cdkc2;*2/CDKC1;RNAi line was a kind gift from Dr. Zhixiang Chen (Purdue University, Indiana, USA).

2.2 Growth conditions of plants

2.2.1 Growth of plants in the greenhouse

Seeds were germinated directly on soil. A small number of seeds were spread over Arabidopsis compost and an 8 cm x 8 cm pot was transferred to a 4^{o}C chamber for seed stratification. After 3-4 days, the pot was transferred into the greenhouse and covered with cling film to retain the soil humid until germinated seedlings are adapted to greenhouse conditions. After seedling emergence, the required number of seedlings were pricked out and transferred into multi-cell trays (40 cm x 25 cm) for further growth.

2.2.2 Growth of plants in growth rooms

When growing plants aseptically in growth rooms, seeds were surface-sterilized using 5% (v/v) sodium hypochlorite (Parozone, UK) for 5 minutes. After washing thoroughly

with sterilized water, seeds were placed on Murashige and Skoog (MS) medium containing 0.8% agar, using a pipette. Petri dishes were sealed with micropore tape (3M Healthcare, Germany), covered in foil and transferred at 4^{o}C for 2 days to synchronize germination. Petri dishes were transferred in a growth room at 20^{o}C with 16h light.

2.3 Growth and maintenance of plant cell cultures

2.3.1 *Arabidopsis thaliana* cell cultures

Landsberg erecta cell cultures were grown on a shaker at 150 rpm under constant light at 22^{o}C. Sub-culturing was done by transferring 7 ml of 7-days-old culture into 100 ml of fresh AT media (4.42 gr MS+GB5 (Duchefa), 30 gr sucrose in 1 litre of distilled water and pH adjusted to 5.7 using 1M KOH). *Columbia-0* (Col-0) cell cultures (Mathur *et al*., 1998) were grown on a shaker at 120 rpm in darkness at 25^{o}C. 15 ml of a 7-days-old culture was added into 35 ml of ATN media [4.3 g MS (Duchefa), 30 g sucrose, 100mg myo-inositol, 1mg nicotinic acid, 1mg pyridoxine, 10mg thiamine supplemented with 1 mg 2,4-Dichlorophenoxyacetic acid (2,4-D)]. The pH of the medium was first adjusted to 5.7 with 1 M KOH and 2,4-D was added after autoclaving the medium.

2.3.2 *Nicotiana tabacum* cell culture

BY-2 cell cultures were grown on a shaker at 150 rpm in darkness at 22^{o}C. Subculturing was by transferring 5 ml of a 7-days-old culture into 100 ml of fresh BY2 media (4.3 g MS salts, 20 µl 1 mg/ml 2,4-D and 3.4 µl of 100 mg/ml KH_2PO_4. The pH was adjusted to 5.8 with 1M KOH and the volume made up to 1 litre with water).

2.4 Molecular biology work

2.4.1 Bacterial cultures

The *E. coli* strains used were DH5a (Hanahan, 1983) and BL21 (Stratagene), both from Invitrogen. *E. coli* cultures were grown on a shaker (200 rpm) at 37^{o}C in LB media (1% (w/v) bacto-tryprone, 0.5% (w/v) bacto-yeast extract, 1% (w/v) NaCl, adjusted to pH 7.0)) supplemented with the required antibiotics.

The *Agrobacterium tumefaciens* strain used was GV3101 (Koncz and Schell, 1986). The GV3101 strain is resistant to rifampicin (20μg/ ml) and gentamycin (25 μg/ ml) and growth of cell cultures was at 28^{o}C at 200 rpm for 24h. When growing Agrobacterium cells on plates, the incubation time was 48 hours.

2.4.1.1 Preparation of chemically competent *E. coli* cells (Inoue *et al*., 1990)

A starter culture was prepared by inoculating a single colony into 5 ml of LB plus appropriate antibiotics, depending on the antibiotic resistance of the plasmid, and grown at 37^{o}C overnight. All 5 ml of the overnight culture was added to 500 ml of LB plus antibiotics and grown at 37^{o}C until the absorbance of the culture at 600nm (OD_{600}) was 0.6. Then cell suspension was incubated on ice for 10 min, spun at 2,500xg for 10min at 4^{o}C in a SS-34 rotor centrifuge. The supernatant was discarded and pellet was gently resuspended, using a plastic Pasteur pipette, into 80 ml of ice-cold transformation buffer (10 mM PIPES, 15 mM $CaCl_2$ and 250 mM KCl, adjusted to pH6.7 with KOH and then $MnCl_2$ added to a final concentration of 55 mM). This suspension was spun as above and finally resuspended using 2 ml of ice-cold transformation buffer. Dimethyl sulfoxide (DMSO) was added to a final concentration of 7% (v/v), mixed and incubated on ice for 10min. The mixture was aliquoted into 50, 100 or 200 ml in 1.5 ml eppendorf tubes, flash frozen in liquid nitrogen and stored at -80^{o}C until needed.

2.4.1.2 Transformation of *E. coli* competent cells (Merrick *et al*., 1987)

Chemically competent *E. coli* cells (-80^{o}C) were thawed on ice for 5-10min. 50 μl of cells and 1 ml of a plasmid or 2 - 5 μl of a ligation reaction were used for a single transformation reaction. Both components were added into an ice-cold eppendorf and the tube was gently flicked to mix and then incubated on ice for 30min, heat-shocked by incubating for 45sec at exactly 42^{o}C and returned to ice for 10min. Then 900 μl of pre-warmed 37^{o}C SOC medium (2% (w/v) bacto-tryptone, 0.5%, (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) were added and the tube was shaken at 150 rpm at 37^{o}C for 1h to allow for induction of plasmid's antibotic resistance gene. After 1h, the tube was spun at 16,200xg and the bacterial

pellet was resuspended into 100 µl of SOC medium and pipetted onto the centre of a Petri dish containing LB medium supplemented with the appropriate antibiotic. Cells were spread on the dish using a sterile glass spreader. Then dishes were grown overnight at 37^{o}C. Individual colonies (2-3) were picked with a sterile pipette tip and tested by PCR. For the positive ones, a 5 ml liquid culture overnight was grown and glycerol stocks prepared (Section 2.4.1.5).

2.4.1.3 Preparation of *Agrobacterium tumefaciens* competent cells

A small amount of frozen cells, scraped from a frozen glycerol stock, was streaked on an LB agar plate (supplemented with 20mg/ ml rifampicin and 25mg/ ml gentamycin) and grown at 28^{o}C for 24 hours. An isolated colony was transferred into 5 ml of LB medium (supplemented by 20µg/ ml rifampicin and 25µg/ ml gentamycin) and grown for 24h at 28 ^{o}C shaking at 150 rpm. From this culture, 4 ml was inoculated into 500 ml of LB medium (containing antibiotics) and allowed to grow for approximately 6h at 28^{o}C, until the O.D. of the culture was 0.6. The culture was placed on ice for 10min and spun at 3,000xg for 5min at 4^{o}C. Pellet was resuspended in 10 ml of 20 mM $CaCl_2$, aliquoted (200 µl or 500 µl) into chilled eppendorf tubes, flash frozen and stored at -80^{o}C.

2.4.1.4 Transformation of GV3101 cells – Freeze-thaw method (An *et al*., 1988)

An aliquot of GV3101 competent cells (from -80^{o}C freezer) was allowed to thaw on ice. (To 100 ml of thawed suspension, 0.5-1µg of plasmid DNA was added and briefly flash frozen in liquid nitrogen, then thawed at 37^{o}C for 5 min. Then, 1 ml of LB medium was added without antibiotics and allowed to grow for 2.5-3 hours at 28^{o}C with 150 rpm shaking. After incubation, the suspension was spun on a bench centrifuge at 16,100xg for 30 sec, and the pellet gently resuspended in 100 ml of LB medium without antibiotics. Using a sterile glass rod, the suspension was spread onto a LB agar plate containing rifampicin and gentamycin (*Agrobacterium* selection) and an appropriate antibiotic selecting for the plasmid and grown at 28^{o}C for 48 hours. After confirmation of positive colonies by PCR, a liquid culture was grown and glycerol stocks prepared and stored as described in Section 2.4.1.5

2.4.1.5 Preparation of glycerol stocks

The same method was used for preparing glycerol stocks for *E. coli* or *A. tumefaciens* cells. Six hundred microlitres of an overnight bacterial culture was added into 200 µl of 80%(v/v) glycerol in 1.5 ml eppendorf tube. The mixture was vortexed vigorously and frozen in liquid nitrogen.

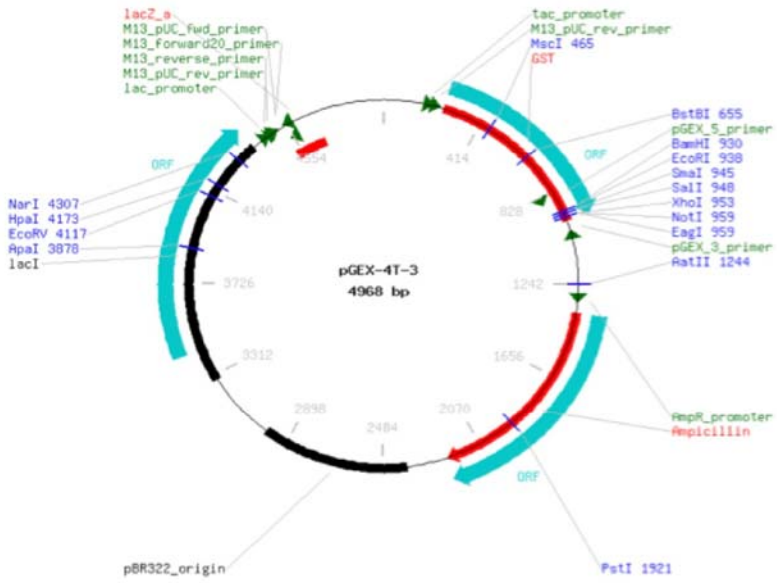
2.4.2 Gene cloning

2.4.2.1 Conventional cloning

Gene cloning using restriction enzymes was used for introducing the cDNA sequence of *CDKC2*, *CDKC2*_D182N, *CDKC2*_F118G, *Magoh*, *Magoh*_T57A, and *Magoh*_T57E into the pGEX4T-3 vector (GE Healthcare; Figure 2.1) to produce N-terminal Glutathione-S-transferase protein fusions. Firstly, primers were designed for amplifying each cDNA sequence and, at the same time, incorporating restriction sites at the 5' and 3' end of the sequence (See Table 2.1). PCR products were purified using the PCR purification kit (Qiagen), following manufacturer's instructions, and I used a 2 µl aliquot to perform a TA cloning to a pGEM-T easy vector (Invitrogen). The principle of TA cloning is based on the T4 DNA ligase-mediated ligation between a PCR product with adeninine (A) overhangs at its 5' ends and a linearized pGEM-T easy vector with thymidine (T) overhangs at its 3' ends. The resulting ligation reaction was transformed into DH5a *E. coli* (Section 2.4.1.2). The transformation reaction was spread onto LB plates containing 100 µg/ml ampicillin, 0.5 mM IPTG (Melford Laboratories) and 80µg/ml X-gal (Melford Laboratories) and grown overnight at 37^{o}C. Plasmids were isolated from white colonies and an aliquot was digested with restriction enzymes to release the cDNA insert. The digest was analysed with horizontal gel electrophoresis. One plasmid was selected on the basis of insert size, and the rest of the aliquot was digested again with the same restriction enzymes to release larger amounts of the cDNA insert for transfer to the pGEX4T–3 vector.

Primer name	Sequence (5' -> 3')	Gene	AGI
BamHI-CDKC2-for	GC**GGATCC**ATGGCGGCTGCGGCTT	CDKC;2	AT5G64960
CDKC2-SalI-rev	CG**GTCGAC**TTACGGTTGCCATCCATA	CDKC;2	AT5G64960
CYCT1;3-XhoI	TCA**CTCGAG**CAGGCTATGGGA	CYCT1;3	AT1G27630
CYCT1;3-NcoI	TA**CCATGG**CTCTGGGTTCA	CYCT1;3	AT1G27630
Magoh-BamHI-for	GCATC**GGATCC**ATGGCCGCGGAAGAAGC	MAGOH	AT1G27630
Magoh-SalI-rev	ACGT**GTCGAC**CTAGATAGGCTTGATTTTGA	MAGOH	AT1G02140

Table 2.1 Primer names and respective sequences. The primers contain restriction enzyme recognition sites that are highlighted in bold within each primer sequence. Modified products were used for conventional cloning using the pGEM-T-easy vector system (Section 2.4.2.1)



pGEX – for : 5’ – GGG CTG GCA AGC CAC GTT TGG TG – 3’

pGEX – rev: 5’ – CCG GGA GCT GCA TGT GTC AGA GG – 3’

Figure 2.1 Map of the pGEX4T – 3 vector.

(Details of vector features are available at www.gelifesciences.com/pGEX. Below the vector, the sequences of primers used for sequencing and PCR screening are given)

Purified digested products were quantified at 280nm absorbance using the nano-drop spectrophotometer ND-1000 (Labtech International) and 60ng from each was ligated to 20ng of digested pGEX4T-3 vector using the Rapid Ligation kit (Roche) following manufacturer's instructions. Five microlitres of the ligation reaction were transformed into *E. coli*, spread on LB plates containing 100 µg/ml ampicillin and grown overnight at 37^{o}C. Colonies were selected and tested using colony PCR with the pGEX-forward and pGEX-reverse primers (Figure 2.1). I isolated plasmid DNA using the Plasmid Miniprep kit (Qiagen) from positive colonies and then sent them for sequencing (Genome Laboratory, JIC) to confirm the presence of the *CDKC2* cDNA sequence, that the sequence is free of errors and in frame with the N-terminal GST tag.

2.4.2.2 Gateway cloning

Generation of protein fusions for expression in cell cultures and plants employed Gateway Cloning (Invitrogen), which uses a two-step strategy for generating constructs containing the gene of interest fused to the desired tag. A flowchart of the procedure is shown in Figure 2.2. Briefly, the first step involves the generation of a PCR product flanked by site-specific recombination sites, attB1 and attB2. This attB-flanked PCR product is then introduced into the entry vector pDONR207 (Invitrogen; Figure 2.2) that harbours a cassette flanked by site-specific recombination sites, called attP1 and attP2 sites. The recombination reaction is performed by the site-specific recombination enzyme, BP clonase. Part of the reaction (half of the amount in my case) was transformed into competent *E. coli* cells (Section 2.4.1.2) and the resultant "entry" clones are then used for performing the LR reaction. The recipient or "destination" vector is, in most cases, a binary vector whose cassette is flanked by attR1 and attR2 sites. Analogous to the BP reaction, the LR reaction is mediated by the LR clonase. After transformation into DH5a cells (Section 2.4.1.1) clones containing recombinant plasmids are selected using an appropriate antibiotic. Colony PCR using recombinant site-specific primers was used to check for insert and positive plasmids used to transform GV3101 cells (Section 2.4.1.4). Destination vectors used during this project contained the cauliflower mosaic virus (CaMV) 35S promoter and a variety of fluorescent or non-fluorescent tags, leading to either N- of C-terminal fusions. Maps of destination vectors used in this project are shown in Figure 2.3.

Figure 2.2 A flow-chart of the GATEWAY cloning system.

An attB site-flanked gene product is converted into an expression clone, through two successive recombination events mediated by the BP and LR Clonase enzymes. For detailed explanation see text in Section 24.3.2 (Adapted by www.invitrogen.com)

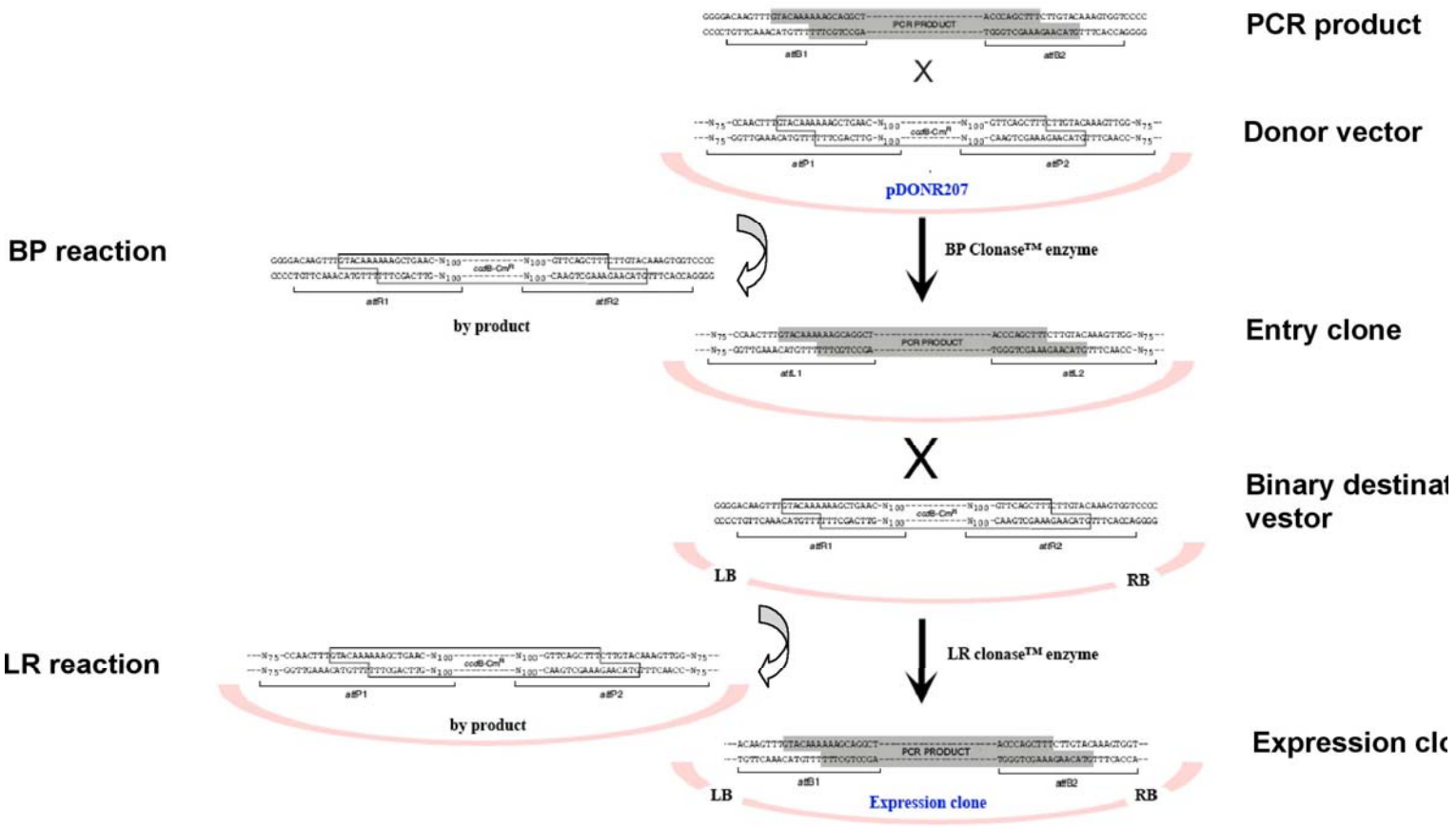

PCR product
Donor vector
pDONR207
BP reaction
BP Clonase™ enzyme
by product
Entry clone
Binary destinat
vestor
LB
RB
LR reaction
LR clonase™ enzyme
by product
Expression clo
LB
Expression clone
RB

Figure 2.3 Maps of Gateway® vector used in cloning experiments.

Vectors shown are donor (A) and binary destination (B-D) vectors. Details about the features of its vector can be found at the respective websites.

(A) pDONR207 was the common vector used for generating entry clones (www.invitrogen.com),

(B) pGFP-N-BIN produces N-terminus GFP protein fusions.

(C) pEarleyGate104 produces an N-terminus YFP protein fusions (http://www.biology.wustl.edu/~pikaardlab).

(D) pH7WGR2.0 produces N-terminus RFP protein fusions (http://gateway.psb.ugent.be/index)

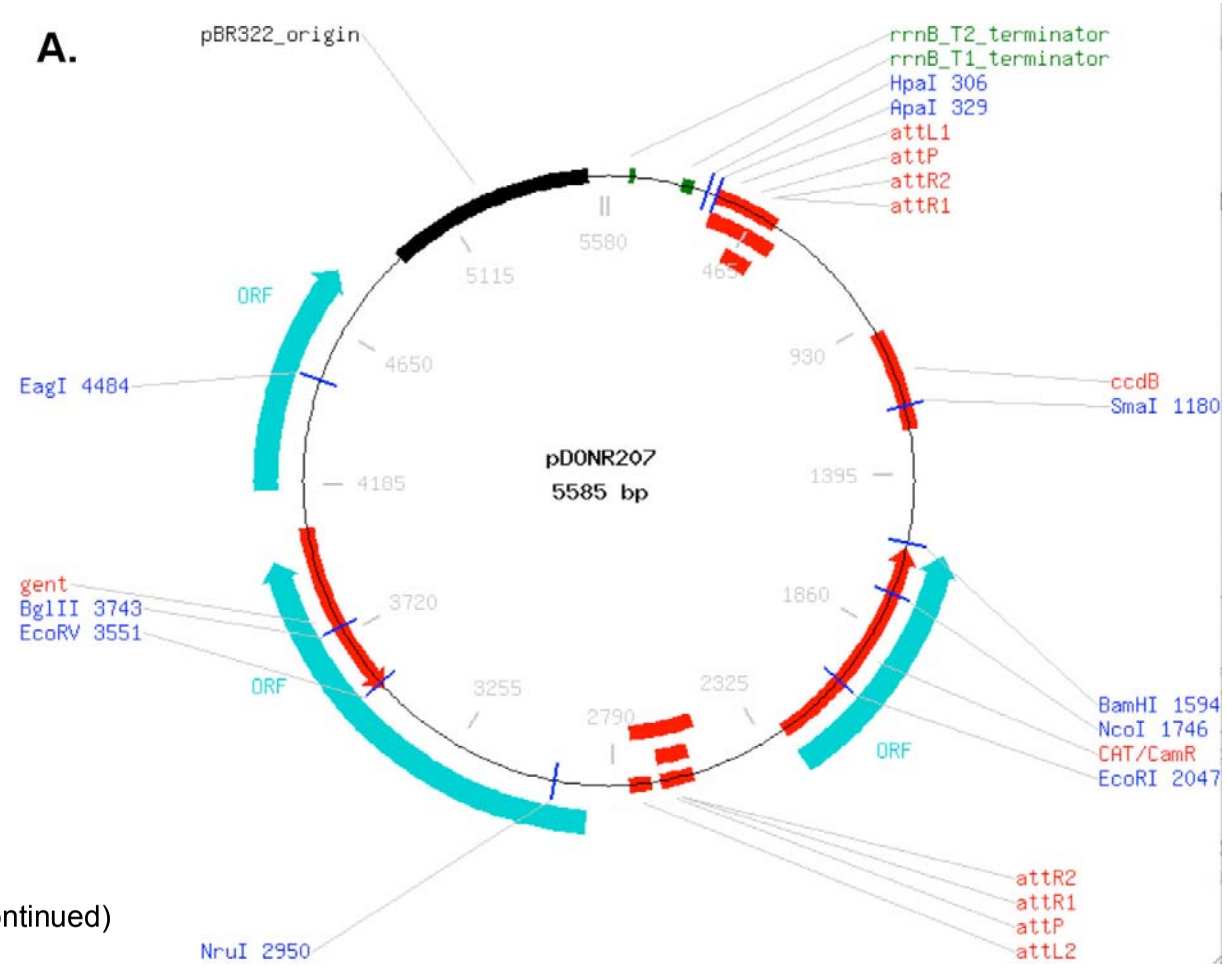


Figure 2.3 (continued)

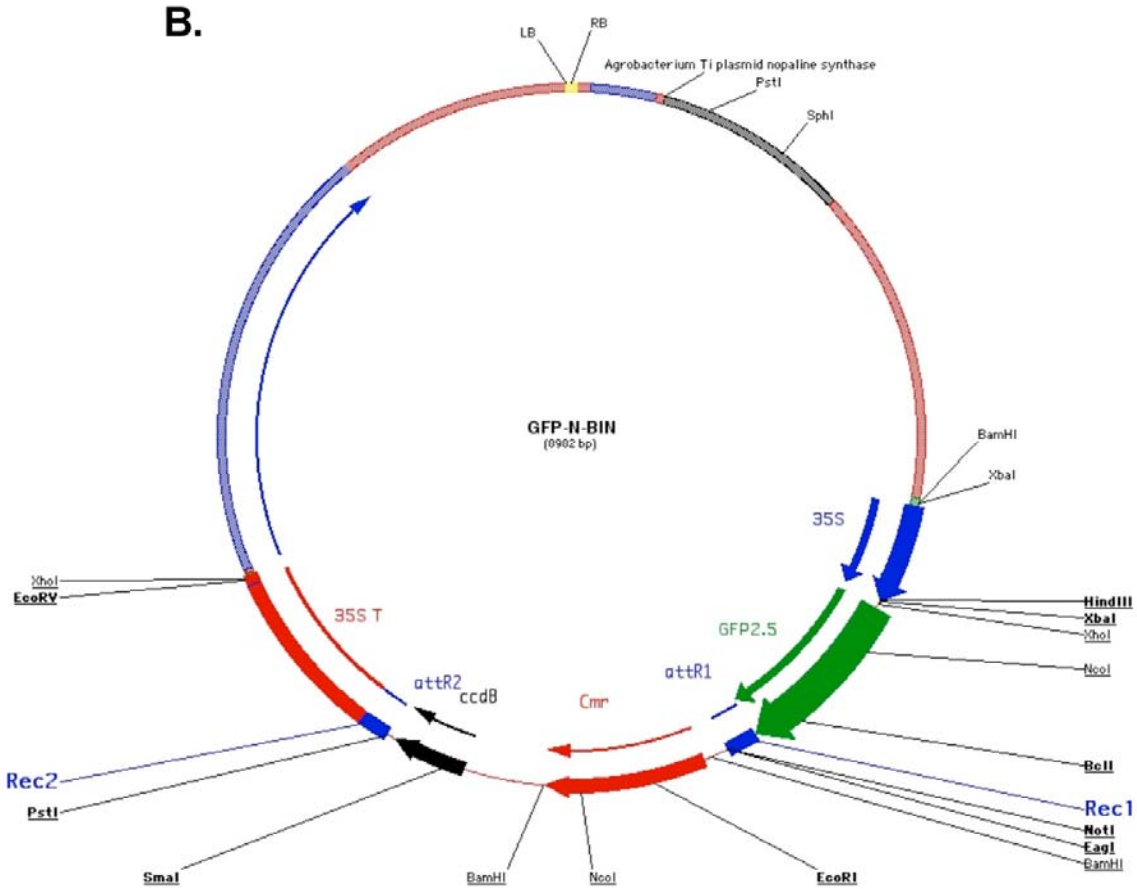

B.
LB
RB
Agrobacterium Ti plasmid nopaline synthase
PstI
SphI
GFP-N-BIN
(0902 bp)
BamHI
XbaI
35S
HindIII
XbaI
XhoI
NcoI
BclI
Rec1
NotI
EagI
BamHI
EcoRI
NcoI
BamHI
SmaI
PstI
Rec2
XhoI
EcoRV
35S T
attR2
ccdB
Cmr
attR1
GFP2.5

Figure 2.3 (continued)

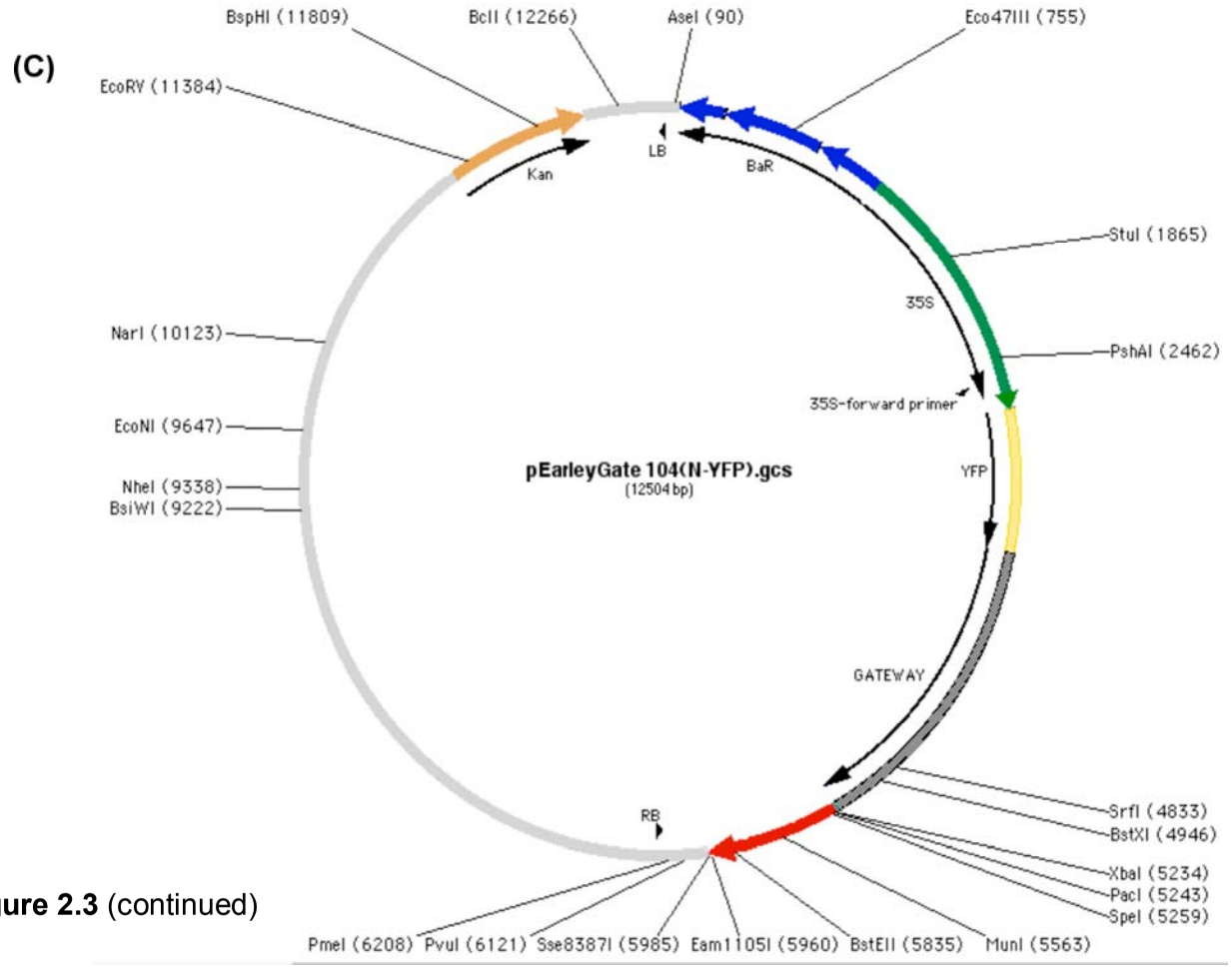


Figure 2.3 (continued)

(D)

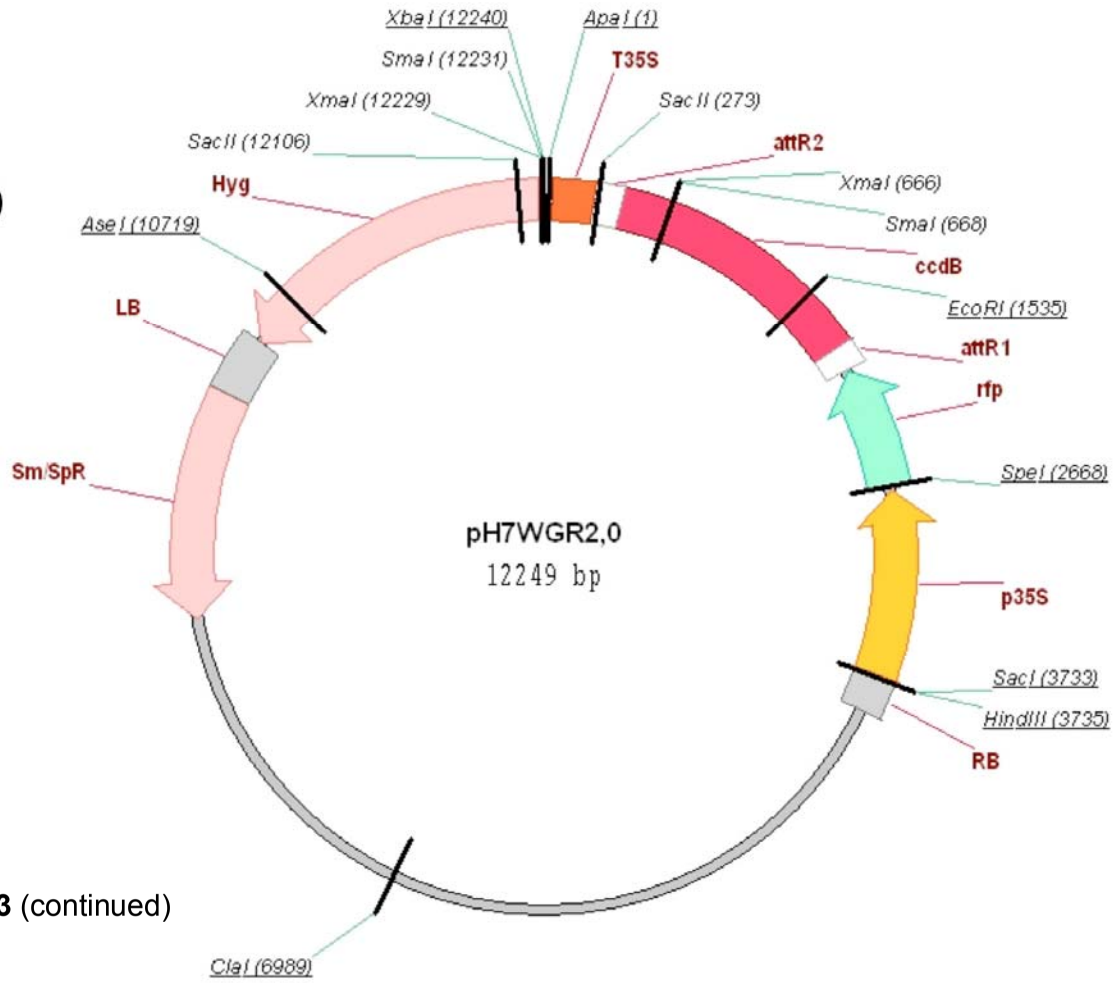


Figure 2.3 (continued)

2.5 Site-directed mutagenesis

Mutations leading to amino acid substitutions on coding sequences were made using the Gene Tailor site-directed mutagenesis kit (Invitrogen). A list of primers used for the mutagenesis reactions is given in Table 2.2.

AGI	Primer name	Sequence (5' -> 3')	Mutation
AT3G48750	CDKA1_D146N-for	CAAACTCACTGAAGCTTGCT**AAT**TTTGGACTGGCC	Asp146 -> Asn146
AT3G48750	CDKA1_D146N-rev	AGCAAGCTTCAGTGAGTTTGTGCGGCGATC	
AT1G76540	CDKB2;1_D164N-for	CAATGAGGCTCAAAATAGCA**AAT**CTTGGTTTAGC	Asp164 -> Asn164
AT1G76540	CDKB2;1_D164N-rev	TGCTATTTTGAGCCTCATTGTCTTGGGATCCA	
AT5G64960	CDKC2_D182N-for	AGGGAAATTTAAAGCTAGCG**AAT**TTTGGGCTTGC	Asp182 -> Asn182
AT5G64960	CDKC2_D182N-rev	CGCTAGCTTTAAATTTCCCTCATTGTCAAT	
AT5G64960	CDKC2_F118G-for	GGTGGAATCTACATGGTC**GGT**GAGTACATGGAT	Phe118 -> Gly118
AT5G64960	CDKC2_F118G-rev	GACCATGTAGATTCCACCCTTGTATTTGTT	
AT1G02140	Magoh_T57A-for	TCCGCAAGGAGGTATTTCTC**GCA**CCCGCCGTGC	Thr57 -> Ala57
AT1G02140	Magoh_T57A-rev	GAGAAATACCTCCTTGCGGATAATAGTATC	
AT1G02140	Magoh_T57E-for	TCCGCAAGGAGGTATTTCTC**GAG**CCCGCCGTGC	Thr57 -> Glu57
AT1G02140	Magoh_T57E-rev	CAGAATTCACATGGCGATTCGGAAGGAGGA	

Table 2.2 Primer names and their respective sequences used for achieving amino acid substitution. The nucleotide triplet corresponding to the new amino acid is indicated in bold in the forward primer.

2.6 Protein overexpression in BL21 *E. coli* cells

For protein overxpression, the BL21(DE3)pLysS strain of *E. coli* was used with the genotype *E. coli* B F– *d cm ompT hsdS*(r_B^- m_B^-) *gal* λ(DE3) (GE Healthcare). This strain is resistant to chloramphenicol and induction of overexpression is achieved with the addition of IPTG into the culture. IPTG activates T7 polymerase promoter that, in turn, drives the expression of the gene of interest. Constructs used to transform the BL21 cells are shown in Table 2.3; including induction conditions and expected protein sizes. Plasmids were transformed into BL21 as described above (Section 2.4.1.2). After

obtaining positive colonies, confirmed by PCR and gene sequencing, I prepared glycerol stocks for future use.

2.6.1 Induction method

For protein expression experiments, liquid cultures were grown in LB liquid medium (supplemented with 100 μg/ml ampicillin and 34 μg/ml chloramphenicol) at 37^{o}C overnight. Two millilitres of overnight culture from each construct was added into 100 ml of LB medium (supplemented with 100 μg/ml ampicillin) and allowed to grow until O.D. was 0.8-1.0. Protein expression was induced by adding the required amount of IPTG for each construct (Table 2.3). Just before adding IPTG, 1 ml of liquid culture was removed to serve as a non-induced control.

2.6.2 Extraction of expressed proteins

I used two methods for isolating the expressed protein. During the pilot experiments, where the optimal conditions for each construct were determined, I boiled bacterial pellets into 1X Sample buffer (1% (w/v) SDS, 10% (w/v) glycerol, 5% (v/v) β-mercaptoethanol, 0.002% (w/v) bromophenol blue, 62.5 mM Tris-HCL, pH 6.6) and monitored protein expression on SDS-PAGE (Section 2.10.3). When optimal induction conditions were determined, I used a protocol for native protein isolation. Cell cultures containing expressed proteins were spun down at 3,000 rpm for 15min and the supernatant discarded. Cell pellets were either used directly for protein extraction or stored at -20^{o}C for future use. For native protein isolation, cell pellets were resuspended completely into ice-cold 1xPBS containing 1 tablet of Complete EDTA-free Protease Inhibitor Cocktail (1 tablet per 50 ml of buffer; Roche) at 1:10 ratio of PBS to initial cell culture. After resuspending pellets, 0.2 mg/ml Lysozyme (Sigma) was added to facilitate bacterial cell lysis. After incubation on ice for 5 minutes, I sonicated protein extracts in SANYO™ Soniprep 150 sonicator on ice for 3x10 sec and 10 sec intervals. This step helps to increase the yield of soluble protein and reduce the viscosity of the solution by shearing the DNA. Sonicated samples were spun using a Sorvall Evolution RC centrifuge, in a SS-34 rotor at 13,000 rpm for 15 minutes at 4^{o}C. The supernatant was transferred into a 15 ml sterile plastic tube (Starlab UK, Ltd) and kept on ice.

Plasmid	Fusion product	Induction			Expected size
		IPTG	time	temperature	
CDKC;2 in pGEX4T-3	N-terminal GST fusion	0.4mM	3h	37°C	82 kDa
CDKC;2 _D182N in pGEX4T-3	N-terminal GST fusion	0.4mM	3h	37°C	82 kDa
CDKC;2 _F118G in pGEX4T-3	N-terminal GST fusion	0.4mM	3h	37°C	82 kDa
CYCLINT1;3 in pET32a*	N-terminal thioredoxin / 6xHis fusion	0.4mM	3h	37°C	57kDa
Magoh in pGEX4T-3	N-terminal GST fusion	0.2mM	4h	37°C	35.6 kDa
Magoh _T57A in pGEX4T-3	N-terminal GST fusion	0.2mM	4h	37°C	35.6 kDa
Magoh _T57E in pGEX4T-3	N-terminal GST fusion	0.2mM	4h	37°C	35.6 kDa

Table 2.3 A summary of plasmids encoding for over-expressed proteins and respective induction conditions used for protein production. Details regarding protein isolation and subsequent purification are presented in the text. * *CYCLINT1;3* in pET32a was a kind gift from Dr. Zhixiang Chen, Purdue University

2.6.3 GST- and His-tagged protein purification

Purification of GST-tagged and His-tagged proteins was performed using Glutathione Sepharose 4B beads (GE Healthcare) and HisPur Cobalt Resin (Pierce), respectively. For the pull downs, I used 200 ml of bead slurry (equal to 150 ml bead volume) per 100 ml of initial cell culture. Beads were washed twice in 1.5 ml of ice-cold 1XPBS, plus protease inhibitors, by spinning on a tabletop centrifuge for 30sec at 13,000 rpm. After the last wash, 1 bead volume of ice-cold 1XPBS was added resulting in 300 □l of bead slurry per 100 ml of initial cell culture. After having equilibrated the beads in PBS, I added the native protein extracts (Section 2.4.5.2) 200 µl of bead slurry per 100 ml of initial cell culture. Mixtures were incubated for 30 min at 4^{o}C on a rocking platform, spun on a RT6000B centrifuge for at 2,500 rpm for 5min at 4^{o}C. The supernatant was discarded and the pellet was resuspended in 15 ml 1XPBS and spun again. This wash was repeated two more times. Finally, bound protein/proteins were eluted by incubating for 10min on ice in 200 ml of GST elution buffer [50 mM Tris (pH8.0), 10 mM reduced glutathione] or His elution buffer (50 mM sodium phosphate, 300 mM sodium chloride, 150 mM imidazole; pH 7.4). Elution was done twice and eluates for each protein were combined, ending up with 400 ml of eluate per expressed protein. Eluates were concentrated using Microcon columns (Milipore) with 10kDa molecular weight cut-off, following manufacturer's instructions. The yield was 5-10 ml of overexpressed protein containing an estimated 100 – 200 mg of protein.

2.7 Nuclei isolation (Adapted from Pendle *et al*., 2005)

A graphical illustration of the procedure is given in Figure 2.3. Two hundred mls of Arabidopsis (*L. erecta*) cells were aliquoted into 50- ml plastic tubes (Starlab, UK) and spun at 800 rpm for 5 min on a RT6000B centrifuge; cell pellets were resuspended into a total of 100 ml of cell wall digestion medium (Table 2.4) and protoplasts were generated by gently shaking for 2-3 hours at 25^{o}C. Cell wall digestion was frequently inspected under an inverted stereo microscope and terminated when the majority of cells were converted to spherical and individual protoplasts, The protoplasts were centrifuged as above and the protoplast pellet was kept on ice. [The digestion medium was removed and stored at -20^{o}C for future use (each batch of medium was reused up to

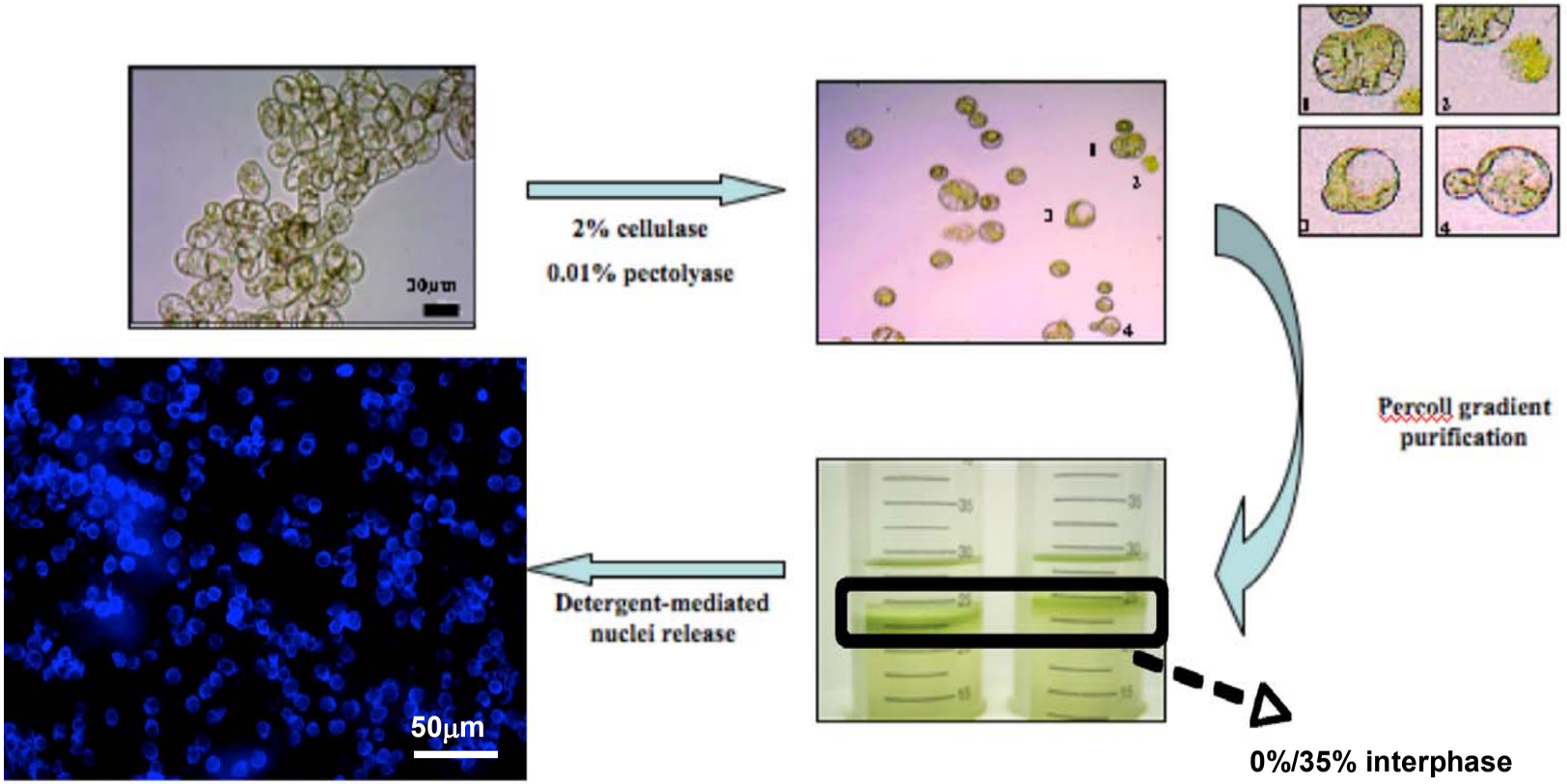
three times)] Percoll gradient density medium (GE Healthcare) was used to manually generate the gradient: protoplasts from each tube were resuspended into 20 ml of floatation buffer (FB; Table 2.4) and gently overlaid with 5 ml of 45% FB, 5 ml of 35% FB and 5 ml of resuspension buffer (RB; Table 2.4). The gradient was spun at 500 rpm for 5min at 4^{o}C and at the end of centrifugation, protoplasts formed a layer at the 0%/35% interface (see Figure 2.4). Protoplasts were removed, using a plastic Pasteur pipette, and transferred to a new 50 ml plastic tube. They went through one round of washing in RB and the resulting pellet was resuspended again in RB. At this stage, protoplasts were kept on ice. To determine their concentration I used a HAE2012 haemocytometer (Scientific Laboratory Ltd.), following manufacturer's instructions. The protoplasts were harvested and resuspended in nuclei isolation buffer (Table 2.4) at a concentration of 1x10^{6} protoplasts per ml of medium.

Buffer	Recipe
Resuspension buffer (RB)	10mM 2-(N-morpholino)ethanesulfonic acid (MES; Sigma), pH5.5 with 1M KOH
	1mM CaCl2
	0.5M Sorbitol
Cell wall digestion medium	2% (w/v) cellulase Onozuka R-10 (Yakult) in RB
	0.01% (w/v) pectolyase Y-23 (Melford Laboratories) in RB
60% Floatation buffer (FB)	10mM MES pH5.5
	1mM CaCl2
	0.44M sorbitol
	60% Percoll (GE Healthcare)
10ml of 45% FB	7.5ml 60% FB
	2.5ml RB
10ml of 35% FB	5.83ml 60% FB
	4.17ml RB
Nuclei isolation buffer (NIB)	10mM MES pH5.5
	0.2M sucrose
	2.5mM EDTA
	2.5mM DTT
	0.1mM spermine (Sigma)
	10mM NaCl
	10mM MgCl2
	0.5mM spermidine (Sigma)
	15mM β-glycerophosphate
	0.5mM sodium orthovanadate (Sigma)
	25mM sodium fluoride (Sigma)
	1mM phenylmethylsulphonyl fluoride (Sigma)
	0.3% Triton X-100 (Sigma)

Table 2.4 Composition of buffers used for the isolation of nuclei from Arabidopsis *L. erecta* cell cultures

Figure 2.4 Schematic representation of the procedure for nuclei isolation in Arabidopsis *L. erecta* cell cultures.

After enzymatic digestion of cell wall, protoplasts are purified in a density gradient and nuclei are subsequently isolated in the presence of detergent. Nuclei integrity is checked using the DNA-specific dye DAPI. A magnification of the 4 protoplasts indicated at the 2nd picture of flowchart, is shown on the top right corner. Number 1 shows fused protoplasts, number 2 shows cell wall debris, number 3 shows a protoplast with local excess in osmotic pressure and number four shows two protoplasts just prior to separation. The protoplast-enriched density interface, resulted after centrifugation of the protoplast suspension, is indicated in the 3rd picture.


30µm
2% cellulase
0.01% pectolyase
Percoll gradient
purification
0%/35% interphase
Detergent-mediated
nuclei release
50µm

The protoplast suspension was incubated on ice for 5-7min and then poured into a stainless steel homogenizer where, by using a close-fitting plastic rod, I homogenized the suspension to rupture protoplasts and release nuclei (this step was further optimised; see Chapter 3). The number of homogenizations depended on the efficiency of nuclei release, which was assessed under an inverted stereomicroscope. After completion of the nuclear release, the suspension was spun at 1000 rpm for 5min at 4^{o}C and the supernatant, containing cytoplasmic proteins, was discarded. The nuclei pellet was resuspended in 5-7 ml of nuclei isolation buffer without any Triton X-100 to reduce remnants of cytoplasmic material. A small amount (50 µl) of nuclei suspension was kept for DAPI staining in order to assess the purity and integrity of nuclear preparation. The rest of the suspension was either aliquoted into 1.5 ml eppendorf tubes, spun down as above, or spun down as a whole. Resulting pellets were flash-frozen in liquid nitrogen and stored at -80^{o}C until needed.

2.8 Genetic transformation of plants and plant cell cultures

2.8.1 Genetic transformation of plants: floral dipping method (Clough and Bent, 1998)

10 ml of LB, supplemented with 20µg/ ml rifampicin, 25 µg/ ml gentamycin and 50 µg/ ml kanamycin, was inoculated with an Agrobacterium colony harbouring constructs containing N-terminal YFP fusions of CDKC2 and CDKC2_D182N. Cultures were shaken at 28 °C for 48 h and 5 ml used to inoculate a new 500 ml of LB culture, supplemented with kanamycin only that was grown for 1 – 18 hours. When the OD_{600} of the culture was 0.8-1.2 cells were spun at 6,000 rpm for 10 min at room temp, using a SLC-1500 rotor in a Sorvall Evolution RC (Thermo Scientific) centrifuge and resuspended in 400 ml dipping solution (5% (w/v) aqueous sucrose in H_2O), to an OD_{600} of 1.0-1.2, 200 µl of Vac-in-Stuff / Silwet L-77 (Lehle Seeds) was added and the mixture poured into a plastic beaker.

Flowering plants were dipped vertically into the solution for 1-3 min, laid horizontically for 2 min on blue towel horizontally and placed in clear plastic autoclave bags (with corners cut off to allow ventilation). After 3 days, the floral dipping was repeated to increase the number of transformed seed. Re-dipped plants were re-sealed for 1 day into

bags and then transferred into the containment glasshouse and allowed to set seeds. Seeds were collected and sowed by the greenhouse personnel on large trays containing F2-based compost. When the two first real leaves appeared, seedlings were sprayed with 0.25% (v/v) Basta (Harvest). Spraying was repeated after one week and, eventually, resistant plants were allowed to set seed.

2.8.2 Transient transformation of Arabidopsis seedlings (Adapted from Marion *et al*., 2008)

Seeds of different genotypes were prepared as described in Section 2.2.2. For the transformation, I used 8-well plates containing MS medium + 0.8% agar and a piece of autoclaved sterile filter (Saatitech, pore size 0.5 mM) sitting on the top of the medium. Sterile seeds were laid on the top of the filter, plates sealed with micropore tape and left at 4^{o}C for 2 days before being transferred to a 20^{o}C growth room with 16h light. Two days later, I set up a liquid culture by growing Agrobacterium in 5 ml LB medium plus appropriate antibiotics at 28^{o}C overnight. One millilitre of the overnight culture was used to inoculate 30 ml of LB medium plus antibiotics early in the morning and allowed to grow until the O.D.$_{600}$ of the culture was 5 (approx. 24h). Cultures were centrifuged at 5000 rpm for 15min; the pellet was resuspended into 2 ml of 5% (w/v) sucrose and O.D.$_{600}$ was determined again. Suspensions were diluted to O.D.$_{600}$ = 2 in 5% sucrose plus 200μM acetosyringone. In the case of co-transformation of seedlings with two constructs, each Agrobacterium strain was diluted to O.D.$_{600}$ 4 and then mixing equal volumes in order to have each construct at O.D.$_{600}$ 2. Four-days-old seedlings were covered with Agrobacterium in 5% (w/v) sucrose and vacuum was applied twice for 2min each using a vacuum pump. Excess Agrobacterium solution was removed, the plates were re-sealed with micropore tape and returned to the culture room. Transiently transformed cotyledons were visualised under a CCD microscope (Section 2.11).

2.8.3 Transformation of plant cell cultures

2.8.3.1 Stable transformation of tobacco BY-2 cell cultures

1.5 ml of LB medium was inoculated with Agrobacterium strain GV3101 containing CDKC2 or CDKC2_D182N in pGFPNBIN, supplemented with rifampicin (20 μg/ ml),

gentamycin (35 μg/ ml) and kanamycin at 50 μg/ ml and grown for 24h at 28 ^{0}C. Agrobacterium and BY-2 cells were co-cultivated in deep Petri dishes (100x20 mM, Falcon), using 4 ml of 3-day-old BY-2 cells and 100 ml of Agrobacterium suspension. Petri dishes were sealed with Micropore tape, covered with foil and transferred to 25^{o}C, for 2 days without agitation. After co-culturing, the BY-2 cells were washed x3 to remove the Agrobacterium (8 ml of BY-2 medium was added into the co-culture and mixed using a Pasteur pipette, was centrifuged at 1000 rpm for 1minute and the supernatant discarded. After the final wash, the BY-2 cell pellet was resuspened in 1 ml of BY-2 medium and spread over a deep Petri dish containing MS + 4% phytagel (Sigma) plus appropriate antibiotics to select transformed calli, 500 μg/ ml carbenicillin (Melford Laboratories) to prevent Agrobacterium overgrowth and 20 μg/ ml rovral (Sigma) as a fungicide. The plates were sealed with Micropore tape and incubated horizontally in the dark at 25^{0}C. After 3-4 weeks, spherical calli were transferred into new phytagel plates containing antibiotics. At the same time, pieces of calli (1-2 cm across) were transferred to 50- ml conical flasks containing 20 ml of BY-2 media plus antibiotics and carbenicillin and allowed to grow at 25^{o}C on a shaker in the dark. When shaken cultures became thick with BY-2 cells (1 – 2 weeks), they were sub-cultured at a rate of 1 – 10 ml in 100 ml fresh media. The amount used was estimated, depending on the density of the initial culture. Carbenicillin selection was maintained for at least 3 rounds of sub-culturing (weekly 3-4 ml in 100 ml) in liquid medium to ensure that Agrobacterium was completely eliminated.

2.8.3.2 Transient transformation of Arabidopsis Col-0 cell cultures

On the day of subculturing a 7-days-old Arabidopsis cell culture, a 5 ml liquid culture of Agrobacterium containing the construct of interest was allowed to grow overnight at 28^{o}C. The next day, 3 ml of the overnight culture was placed into 100 ml of LB medium and grown at 28^{o}C until O.D. = 0.6. When the bacterial culture (or cultures) had reached the required O.D., it was spun down at 3,000 rpm for 10min and supernatant discarded. The bacterial pellet was resuspended into 1 ml of transformation buffer [10 mM MES, pH5.5 (with 1M KOH), 10 mM $MgCl_2$, 200μM acetosyringone (Sigma)] per 50 ml of bacterial cell culture. The resultant 1 ml suspension was added into 50 ml of 1-day-old

Col-0 culture and allowed to grow for 3 days as described above (Section 2.3.1). In the case of co-transformation, where two constructs were transformed into the same Col-0 culture, 500 µl of each bacterial suspension was used.

2.9 Immunolabelling of BY-2 cell cultures

2.9.1 Cell fixation

500 µl of a 3-days-old BY2 cell suspension was fixed with an equal volume of 4% (w/v) paraformaldehyde (PFA) in microtubule stabilizing buffer [MSB; 50 mM 2-[4-(2-sulfoethyl)-piperazin-1-yl]ethanesulfonic acid (PIPES; Sigma), 5 mM EGTA (Sigma), 5 mM MgSO4, pH 7.0, adjusted with KOH] for 25 minutes. Cells were collected by centrifugation at 1,000 rpm for 30 sec on a tabletop centrifuge, resuspended in 500 µl of 4% PFA in MSB and aliquots of the suspension were transferred onto 8-well slides (ING Biomedicals Inc, Ohio, USA). Slides were placed inside a square Petri dish on wet paper and incubated for 35 min at room temperature. Following fixation, samples were washed with MSB / 0.1% (v/v) TritonX-100 (Sigma) 5x10 min and with distilled water (5x10 min).

2.9.2 Enzymatic digestion and DMSO treatment

Cell-wall degrading enzymes (2% (w/v) cellulase Onozuka R-10 and 0.5% (w/v) Pectolyase Y-23) were dissolved in MSB / 0.1% Triton X-100 and spun at 13,000 rpm for 1 min to remove any non-dissolved particulates. Samples were digested for 20 min at RT and then washed 5 times, 10 min each wash, with MSB / 0.1% Triton X-100.

After cell wall digestion, cell were incubated in 10% (v/v) DMSO / 3% (v/v) NP-40 in MSB for 50 minutes and then washed thrice in MSB / 0.1% Triton X-100; 10 min each wash.

2.9.3 Antibody labelling and signal detection

Cells were incubated with 2% (w/v) BSA in MSB for 30 min at 32^{o}C, then in the primary antibody; anti-PolII (Abcam) at 1:100 dilution in 3% (w/v) BSA / MSB for

overnight incubation at 4°C, washed C. 5 times for 10 min each with MSB / 0.1% Triton X-100 and then incubated in the secondary antibody; 1:200 dilution of a mouse Alexa Fluor 488 (Sigma) in 3% BSA/ MSB for 2h at room temperature. After washing and mounting in antifade solution Citifluor (Citifluor Ltd), signal was visualized using a CCD microscope comprised of Nikon Eclipse 600 microscope coupled to a Hamamatsu Orca HQ cooled CCD digital camera. GFP was viewed using a standard FITC filter block and Alexa Fluor 488 using a TRITC filter (Section 2.11).

2.10 Protein methods

2.10.1 Protein extraction

2.10.1.1 Plant cell cultures

Cell suspensions were spun at 1,000 rpm for 5 min and the pellet was briefly dried on Whatman filter paper. Dried pellets were weighed and homogenized in a mortar with pestle and liquid nitrogen. Homogenized powder was transferred either into a 1.5 ml eppendorf tube or a 15 ml tube. 1 ml or 1.5 ml of protein extraction buffer [100 mM Hepes (pH7.5), 5% (v/v) glycerol, 50 mM KCl, 5 mM EDTA, 5 mM NaF, 0.1% (v/v) Triton-X 100, 15 mM sodium b-glycerophosphate, 0.5 mM sodium vanadate; plus protease inhibitor cocktail and 1 mM DTT just before use] was added per 1gr of dried powder, and the mixture allowed thawing on ice (around 30min). After thawing, the mixture was homogenised with a glass or plastic rod attached to an electric drill head, (still on ice). The homogenate was spun at 13,000 rpm at 4°C for 30 min. The supernatant was filtered using a 0.2 μm syringe filter (Sartorius UK, Ltd) to further clarify the extract. Aliquots of the extract were prepared, flash-frozen in liquid nitrogen and stored at -80°C until needed.

2.10.1.2 Nuclear protein isolation

Frozen nuclear (Section 2.6) pellets (-80°C) were placed on ice and 600-700 μl of extraction buffer [0.1M HEPES pH7.5 (adjusted with 1M HCL)], then high salt/detergent solution (7M urea, 0.3M NaCl, 50 mM NaH_2PO_4, 1 mM PMSF, 1 mM DTT) were added I and the thawed contents were homogenized on ice for 10-15 sec

using a plastic pestle attached to a mechanical pestle homogenizer. The whole mixture was transferred into another eppendorf containing another nuclear pellet and the procedure was repeated until all nuclear pellets were homogenized. The homogenate was incubated for 30 min on a rocking table at 4^{o}C (to increase protein solubilization). Following incubation the homogenate was spun on a tabletop centrifuge at 13,000 rpm for 30 min at 4^{o}C. The supernatant was recovered and used for protein precipitation, whereas the resulting insoluble pellet was stored at -80^{o}C,

Removal of urea was imperative for the downstream proteomics analysis. To 600-700 µl of extract, 4 volumes of methanol was added, vortexed well, then 4 volumes of chloroform was added and vortexed again and finally. Three volumes of sterile distilled water were added followed by further vortexing. The mixture was centrifuged for 1 min at 3,000 rpm on a RT6000B centrifuge and the top aqueous layer removed and discarded. Then 4 volumes of methanol was added, the mixture vortexed and spun for 5 min at 3,000 rpm. The resulting pellet was re-suspended (as a suspension) in 5 volumes of ice-cold 80% (v/v) acetone and spun for 5 min at 3,000 rpm and the supernatant discarded. I repeated the acetone wash 3-4 times and finally re-suspended the protein pellet in 80% (v/v) acetone. The acetone suspension was stored at -20^{o}C until used for proteomics analysis (Section 2.12)

2.10.2 Bradford assay

The concentration of extracted proteins was estimated with the Bradford assay (Bradford, 1976). A standard curve for protein concentration was created using BSA (Albu min Bovine Serum). BSA solution was diluted with sterile water to obtain a concentration of 1mg/ ml. In a plastic 96-well plate 0 – 8 µg of BSA were added, in 1 µg step increase, in duplicates. For the experimental samples 1 µl of protein extract was used in duplicates. Two hundred microlitres of Bradford reagent (BioRad) was added in each well and the absorbance was measured at 595nm on a 96-well compatible spectrophotometer.

2.10.3 SDS polyacrylamide gel electrophoresis (Laemmli, 1970)

Protein extracts were separated using the Discontinuous Sodium dodecyl sulphate (SDS) – Polyacrylamide Gel Electrophoresis" (SDS-PAGE) (Laemmli, 1970). The system is called "discontinuous" because it uses two different types of buffers; the stacking gel buffer and the resolving gel buffer. The SDS-coated proteins in the sample are pushed through a gel of high porosity (stacking gel) by a moving boundary created when an electric current is passed between electrodes. At the end of the stacking gel, SDS-protein complexes accumulate as a very thin line on the surface of resolving gel allowing high degree of resolution based on protein sizes

2.10.3.1 Preparing and running the SDS-PAGE gel

Acrylamide was purchased as a 30% (w/v) solution, mixed with N,N'-methylene-bis-acrylamide (bis-acrylamide) at a ratio 37.5:1 (Severn Biotech Ltd.). The donor of oxygen free radicals for polymerization of acrylamide is a mMonium persulfate (APS) and the catalyst used for the production of radicals is N,N,N',N'-tetramethylethylenediamine (TEMED).

Glass plates and spacers were washed with distilled water and carefully cleaned with ethanol. They were sandwiched together and fir mly fitted in a stable base. Firstly, the resolving gel was prepared and then the stacking gel (Table 2.5; indicated amounts for preparation of 10% resolving gel and 4.5% stacking gel). After pouring the resolving gel between two plates it was overlaid with water-saturated butanol to prevent oxygen from diffusing into the gel and inhibiting polymerization. The resolving gel was poured so that at least 1 cm was left available for the stacking gel (mini-gel system; Bio Rad). After polymerization of the resolving gel, butanol was removed by multiple washes with distilled water. A five percent stacking gel was poured on the top of resolving gel and the comb was inserted carefully into the stacking gel making sure not to trap air bubbles. After polymerization, the comb was removed and the assembled gel inserted into the electrophoresis tank. The inner and outer chambers of the tank were filled with 1X running buffer [0.33M Tris-base, 1.92M Glycine and 10% (v/v) SDS]. Mini-gels were always run at 200V for 45-60 minutes.

Reagent	Resolving gel (10%)	Stacking (4.5%)
Acrylamide	6 ml	900 µl
1.5M Tris pH 8.8	4.5 ml	-
0.5M Tris pH 6.6	-	1.5 ml
Distilled water	7.5 ml	3.6 ml
APS	80 µl	20 µl
TEMED	10 µl	10 µl

Table 2.5 Chemical reagents and their respective amounts for preparing 10% resolving and 4.5% stacking gel.

2.10.3.2 Gel staining

Coomassie-based dyes, Coomassie Brilliant Blue R-250 (Sigma) and colloidal SimplyBlue Safestain coomassie (Invitrogen), were used for staining of protein gels. A solution of the first dye was prepared by adding 25 mg of the powder in 100 ml of 40% (v/v) methanol / 10% (v/v) acetic acid and filtering the solution using Whatman filter paper. The colloidal Coomassie was purchased as a ready-to-use solution from Invitrogen**.** When Coomassie Brilliant Blue R-250 was used for gel staining, the gel was transferred into a square Petri dish containing 50-75 ml of the dye solution and allowed to stain for 1.5 hours with gentle agitation. After that, the dye solution was poured into a container and re-used up to 3 times. For destaining, the gel was soaked in a solution of 30%methanol / 10% acetic acid in the presence of a piece of foam rubber, used for absorbing the dye. When most of the background staining was removed, the gel was washed thoroughly in distilled water.

In the case of colloidal Coomassie, the dye was poured on the gel and allowed to stain for 1 hour. The gel was destained with distilled water. The gel was then either photographed or used for blotting

2.10.4 Western blot

Western blot (Tobin *et al*., 1979) was used for transferring electrophoretically separated proteins from a gel to a membrane, usually nitrocellulose or polyvinylidene difluoride (PVDF), where they could be probed using specific antibodies.

After having separated proteins using SDS-PAGE, the unstained gel was washed in 1X blotting buffer (37.5 mM Trizma base, 0.3M glycine, 20% (v/v) methanol) for 30 minutes. After equilibration in blotting buffer the gel was assembled in the blotting apparatus with the order shown in Figure 2.5. The membrane I used was a Protran BA85 nitrocellulose membrane (Whatman, Schleicher and Schuell). After assembling the apparatus, the case was closed and put vertically into the case holder and both put into the buffer tank. A plastic box containing ice was placed touching the tank in order to prevent the buffer from over-heating. The buffer tank was filled with 1X blotting buffer and a magnetic stirrer was put at the bottom of the tank to ensure that the temperature of the buffer does not fluctuate. Protein transfer was carried out at 200 Volts for 60 minutes. After transfer, the membrane was removed and washed twice for 5 min each in distilled water. Protein transfer was verified by staining with PosceauS solution [0.1% (w/v) PonceauS (Sigma) in 5% (v/v) acetic acid] for 5 min. Destaining was done by washing in distilled water until sufficient background was removed.

The destained membrane was incubated in 5% (w/v) blocking solution [Marvel® milk powder in 1xTBS (50 mM Tris-HCl, pH 7.4, 150 mM NaCl)) / 0.01% Tween-20] for 30 min to block non-specific antibody binding sites. During the blocking, I prepared 10 ml of an anti-mouse or anti-rabbit primary antibody at 1:1000 dilution in 1% (v/v) blocking solution, supplemented with 0.01% NaN_3 (Sigma) to prevent microbial outgrowth. Antibody solution was poured over the membrane and allowed to bind to target protein overnight at 4^{o}C on a shaker, making sure that the membrane is completely covered by the solution. Next day, antibody was returned into a plastic tube and stored at 4^{o}C for future use. The membrane was washed thrice with 1xTBS/0.01% Tween-20, 5 min each wash. I added the secondary antibody (mouse or rabbit HRP-conjugated) at 1:1000 dilution in 1% (v/v) blocking solution and incubated the membrane for 1 hour. After incubation, the membrane was washed again as above and then developed using the SuperSignal West Pico kit (Pierce); equal amounts (400 µl) from the peroxide and luminal enhancer solutions were pipetted up down on the whole surface of the

membrane for 5 minutes. The membrane was placed between two acetate sheets, sealed around with tape and transferred into a photographic cassette and placed Hyperfilm™ autoradiography film (GE Healthcare) on the membrane and exposed until the band/s of interest appeared. Exposure times varied for 5 min to 2 hours, depending on the amount of target protein present or the amount of antibody used.

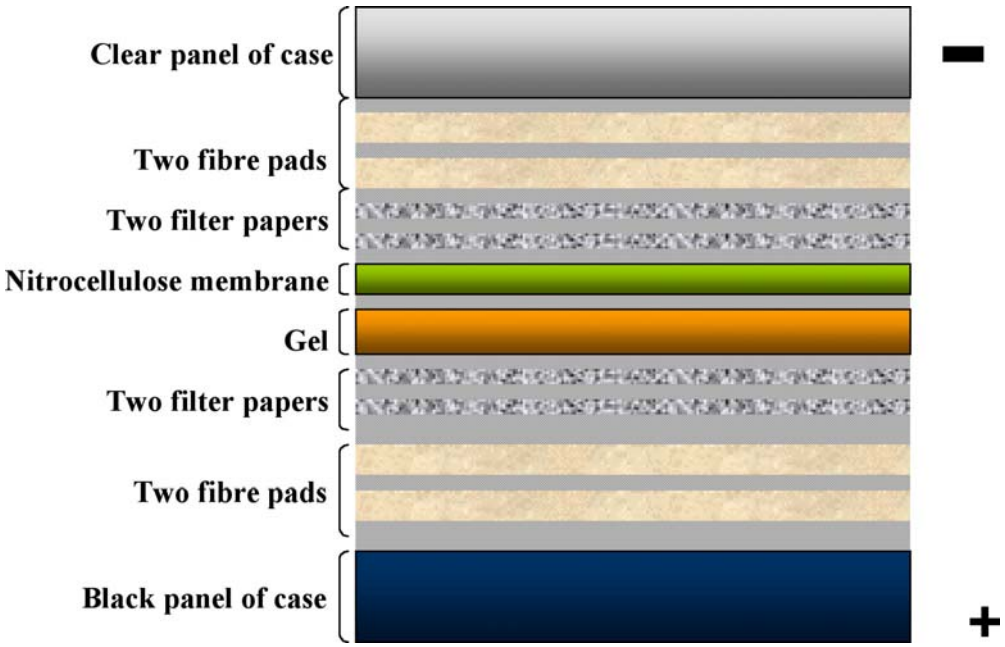


Figure 2.5 Illustration of the western blot transfer apparatus.

2.10.5 Kinase assay

Kinase assays were performed using expressed proteins produced as described in Sections 2.7.5.1 and 2.7.5.2. The components used for the assay and respective amounts are given at Table 2.5. The composition of kinase buffer (KB) was: 20 mM HEPES-pH7.5, 15 mM $MgCl_2$, 5 mM EGTA and 1 mM DTT. For preparation of the reaction

cocktail (RC) 200 µl KB, 2 µl 10 mM ATP was mixed with 4 µl γ-^{32}P-ATP (Amersham). Each component was added into individual 0.5 ml PCR tubes. After preparing the mixtures PCR tubes were transferred into a PCR machine and incubated at 30^{o}C for 30 minutes. After the kinase assay was completed, the content of each PCR tube was transferred into respective screw-top eppendorfs containing 2X sample buffer, and boiled at 95^{o}C for 5 minutes. Boiled samples were spun at 13,000 rpm for 1 min and either stored at -20^{o}C or used directly on SDS-PAGE. All of each sample (25 µl) was loaded on a 15% gel and run at 200V for 1 hour. After running, gel was dried overnight using the Gel Drying Kit (Promega) and following manufacturer's instructions. Briefly, the gel was soaked into 40% (v/v) methanol, 10% (v/v) glycerol and 7.5% (v/v) acetic acid for no more than 5 minutes. Then, the gel was placed between two moistened sheets of clear cellulose film and the sandwich was clamped between two plastic frames. The gel was allowed to dry overnight in the fume hood and behind a crop protective Plexiglas screen. The next day, the dried gel was cut out and placed into a photographic cassette for exposure to photosensitive photographic film. In order to increase the sensitivity of the detection, the dried gel was covered with a Biomax Transcreen HE intensifying screen (Kodak) and the photographic film (Kodak) was put between the two sheets of the screen. The amplification screen receives the signal transmitted by the gel, it amplifies it and transfers the much stronger signal to the photographic film. When the film was to be exposed for more than 2 hours, the whole cassette was put in a bag and stored at -80^{o}C.

2.11 Microscopy and image handling

Plants, BY2 and Col-0 cells expressing GFP and mRFP fusions were visualised using the Cairn CCD system, consisting of a Nikon Eclipse 600 microscope with a Hamamatsu Orca HQ cooled CCD digital camera. GFP was viewed using a standard FITC filter block, while mRFP was visualized using the TRITC filter block. Images were recorded with a mounted Nikon E995 digital camera. Acquired images were deconvolved using the Metamorph deconvolution software. Further processing was done using the ImageJ (version 1.38x) and Adobe Photoshop (CS3 extended) software.

2.12 Proteomics methods

2.12.1 Proteolytic digestion of nuclear extracts

Tryptic digestion was performed on protein extracts initially resuspended in 80% (v/v) acetone (Section 2.9.1.2). For tryptic digest, the protein suspensions (described in Section 2.10.1.2) were spun at 16,200xg for 5 min to pellet the protein and the acetone was removed. The pellet was then dissolved in 100 µl of 0.5 % (v/v) *Rapid*Gest *SF*™ (Waters Ltd), 0.2 M triethyla mMonium bicarbonate (TEAB), pH 8.0 and mixed by pipetting. The suspension was heated in a water bath at 80-95 ºC for 10 min; 10 µl DTT reducing agent was added; the solution was incubated at 60 ºC for 1 h and then 5 µl of 200 mM iodoacetamide (IAA) was added. This was incubated for 30 min at room temperature in the dark. Trypsin was prepared by dissolving 20 µg in 20 µl of water, and 8 µl was used for the digestion. Digestion was allowed to continue overnight at 37 ºC and then concentrated in a SpeedVac™ down to 30 µl and used directly for MS analysis or used for iTRAQlabelling (Section 2.12.2).

2.12.2 iTRAQ labelling

For quantatitive labelling, an iTRAQ Reagents kit was used (Applied Biosystems). First, iTRAQ reagents were equilibrated to room temperature and then dissolved in 70 µl of 100% ethanol. The reagent was added to the 30 µl concentrated digest (Section 2.11.1) and the mixture left to stand for 1 h at room temperature. After labelling, the solution was dried using a Speedvac™. The dried sample was redissolved in 0.5% (v/v) trifluoroacetic acid (TFA), left for 10 min at room temperature and then spun using a benchtop microcentrifuge at maximum speed to remove any precipitate. The supernatant was transferred to another tube and stored at -20 ^{o}C until used.

2.12.3 Liquid chromatography and MS (undertaken by Dr. Mike Naldrett and Dr. Gerhard Saalbach)

2.12.3.1 Strong cation exchange

To reduce the complexity of peptides loaded on the mass spectrometer, strong cation exchange (SCX) chromatography was performed using a PolySULFOETHYL A™ column (5 µm, 200 Å, 1 mM i.d. x 150 mM; PolyLC, Columbia, MD, USA). Digested

peptides were equilibrated in buffer A (10 mM KH_2PO_4, 20 % acetonitrile, pH 2.7) and elution was performed by increasing the amount of buffer B (10 mM KH_2PO_4, 20 % acetonitrile, 1 M KCl, pH 2.7) using the following gradient: 0 % B over 20 min, 0-6 % B over 5 min, 6-16 % over 16 min, 16-50 % over 5 min, hold at 50 % for 5 min, then 50-100 % over 1 min and finally hold at 100 % for 5 min before re-equilibrating at 0 % B for 27 min before reinjection. The flow rate was 50 µl/min. 2 min fractions were collected and stored until needed for MS analysis.

2.12.3.2 Sample loading and MS analysis of iTRAQ™-labelled samples

The ultraviolet absorbance of each SCX fraction, calculated from the integrated chromatogram, was used to normalise the amount of material analysed. Nano-LC-MS/MS experiments were performed by Dr Gerhard Saalbach (JIC Proteomics Facility) on an LTQ-Orbitrap XL™ mass spectrometer (Thermo Fisher Scientific Inc) coupled to a nanoAcquity UPLC™ (Waters Ltd). The LC system was run at a flow rate of 250 nl/ min and coupled to the mass spectrometer via an ion source (Proxeon) with a nanospray emitter (SilicaTips™, 10 µm, New Objective). Calculated portions were dissolved in 0.1 % trifluoroacetic acid **(**TFA) and individually injected onto a trap column (Sy mMetry® C18, 5 µm, 100 Å, 180 µm x 20 mM, Waters) and desalted. The trap was then switched in-line and the concentrated sample was separated on a reverse-phase nano column (BEH C18, 1.7 µm, 130 Å, 75 µm x 250 mM, Waters Ltd). Peptides were separated and eluted with a gradient of 5-45 % acetonitrile in water/0.1 % formic acid at a rate of 0.5 %/min.

The mass spectrometer was operated in positive ion mode at a capillary temperature of 200 °C. The source voltage and focusing voltages were tuned for the transmission of MRFA peptide (m/z 524) (Sigma). The Orbitrap™ was run with a resolution of 30,000 over the MS range from m/z 350 to m/z 2000 and an MS target of 10^6 and 1 s maximum scan time. Data-dependent analysis of iTRAQ™-labelled samples was carried out in Orbitrap-IT parallel mode using PQD fragmentation on the 6 most abundant ions in each cycle.

For PQD, settings were chosen according to Bantscheff *et al*. (2008). The collision energy was set to 29, the activation Q to 0.55, the activation time to 0.4 s, and an

isolation width of 2 was used. The MS2 was triggered by a minimal signal of 2000 with an AGC target of $5x10^4$ ions and 200 ms scan time. For selection of 2+ and 3+ charged precursors, charge state and monoisotopic precursor selection was used. Dynamic exclusion was set to 2 counts and 90 s exclusion time with an exclusion mass window of ±30 ppm. MS scans were saved in profile mode while MS2 scans were saved in centroid mode.

2.12.4 Data analysis

2.12.4.1 Database searching

Tandem mass spectra were extracted and the charge state determined by BioWorks™ (version 2.0; ThermoElectron). De-isotoping was not performed. All MS/MS samples were analyzed using Mascot (version 2.2; Matrix Science, London, UK;) and SEQUEST (version 27; ThermoElectron, San Jose, CA, USA). Mascot and SEQUEST were set up to search the "atg_contami_20080413b" database (33024 entries) with the digestion enzyme as trypsin. Mascot was searched with a fragment ion mass tolerance of 0.50 Da and a parent ion tolerance of 5.0 PPM. SEQUEST was searched with a fragment ion mass tolerance of 0.50 Da and a parent ion tolerance of 0.0071 Da. The iodoacetamide derivative of cysteine (carboxyamidomethyl cysteine) was specified in Mascot and SEQUEST as a fixed modification. Oxidation of methionine and the iTRAQ label were specified in Mascot and SEQUEST as variable modifications.

2.12.4.2 Criteria for protein identification

Scaffold (version Scaffold_2_04; Proteome Software Inc, Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the PeptideProphet algorithm (Nesvizhskii, 2003; Keller, *et al*., 2002). Peptide identifications were accepted only for those that had probability of being correct equal or more than 99%. Protein identifications were accepted if they could be established at greater than 95.0% probability and if they contained at least 2 identified peptides. Again, protein probabilities were assigned by ProteinProphet algorithm. Proteins that

contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

2.13 Gene ontology enrichment analysis

Classification of identified proteins in Gene Ontology (GO) groups was done using the BiNGO plugin (version 2.3; Maere *et al*., 2005) within the Cytoscape software (version 2.6.2; Shannon *et al*., 2003). The enrichment analysis was performed using the hypergeometric statistical test. Correction was applied using Benjamini and Hochberg False Discovery Rate (Benjamini and Hochberg, 1995) at 0.05 significance level. To minimise the impact of applying multiple testing, fewer GO categories were tested by using the GOSlim_Plants category; a slimmed down version of the full GO hierarchy. BiNGO tool retrieves default annotations from the GO database available in NCBI (www.ncbi.nlm.nih.gov/Ftp/).

Chapter 3 – Proteomics analysis of the *Arabidopsis* nucleus

3.1 Aims

The aim of this chapter is to optimize the nuclei isolation protocol from *Arabidopsis* cell cultures of different ages and then characterize the nuclear proteome using in-solution peptide separation and MS-based protein identification. Identification of the nuclear proteome will provide a reference map of the protein composition of this subcellular organelle in *Arabidopsis*, and possibly reveal new components of known nuclear processes or even discover possible new functions of this organelle. The generation of a high-resolution nuclear proteome, combined with available technologies for differentially labelling proteins (see Section 1.4.4), should allow us to initiate studies on quantitative changes of the whole nuclear proteome (see Chapter 4).

3.2 Optimisation of nuclei isolation from *Arabidopsis* cell cultures

The basic protocol for *Arabidopsis* nuclei isolation is described in Chapter 2, Section 2.7. Here I will present problems encountered in the initial isolation protocol (Pendle *et al*., 2005) and methods that were used to circumvent them.

3.2.1 Optimising the release of protoplasts

The initial stage of the protocol involves isolating protoplasts by means of cell wall degradation. The enzymes used initially were cellulase Onozuka R-10 (hereafter called cellulase) and pectolyase Y-23 (hereafter called pectolyase) at 2% (w/v) and 0.04% (w/v), respectively. At the beginning of the project, application of the protocol with the aforementioned enzyme concentrations yielded good quality protoplasts that were uniformly spherical and well separated. However, different batches of pectolyase clearly vary in potency: a new batch of pectolyase used at 0.04% (w/v) lysed protoplasts within 30 min. Therefore, each batch of enzyme needs to be titred in order to determine the optimal concentration of pectolyase: different dilutions of the new enzyme stock were incubated with cell cultures for different periods (Figure 3.1, 3.2 and 3.3) to

determine which treatment produced the best batch of protoplasts. In parallel, I included different dilutions of the enzyme Driselase. Figure 3.1A shows the effect in digestion of three different dilutions of pectolyase [0.001%, 0.005 and 0.01% (w/v)] in the presence of 2% (w/v) cellulase after 30 min at room temperature. Cell clumps were present for all the enzyme dilutions, something expected for this short period of digestion. However, in 0.001% (w/v) pectolyase, vacuoles were released from cells at the periphery of the clumps (Figure 3.1B). Vacuole release suggests that the concentration of pectolyase is not sufficient to separate the cell clumps before the cell membranes are damaged. For the remaining two concentrations of pectolyase, 0.005% and 0.01%, no apparent differences could be detected. Figure 3.1C shows the effect of 0.01% pectolyase and 0.01% Driselase in the presence of 2% cellulase after 30 min of digestion. As for the two higher concentrations of pectolyase, no apparent differences could be seen between the two concentrations of Driselase.

After 90 min of digestion, differences among the different treatments were clear (Figure 3.2). In 0.001% pectolyase, the vast majority of cells/protoplasts were lysed or appeared deformed (Figure 3.2A, i) so the sample was discarded. In 0.005% pectolyase, low numbers of lysed cells were observed, but the majority of cells/protoplasts collapsed (Figures 3.2A, ii & 3.2B), suggesting that cells are highly stressed and not fit to use in downstream steps. Pectolyase at 0.01% did not seem to have an adverse effect on the cells and the numbers of individual and round protoplasts increased considerably after 90 min incubation (Figure 3.2A, iii).

Figure 3.1 Optimisation of digestion: 30 minutes.

Representative images from each treatment show progression of cell wall digestion after 30min. (A) and (C) both shown at the same magnification: Scale bar: 50 μm.

(A) *Arabidopsis* cells digested with pectolyase in the presence of 2% cellulase. The different concentrations of pectolyase are indicated on the image (0.001%, 0.005% and 0.01% [all as (w/v)]). Dashed rectangles 1 and 2 are shown magnified in (B).

(B) Magnification of rectangles shown in (A) showing the release of vacuoles (arrows) from protoplasts due to over-digestion.

(C) *Arabidopsis* cells digested with two different concentrations of Driselase (0.01% and 0.1% (all as (w/v) plus 2% (w/v) Driselase.

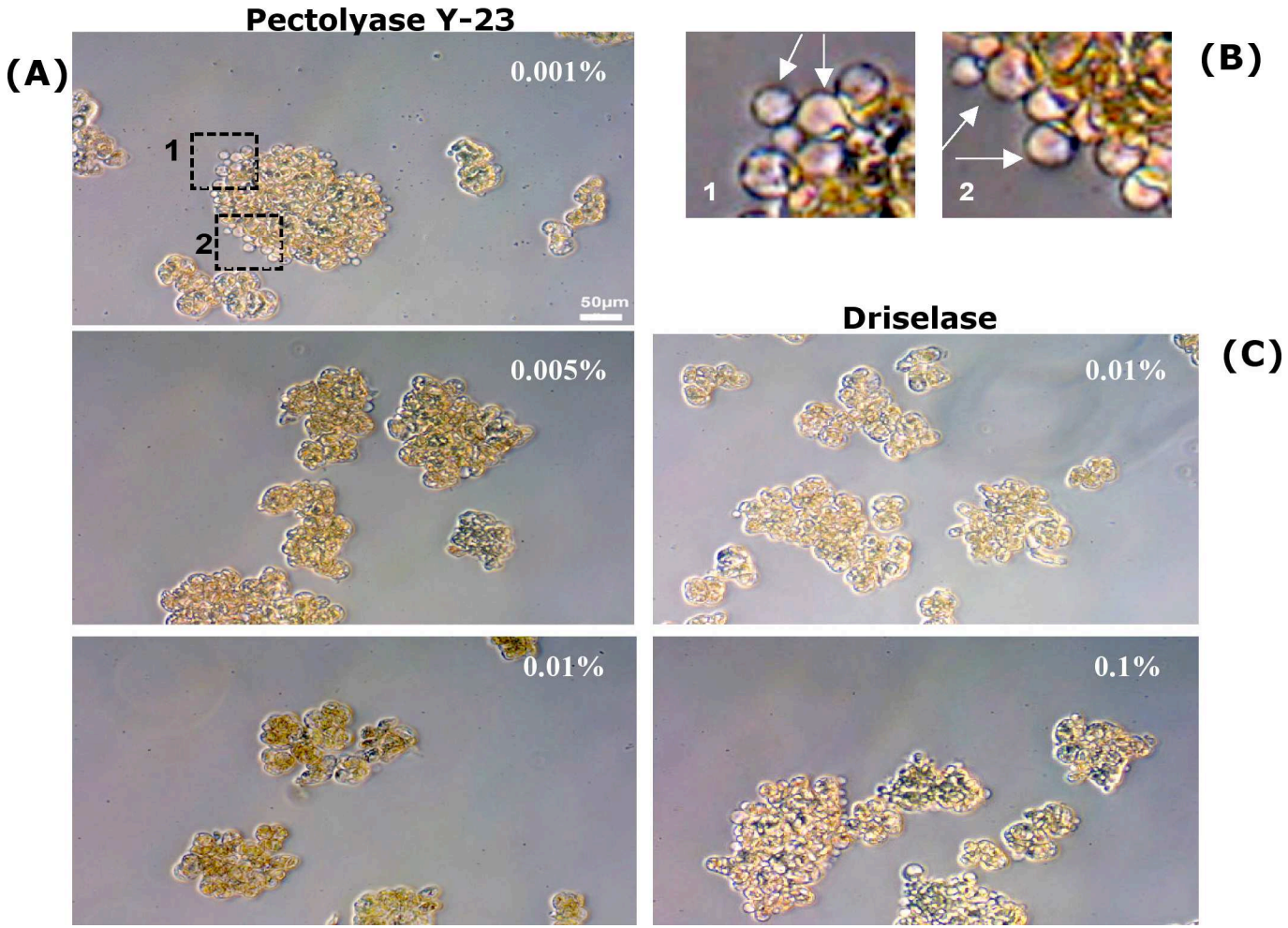

Pectolyase Y-23
(A)
0.001%
1
2
50µm
0.005%
0.01%
(B)
1
2
Driselase
(C)
0.01%
0.1%

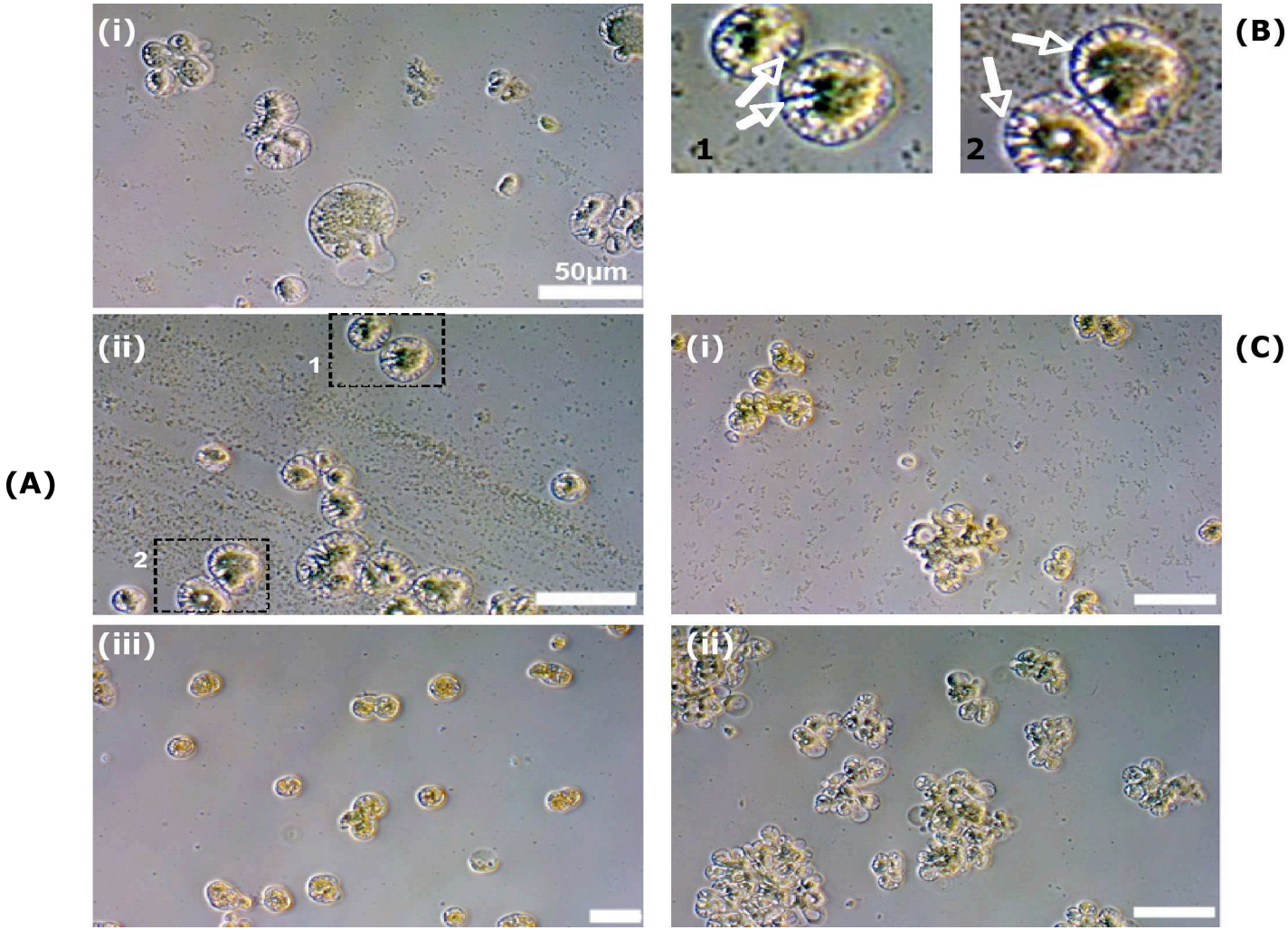
Figure 3.2. Optimisation of digestion: 90 minutes.

Representative images from each treatment (as shown in Fig 3.1) show progression of cell wall digestion after 1h 30min. Scale bars = 50µm

(A) Cells incubated in the presence of 2% cellulase (w/v) and different concentrations of pectolyase; (i) 0.001% (w/v), (ii) 0.005% (w/v) and (iii) 0.01% (w/v)

(B) Magnifications of dashed areas 1 and 2 shown in (Aii). Arrows indicate initial stages of cytoplasm collapse

(C) Cells incubated in the presence of 2% cellulase (w/v) and different concentrations of Driselase; (i) 0.01% (w/v) and (ii) 0.1% (w/v). Scale bar: 50 µm.


(A)
(i)
50µm
(ii)
1
2
(iii)
(B)
1
2
(C)
(i)
(ii)

In the case of 0.01% Driselase, no protoplast lysis could be seen but cell clumps could still be observed after 90 min (Figure 3.2C, i). At 0.1% driselase, cells still appeared in clumps after 90 min but some lysed protoplasts and vacuole release could also be seen (Figure 3.2C, ii).

After 2 hours of treatment, the optimal enzyme and respective concentration for protoplast production was determined. Figure 3.3 shows the appearance of cells/protoplasts after 2 h of incubation with 0.005 % or 0.01 % pectolyase and 0.01 % or 0.1 % Driselase, in all cases in the presence of 2% cellulase. In the case of low concentrations of pectolyase, a considerable number of cells had formed protoplasts but the products were not as round as for the higher concentration of pectolyase, where the vast majority of protoplasts were single and round (Figure 3.3A). When Driselase was used at the lower concentration, cell clumps could still be observed after 2 h, whereas at the higher concentration the majority of cells were in the form of isolated and individual protoplasts (Figure 3.3B), though the shape of protoplasts was not as uniformly spherical as with the 0.01% pectolyase / 2% cellulase treatment. Based on these observations, the composition of the digestion mixture used for the subsequent experiments was adjusted to 2% cellulase and 0.01% pectolyase.

3.2.2 Optimisation of the isolation of nuclei from purified protoplasts

Separation of nuclei from purified protoplasts is the second crucial step in the isolation protocol, as inadequate release of nuclei would result in preparations of low yield and reduced purity. Following protoplast lysis in the presence of detergent (see Section 2.7), the majority of nuclei were surrounded by substantial cytoplasmic material, (Figure 3.4A). To reduce this cytoplasmic contamination, I passed the crude nuclear suspension through a 23-gauge needle using a syringe, after mild homogenisation in a glass-on-glass homogenizer. This method removed most of the cytoplasmic remnants from the nuclei, producing “clean” nuclei as judged by microscopy (Figure 3.4B).

Figure 3.3 Optimisation of digestion: 180 min.

Representative images from each treatment (as shown in Figure 3.1) show progression of cell wall digestion after 1h 30min. Scale bars = 50µm

(A) Cells incubated in the presence of 2% cellulase (w/v) and different concentrations of pectolyase; (i) 0.005% (w/v) and (ii) 0.01% (w/v)

(B) Cells incubated in the presence of 2% cellulase (w/v) and different concentrations of Driselase; (i) 0.01% (w/v) and (ii) 0.1% (w/v).

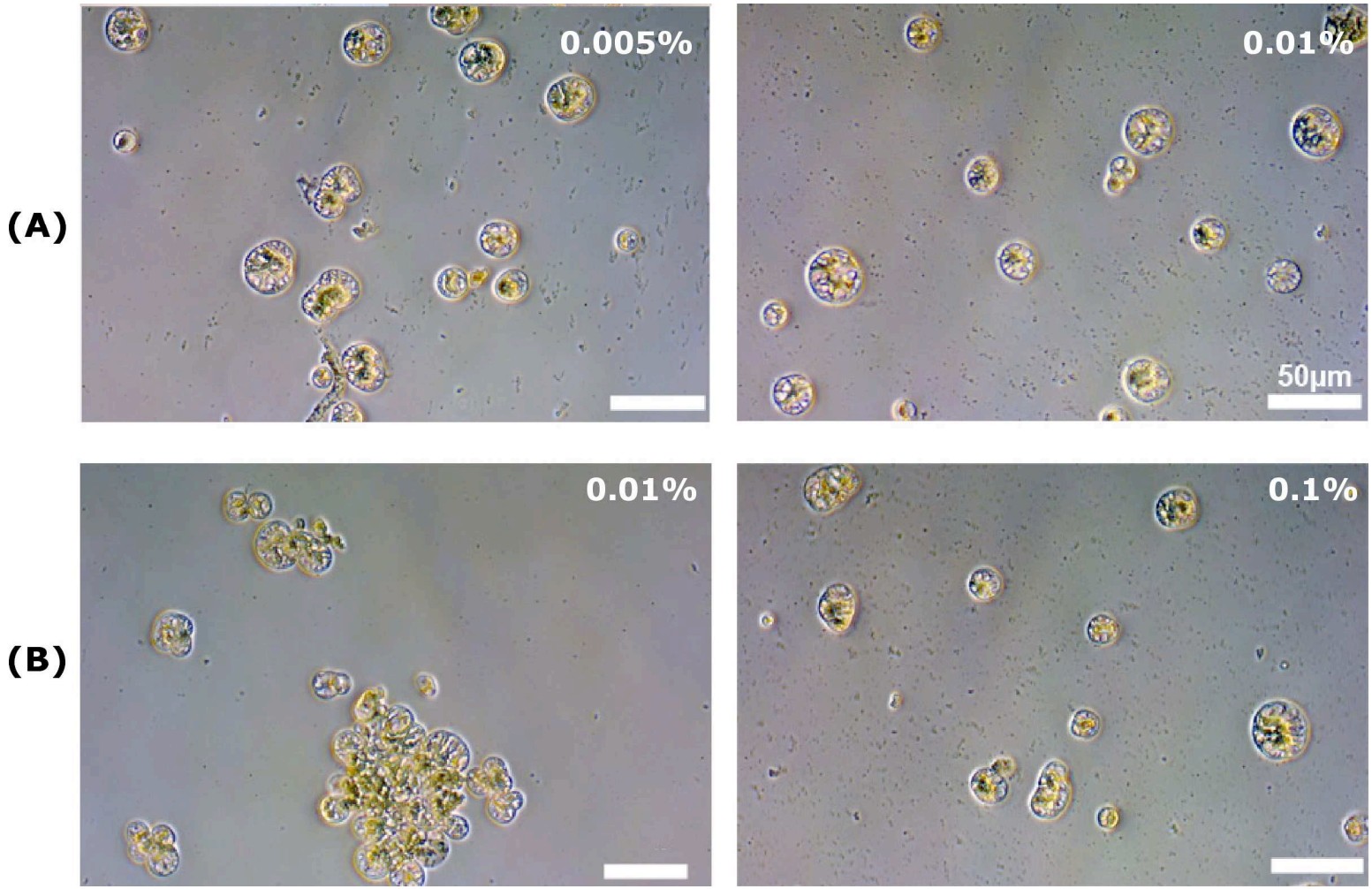

(A)
0.005%
0.01%
50µm
(B)
0.01%
0.1%

Figure 3.4. Optimisation of nuclear isolation.

Nuclei released from *Arabidopsis* protoplasts (see Figure 3.3A) by lysis in NIB and then gently broken using either:

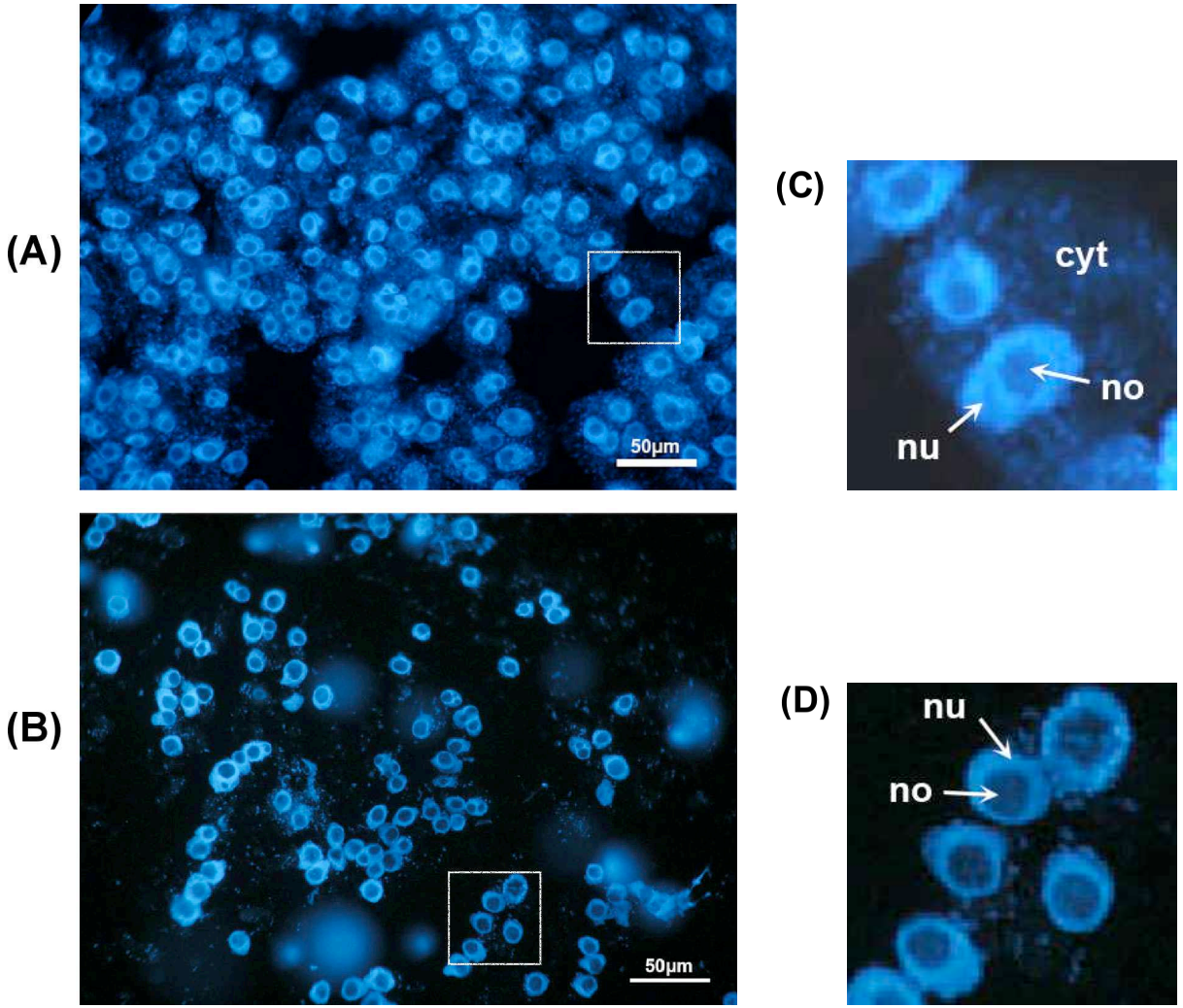
(A) A stainless steel homogeniser

or

(B) A 23G-gauge needle followed by stainless steel homogenisation.

(C) - (D) Magnification of the boxed areas in (A) and (B), respectively. nu: nucleoplasm, no: nucleolus, cyt: cytoplasmic contaminant

Scale bar = 50 μm


(A)
50µm
(B)
50µm
(C)
cyt
no
nu
(D)
nu
no

Application of this optimized nuclear isolation protocol using 2-day-old cell cultures gave consistently good nuclei preparations with minimal cytoplasmic contamination when observed under the microscope. However, nuclear preparations from 8-day-old cell cultures suffered from low yields of nuclei and most importantly, the vast majority of nuclei appeared to be damaged or lysed (Figure 3.5B). I reasoned that nuclei lysis might be due to the age of nuclei in combination with the high concentration of non-ionic detergent (0.1% (v/v) Triton X-100). Reduction of Triton X-100 0.05% and 0.01% did not improve nuclear morphology, suggesting that the main reason could be the age of the material/nuclei. To test this, a 6-day-old culture with a mitotic index of 0.3% (close to the 0.1% of 8-day-old cultures) was used and the yield of morphologically "normal" nuclei increased considerably, as judged by DAPI staining (compare Figure 3.5A with Figure 3.5B). Following these results I decided to use nuclei from 6-days-old cells for subsequent experiments. The problem of isolating high quality nuclei from older cells remains to be solved.

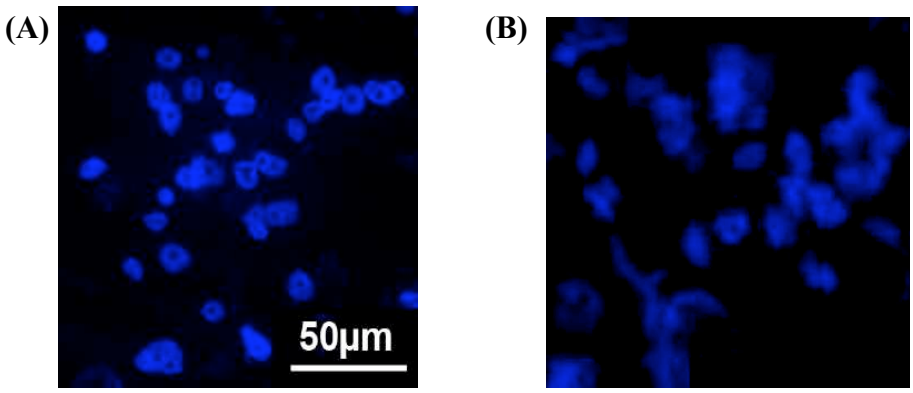


Figure 3.5 Isolated nuclei from *Arabidopsis* cells at early and late stages of culture growth.

Nuclei were isolated using the optimized isolation protocol. DAPI stained fixed nuclei from 6-days-old (A) and 8-days-old (B) cell cultures are shown. Scale bar = 50 µm.

3.3 Phenotypic characterisation of nuclei isolated from cells at day2 and day6 of culture growth

Optical sections across the centre of nuclei originating from 2-days-old and 6-days-old cultures indicated that nuclear size decreases as the cells progress through the growth cycle (data not shown). Measurements of at least 50 nuclei from each sample showed that, on average, the diameter of nuclei from 2-days-old cells was almost 3.5 times bigger than that from 6-days-old cells (Figure 3.6A) and, if one assumes that the nucleus is a regular sphere, the change in nuclear volume is approximately 6.5-fold.

An even greater change was also true for nucleoli. Estimation of nucleolar diameter indicated that a 4.6-fold change occured between day2 and day6 (Figure 3.6B). Again, if the nucleolar compartment is considered as spherical the volume change is in the range of 10-fold. Therefore, both the nucleus and the nucleolus undergo a marked reduction in size as cells progress from proliferation to stationary growth phase.

Although extensive size variability has been observed for various organelles, such as ER and mitochondria, cells appear to retain a roughly constant karyoplasmic ratio [nuclear volume to cell volume ratio (Cavalier-Smith, 2005; Wilson, 1925)]. In yeast cells, the karyoplasmic ratio is maintained around 0.08 (Jorgensen *et al*., 2007; Neumann and Nurse, 2007). The observed difference in nuclear volume in *Arabidopsis* cell cultures, prompted me to examine whether the findings in yeast could apply in *Arabidopsis* cell cultures as well. After calculating cell sizes from day2 and day6 cell cultures (Figure 3.6), the karyoplasmic ratio under these specific culture conditions did not differ significantly between the two stages (Figure 3.6D).

Figure 3.6 Size variations of cell, nuclei and nucleoli between proliferating and stationary phase cell populations.

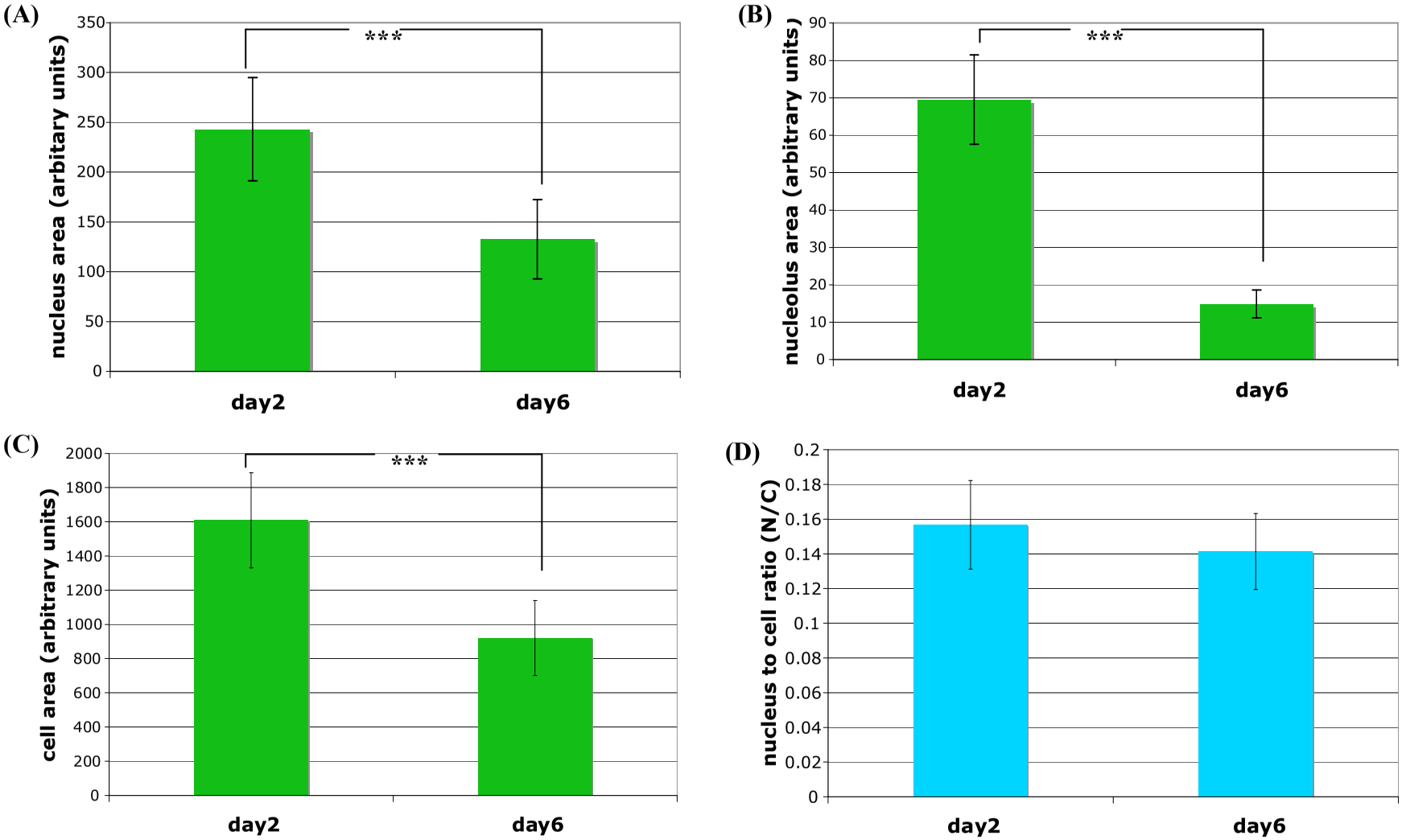
Measurements were performed on optical sections through the centre of the nucleus, in all cases. At least 50 cellular particles were measured in each case. Asterisks indicate significant differences at the level of $P<0.001$.

(A) Average nuclei area from day2 and day6 cells

(B) Average nucleoli area from day2 and day6 cells

(C) Average area of day2 and day6 cells

(D) Karyoplasmic ratio (N/C) in day2 and day6 cells


(A)
nucleus area (arbitary units)
350
300
250
200
150
100
50
0
***
day2
day6
(B)
nucleolus area (arbitrary units)
90
80
70
60
50
40
30
20
10
0
***
day2
day6
(C)
cell area (arbitrary units)
2000
1800
1600
1400
1200
1000
800
600
400
200
0
***
day2
day6
(D)
nucleus to cell ratio (N/C)
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
day2
day6

3.4 Assessment of the purity of nuclei preparations using biochemical methods

Microscopic observation of nuclear preparations gave an initial rapid assessment of nuclear purity. A more rigorous assessment used two antibodies against proteins from different parts of the cell: one against the cytoplasmic eukaryotic elongation factor isoform 4G (eIF4isoG); and a second one against the C-terminal domain of the large subunit of RNA polymerase II, to check for enrichment of nuclear proteins. To follow enrichment, four protein samples were analysed: (i) proteins from homogenised cells (soluble cell extract; SCE), (ii) proteins isolated from the protoplasts lysed in the NIB buffer (cytoplasmic extract, CytE), (iii) urea-soluble nuclear extract from day2 cells (N2) and (iv) urea-soluble nuclear extract from day6 cells (N6). Figure 3.7 shows the SDS-PAGE profiles and western blots with the aforementioned antibodies. Obvious differences in the banding pattern between the nuclear (N2 and N6) and non-nuclear (SCE and CytE) samples are marked with double-headed arrows (Figure 3.7A). This indicates that there has been a selective enrichment of proteins in the nuclear preparations and these most probably correspond to nuclear proteins.

Western blot analysis of the four samples against cytoplasm-specific and nucleus-specific proteins confirmed the enrichment and was routinely used to assess the purity of nuclear preparations. Using the antibody to eIFiso4G no protein was detected in the nuclear samples whereas eIFiso4G was present in non-nuclear extracts, SCE and CytE (Figure 3.7B). In contrast, RNA Pol II could be detected only in the nuclear extracts (Figure 3.7C). The presence of extra bands in the nuclear extracts when using the antibody against RNA Pol II is probably due to partial protein degradation.

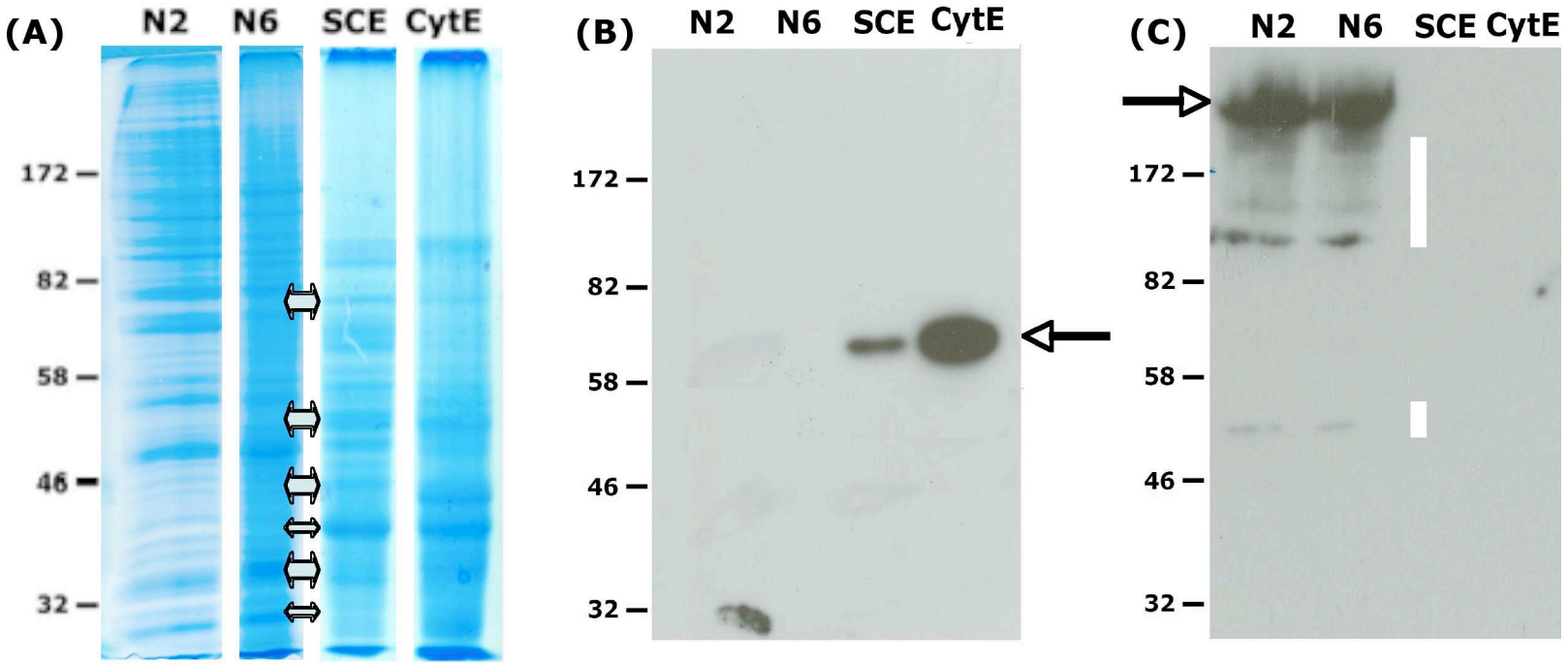
Results from the biochemical analysis suggest that nuclei preparations are enriched for nuclear proteins and depleted in contaminating cytoplasmic proteins (Figure 3.7B, C), and are therefore suitable for proteomic analysis.

Figure 3.7. Assessment of nuclear protein extracts using compartment-specific antibodies.

(A) Protein extracts from whole cells (SCE), cytoplasmic protein extracts (CytE) and nuclear extracts from day2 (N2) and day6 (N6) cells were run on a 10% SDS-PAGE gel and stained with colloidal Coomassie Blue stain.

(B) - (C) Protein samples shown in (A) were transferred onto a nitrocellulose film and hybridised with antibodies against eIFiso4G (B) and the C-terminal domain of RNA Pol II **(**C).

Arrowheads indicate the position of the expected band. Double-headed arrows in (A) indicate differences in protein composition between nuclear (N2, N6) and non-nuclear (SCE, CytE) samples, or *vice versa*. Arrowheads in (B) and (C) indicate the position of the expected band. Lines on (C) indicate protein degradation products. The position of molecular weight markers is shown to the left of each figure.


(A)
N2
N6
SCE
CytE
172
82
58
46
32
(B)
N2
N6
SCE
CytE
172
82
58
46
32
(C)
N2
N6
SCE
CytE
172
82
58
46
32

3.5 Proteomic analysis of nuclear extracts

Three major independent proteomics experiments have been performed; namely K003, K004 and K005, presented in chronological order. The aim of all three experiments was to identify the *Arabidopsis* nuclear proteome and, in parallel, to perform a quantitative analysis between protein extracts from proliferating and quiescent cell populations (see Chapter 4 for quantitative nuclear proteomics). Due to reasons explained in Chapter 4, labelling efficiency of peptide mixtures for quantitative analysis was poor for the K003 and K004 experiments, but good for the experiment K003. Nevertheless, the protein identification data from the first two experiments are useful and were used, together with data from the K005 experiment, for assessing the purity of my nuclear preparations among the different experiments. K003 and K004 also provide useful replicates to K005 confirming the proteins identified.

The first two experiments, K003 and K004, were conducted in collaboration with Dr. Alex Jones from The Sainsbury Laboratory, Norwich, UK. The K003 and K004 experiments were composed of nuclear extracts from day2 (N2) and day8 (N8) cell cultures. Mass spectrometry (MS) analysis was undertaken using an LTQ ion trap mass spectrometer (Thermo Finnigan) with parameters set by Dr. Jones. Analysis of the K005 experiment was carried out in collaboration with Drs. Gerhard Saalbach and Mike Naldrett from the John Innes Centre Proteomics Facility, Norwich, UK. The K005 sample was composed of N2 and N6 and the MS analysis was done on an LTQ-Orbitrap XL™ hybrid mass spectrometer (Thermo Fisher Scientific) with parameters set by Dr. Saalbach. Table 3.1 summarises the above information.

Protein identifications presented in this chapter originate from iTRAQ™-labelled peptide mixtures, but protein quantification data will be analytically presented in Chapter 4.

Experiment name	Biological material	Mass spectrometer	Puspose of the experiment
K003	2-days-old and 8-days-old cultures	LTQ ion trap	identification of nuclear proteins; quantitative analysis between proliferating and quiescent cell populations
K004	2-days-old and 8-days-old cultures	LTQ ion trap	
K005	2-days-old and 6-days-old cultures	Orbitrap	

Table 3.1 Proteomics experiments.

Summary of proteomics experiments. Names are given in chronological order, with K003 being the earliest and K005 the latest.

In the first part of the chapter, I assess the similarity of crude chromatographic profiles from N2 and N6 samples in order to obtain an impression of the similarity of protein composition between the two samples. In the second part of this chapter, I deal with the protein identifications from SCX fractionated iTRAQ-derivatised tryptic peptides of N2 and N6 samples (before dealing separately in Chapter 4 with the quantitative aspect of the K005 experiment).

3.5.1 Label-free analysis on N2 and N6 samples

Chromatograms of unlabelled digested nuclear extracts of N2 and N6 samples for K005 experiment revealed peptide profiles as they elute from a reverse-phase column over a 120 min gradient (Figure 3.8). A gross level of inspection of the profiles from the two samples suggests that they are similar (Figure 3.8A). The high intensity peaks observed at the end of the gradient will consist of a mixture of partially undigested material and various common chemical contaminants. [Figure 3.8A; shaded area (I)]**.** Comparison of the elution profiles from N2 and N6 samples during the first 70 min showed that, even though they appeared to be similar, there were areas that differed considerably (Figure 3.8B; shaded areas (II), (III) and (IV)). These differences may reflect technical variation between the two HPLC runs (Prakash *et al*., 2007), sample preparation issues or more likely may reflect actual inherent differences between the two samples due to their

different biological backgrounds, given that the MS sample preparation conditions had been standardised.

One way of assessing the similarity of the two samples was to compare protein or peptide datasets generated after samples were run on the LTQ-Orbitrap XL™ mass spectrometer (the run was performed by Dr. Saalbach). Using Mascot software (Section 2.12.4.1) (www.matrixscience.com) 556 and 678 proteins were identified for the N2 and N6 samples, respectively. The threshold parameters applied by Scaffold (Section 2.12.4.1) were a 99% probability for peptide identification not being random, at least two peptides assigned to the identified protein, and an overall protein probability of 95%. From the above protein numbers, 283 proteins were common to both samples. This number of common proteins corresponds to 56% and 42% of the total number of proteins in N2 and N6 samples, respectively.

Running N2 and N6 unlabelled tryptic peptides separately, provided a trial run to examine the issues associated with label-free quantitative proteomics. This approach is considered viable, within the proteomics community, but is clearly software intensive and few of the world's leading groups including the Mann group have yet to adopt it over other methods.

An issue of great importance for label-free quantitative proteomics analysis is that elution times of the peptides, between the different samples, must be very similar if not exactly the same. For this reason, samples of interest are always run consecutively to avoid variations in ion intensities due to differences in the experimental setup (column properties, temperatures). Here, to initially assess the technique, common peptide pairs for the N2 and N6 samples were identified (using Mascot) and their elution times extracted using the Mascot Distiller tool (www.matrixscience.com/distiller.html) with the help of Dr. Gerhard Saalbach. Comparison between elution times gives an idea of the ease of using the label-free approach for protein quantification. Between the two consecutive runs of the N2 and N6 samples, the HPLC column was washed thoroughly to remove any traces of sample from the previous rus. Following washes, the column was re-equlibrated to its pre-run condition and was then used for loading the second digested protein sample. Since the samples were run consecutively and system setup variation was kept to a minimum, one would expect that elution times for the same peptide in different runs would be the same irrespective of which sample it came from.

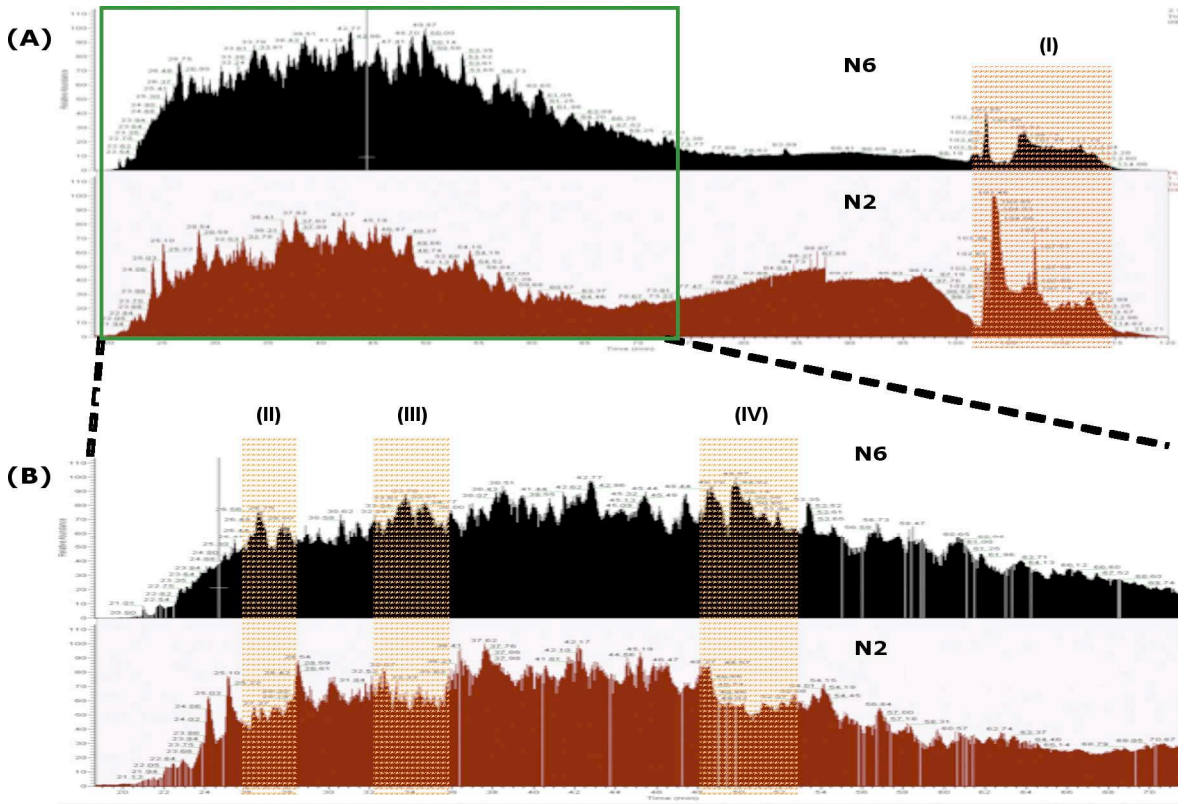


Figure 3.8 HPLC analysis of unlabelled digested N2 and N6 samples.

(A) Full scale chromatogram profiles of N6 (top) and N2 (bottom) samples.

(B) Magnification of the area marked by the green square in A. Shaded areas (I), (II), (III), IV) highlight differences observed between the two samples (see text for details)

From the 2855 and 2632 peptides identified in the N2 and N6 samples, respectively, only 586 were common between the two samples. Elution times for each of the common peptides were Log_2 transformed and each pair of values [Log_2(N2) and Log_2(N6)] were plotted as a scatter plot in order to determine the degree of correlation between the two elution times (Figure 3.9). As judged by the R-squared value (R^2=0.12), the correlation between two elution times was small, indicating considerable differences between the elution times for the majority of peptides present in both N2 and N6 samples, even though for a certain number the correlation was close to 1 (Figure 3.9; blue diagonal line). This result highlights the problem of technical reproducibility and software issues in label-free experiments, necessitating optimization of the system setup as well as the parameters of software used for extracting the respective elution times. This approach was judged as being too time-consuming, and it was suggested instead to use the iTRAQ stable isotope labelling technology, thereby substantially overcoming technical variation issues encountered in the label-free approach (see Chapter 4).

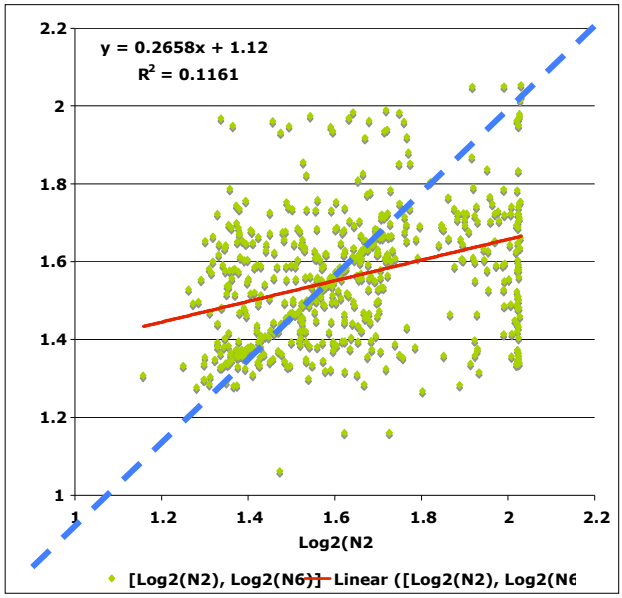


Figure 3.9. Correlation analysis of elution times in common peptides between N2 and N6 samples. The graph represents a scatter plot of log_2 transformed elution times of peptides present in N2 and N6 samples. The blue diagonal line indicates the area of the graph where the ratio between the Log_2(N2) and Log_2(N6) values are 1.

3.6 Analysis of the *Arabidopsis* nuclear proteome

3.6.1 Assessing variation between *Arabidopsis* nuclei preparations

Analysis of K003, K004 and K005 samples was performed at different times (Section 3.5) and each one was initially analysed separately, using different search and probability parameters for protein identification. As expected, the results of each analysis show common protein identifications but also substantial differences (data not shown). To clarify the search results, since the raw data from all experiments were not analysed using the same MS and software parameters, the raw spectral data from the experiments on samples K003, K004 and K005 were re-analysed all together using the proteomics software Scaffold (version 2.04.00; Proteome Software Inc.) with a common set of parameters (see Section 2.12.4.2). All samples gave >500 protein identifications at the ≥95% probability level (Table 3.2A) with K003, K004 and K005 experiments identifying 756, 536 and 659 proteins, respectively. Since these proteins were identified from purified nuclei, they are considered as *a priori* nuclear proteins, and hereafter I refer to them as "nuclear-putative" ($nuclear_{PU}$) proteins.

To assess how many proteins from each dataset were assigned a nuclear localisation in the literature or by *in silico* prediction, I queried a GO-annotated gene list from TAIR (TAIR8; www.arabidopsis.org) and SUBA-II, a subcellular protein localisation database (Heazlewood *et al*., 2007). The TAIR8 database was queried to identify proteins that have been previously experimentally annotated as nuclear, whereas SUBA-II was queried to identify proteins that contain a putative nuclear localisation signal (NLS) in their sequence and, thus, would be annotated as "putative nuclear" proteins. The SUBA-II database employs different subcellular localization predictors for assigning a putative localisation to a specific protein.

Proteins that had been experimentally observed to localise in the nucleus (using the TAIR8 database) were termed "nuclear-verified" ($nuclear_{VE}$), whereas those predicted to be nuclear localized by at least two predictors (within the SUBA-II database) were termed "nuclear-predicted" ($nuclear_{PR}$). The sum of $nuclear_{VE}$ + $nuclear_{PR}$ comprised the "nuclear-annotated" protein dataset ($nuclear_{AN}$). The K003 dataset contained a total of 756 $nuclear_{PU}$ proteins, 425 (57%) of which were assigned as $nuclear_{AN}$, 224 as $nuclear_{VE}$ and the rest 201 as $nuclear_{PR}$ (Table 3.2B). The K004 dataset contained a total of 536 $nuclear_{PU}$ proteins, 324 (61%) of which were assigned as $nuclear_{AN}$, with 183

$nuclear_{VE}$ and 141 $nuclear_{PR}$. The K005 dataset contained 659 $nuclear_{PU}$ proteins, of which 445 (67.5%) were assigned as $nuclear_{AN}$, with 247 being $nuclear_{VE}$ and 198 being $nuclear_{PR}$ (Table 3.2B). Thus, the K005 sample appears to be the most "nuclear protein-enriched" sample.

(A)

Experiment	Scaffold score		Total number of $nuclear_{PU}$
	protein probability of < 95%	protein probability of >=95%	
K003	157	756	756
K004	227	536	536
K005	150	659	659

(B)

Experiment	Annotation source		Total number ($nuclear_{AN}$)
	TAIR8 ($nuclear_{VE}$)	SUBA-II ($nuclear_{PR}$)	
K003	224	201	425 (56%)
K004	183	141	324 (60.5%)
K005	247	198	445 (67.5%)

Table 3.2 Number of proteins identified from three independent biological experiments.

(A) $Nuclear_{PU}$ proteins identified using the Scaffold software were categorised based on their Scaffold score, using the 95% confidence threshold for distinguishing between significant and not significant protein identification.

(B) Proteins from all three experiments, with protein score equal to or larger than 95%, were annotated for their subcellular localisation. The TAIR8 experimentally verified dataset ($nuclear_{VE}$) and SUBAII prediction tool ($nuclear_{PR}$) were used to determine the number of $nuclear_{AN}$ proteins. Percentages show the proportion of $nuclear_{AN}$ proteins in the total number of $nuclear_{PU}$ with a probability score of ≥ 95% (A, column 2).

A list of genes from the K003, K004 and K005 experiments of $nuclear_{PU}$ and $nuclear_{AN}$ proteins is given in Appendix-I.

Next, to assess the reproducibility of the nuclei isolation protocol, I asked how many nuclear$_{PU}$ proteins were in common between the three experiments. Firstly, I checked the extent to which the confidently identified nuclear$_{PU}$ proteins (Scaffold score ≥95%) overlapped between experiments (see the Venn diagram in Figure 3.12A). Overlap was also seen for the nuclear$_{AN}$ protein datasets as well (47%, 61.5% and 44% for the K003, K004 and K005, respectively; Figure 3.12B).

In the K004 dataset, the majority of nuclear$_{PU}$ or nuclear$_{AN}$ proteins are shared with either K003 or K005 datasets. Almost 90% of the nuclear$_{PU}$ proteins in the K004 dataset are shared with only K003, K005 or both (Figure 3.12A), whereas in the case of the nuclear$_{AN}$ protein dataset, the vast majority of proteins from the K004 experiment (89.2%) are shared by only K003, K005 or both experiments (Figure 3.12A). On the other hand, less overlap is observed between the K003 and K005 experiments. The proportion of proteins common between K003 and K005 datasets is a small fraction (13.5% and 15%, respectively) of the nuclear$_{PU}$ protein dataset for each of the two experiments (Figure 3.12A) and the same is true when only the nuclear$_{AN}$ proteins are considered (13.5% and 14.5%, respectively; Figure 3.12B).

In conclusion, the K005 dataset provides a high number of both putative nuclear proteins and nuclear-annotated proteins, indicating that this is likely to be the dataset of the highest quality.

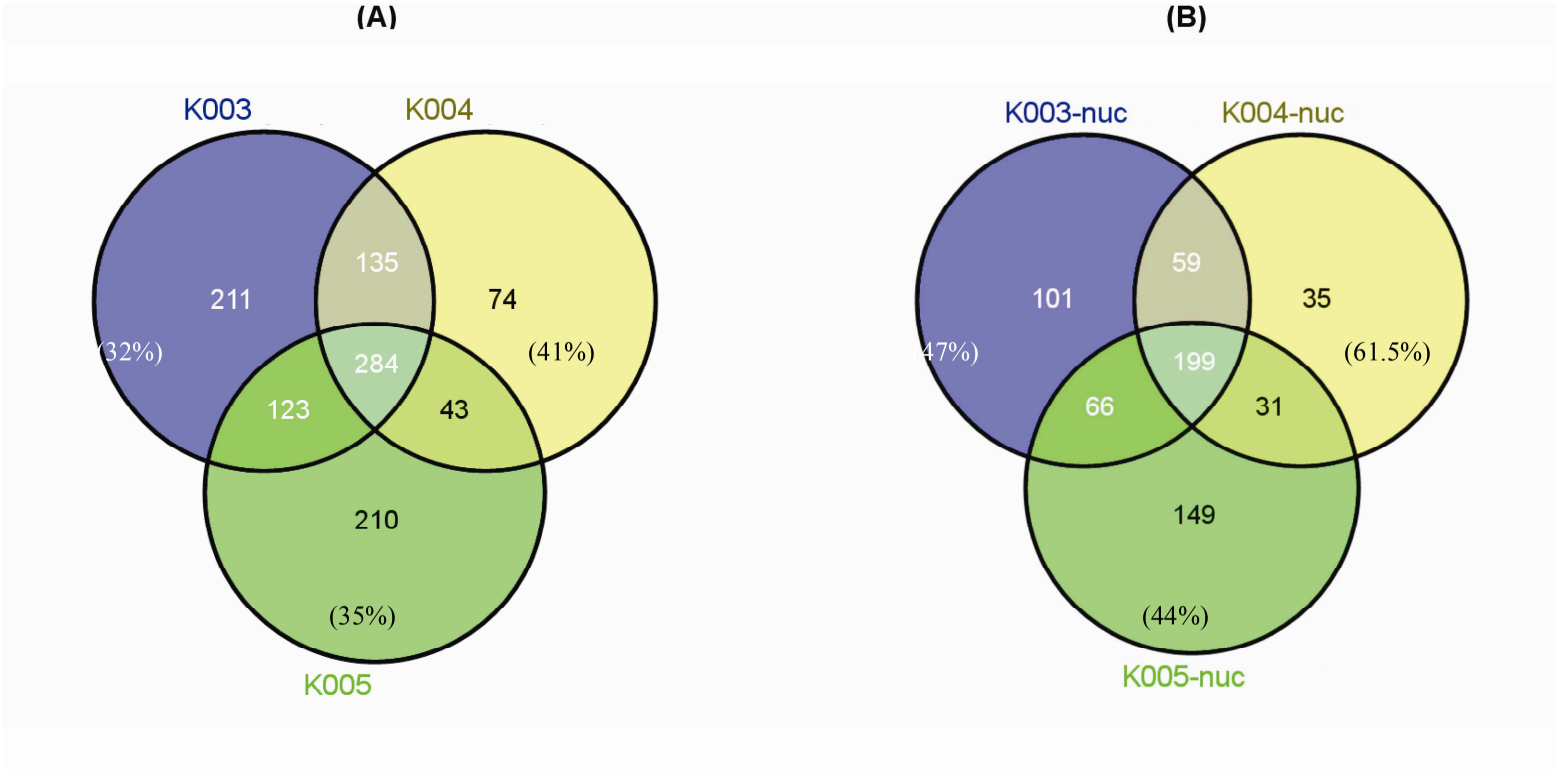


Figure 3.12 Overlap analysis of the K003, K004 and K005 experiments.

(A) All the nuclear$_{PU}$ proteins identified in each experiment were used for generation of the Venn diagram

(B) Only nuclear$_{AN}$ proteins from each experiment were used for the Venn diagram. Numbers shown in parentheses in (A) and (B) indicate the percentage of nuclear$_{PU}$ proteins in each experiment that belong to the group of common proteins.

3.7 Analysis of the nuclear proteome

3.7.1 Functional enrichment analysis of identified nuclear$_{PU}$ proteins

To assess which functional categories were over-represented across the nuclear$_{PU}$ datasets, I performed a gene ontology (GO) enrichment analysis. This analysis identifies enriched categories by comparing the number of proteins belonging, for example, to a functional category with the number of proteins belonging to the same category within a randomly generated protein dataset. The larger the former number is, as compared to the latter, the more significant the specific functional category will be. Graphical representation of the results and the relation between the enriched categories was achieved by using the BiNGO tool in the Cytoscape software (see Section 2.13). Usage of the "GOSlim_Plants" annotation category on the K003, K004 and K005 datasets grouped the genes into three GO categories; "cellular component", "molecular function" and "biological process".

Regarding the GO enrichment based on "biological process", over-fragmentation led to large numbers of processes that were not amenable to statistical analysis. For the above reason, I present the GO annotations for "molecular function" only. Figure 3.13(A-C) shows GO enriched categories of nuclear$_{PU}$ proteins from K003, K004 and K005 experiments.

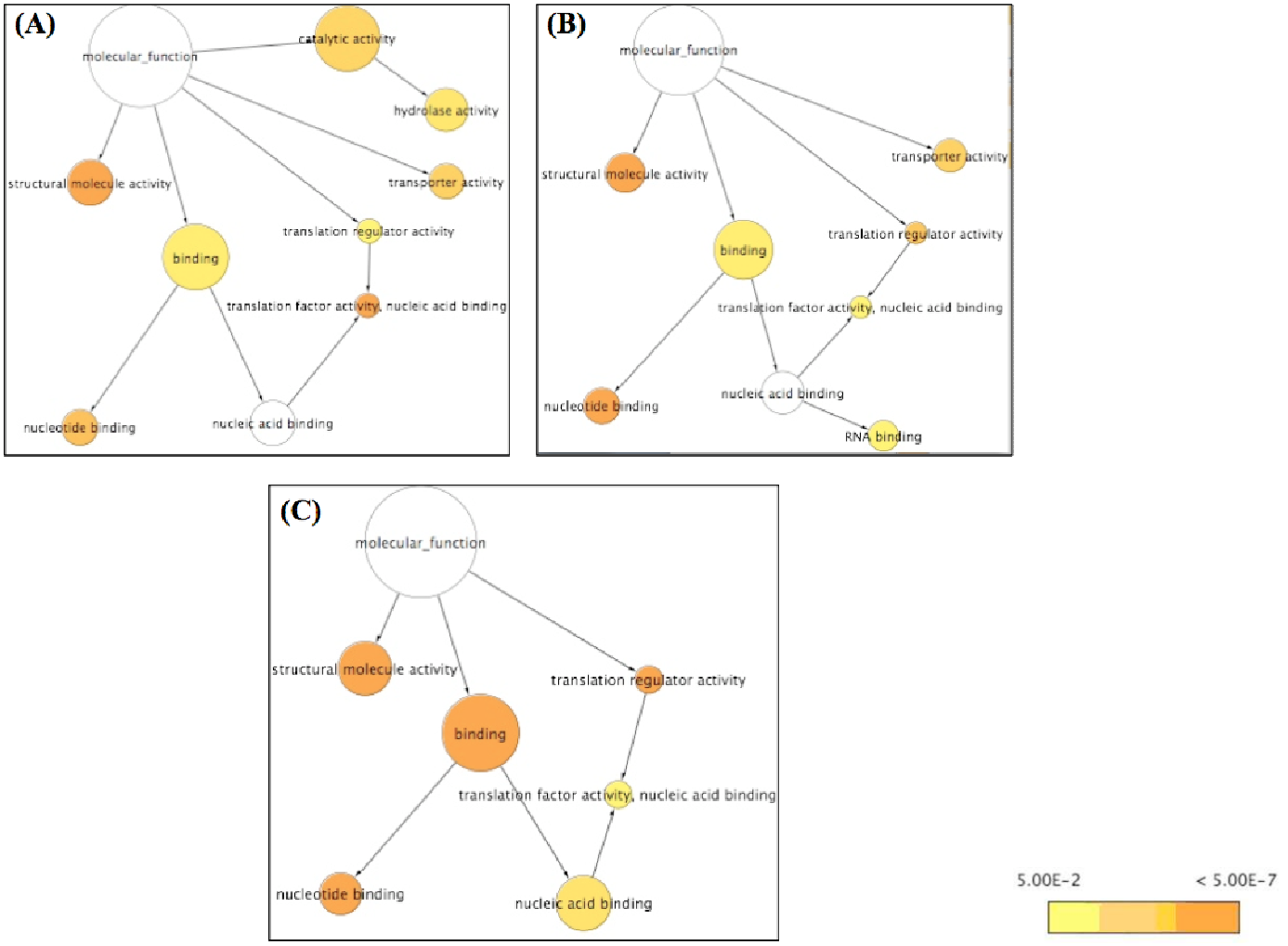
Generated networks for the three experiments showed the same overall structure with one major difference. The two additional nodes that are present at the K003-based network, grouping proteins involved in catalytic and transporter activities (Figure 3.13A) are gradually disappearing as we move towards the K005-based network (Figure 3.13B and C). This change could be due to the decrease at the number of proteins that belong to these functional categories, falling below the threshold necessary for obtaining an enriched GO category. The gradual disappearance of these GO categories coincides with the improvement of the nuclei isolation protocol (Section 3.2.2 and Table 3.1). Thus, disappearance of the specific GO categories probably reflects the increase at the purity of my nuclei preparations as could be seen in the case of nuclear$_{AN}$ proteins, as a portion of nuclear$_{PU}$ proteins (Table 3.2)

Figure 3.13. BiNGO enrichment analysis of nuclear$_{PU}$ protein datasets based on molecular function.

(A) - (C) Nuclear$_{PU}$ proteins from K003 (A), K004 (B) and K005 (C) experiments (Table 3.2B) were used for performing BiNGO enrichment analysis and generating the respective networks.

(D) Corrected *p*-values for over-represented GO categories in the nuclear$_{PU}$ protein datasets for K003, K004 and K005 experiments.

The output of the BiNGO analysis consists of a network with three major branches; one refers to the over-represented GO categories based on the cellular compartment, one on biological process and one on the molecular function. I present only the network based on the molecular function, for all the three nuclear protein groups. The colour of each circle represents the statistical significance for over-representation of the respective GO category, according to the scale, and the size of the circle is proportional to the number of genes that belong to the GO category.
The left end of the color scale indicates significance of a $p<0.01$ and the right end indicates significances of a $p<1x10^{-7}$. n.s.: not statistically significant.


(A)
molecular_function
catalytic activity
hydrolase activity
structural molecule activity
transporter activity
translation regulator activity
binding
translation factor activity, nucleic acid binding
nucleotide binding
nucleic acid binding
(B)
molecular_function
transporter activity
structural molecule activity
translation regulator activity
binding
translation factor activity, nucleic acid binding
nucleotide binding
nucleic acid binding
RNA binding
(C)
molecular_function
structural molecule activity
translation regulator activity
binding
translation factor activity, nucleic acid binding
nucleotide binding
nucleic acid binding
5.00E-2
< 5.00E-7

3.7.2 Functional categorisation of K005 nuclear$_{PU}$ protein dataset

Since the K005 sample appeared to be the most enriched in putative nuclear proteins among the three independent experiments (Section 3.7.1), I wanted to further refine the percentage distribution of functional GO categories across the whole dataset, without accounting for enrichment. I assigned all the identified nuclear$_{PU}$ proteins to a GO category using the TAIR8 database (Figure 3.14). Proteins belonging to the DNA/RNA binding GO category formed the largest group (19%) and included transcription factors, histones and histone-binding proteins, RNA binding and processing proteins. The second largest category (13%) comprised ribosomal proteins. The third GO category contained proteins involved in general protein binding (13%), whereas proteins of unknown function formed the fourth largest category (10%). Representative members of the top three categories, based on the number of unique peptides, are presented at Table 3.4.

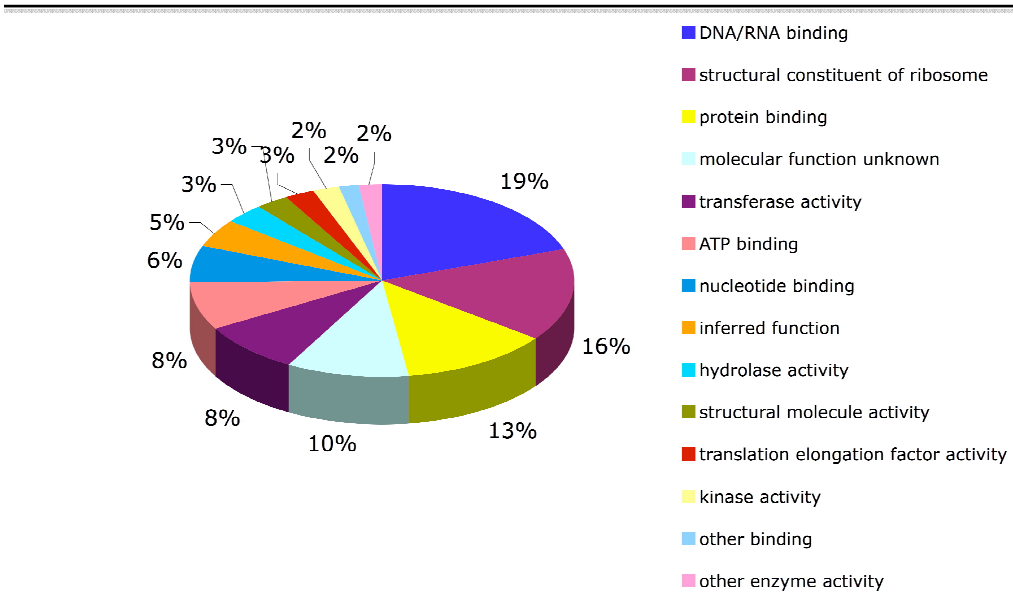


Figure 3.14. GO annotation of proteins identified in the K005 sample.

Functional annotation of nuclear$_{PU}$ proteins was done using the TAIR8 database. GO categories are showed at the right of the pie.

GO category	AGI	Gene Name	no. of unique peptides
DNA/RNA binding	AT1G06760	histone H1, putative	15
	AT1G48920	NUC-L1	15
	AT3G05060	SAR DNA-binding protein	14
	AT3G16810	PUMILIO24	14
	AT4G25630	FIB2	14
protein binding	AT5G46070	GTPase	20
	AT3G44110	PUMP1	11
	AT1G52360	WD40-binding	10
	AT2G02560	CAND1	10
	AT3G54110	ATJ3	10
structural molecule activity	AT3G09630	RPL4A	18
	AT2G17360	RPS4A	17
	AT1G43170	RP1	16
	AT4G34450	RPL6A	14
	AT2G01250	RPL7B	13

Table 3.4. Top three GO categories identified for the K005 protein dataset.
Proteins with the highest number of unique peptides identified in each GO category are presented. A full list of the protein groups in each category and unique peptides for each protein are given in Appendix-I

3.7.3 Nuclear$_{PU}$ proteins of unknown function

Annotation of nuclear proteins from the K005 group based on their molecular function led to the identification of 63 proteins of unknown function. As a first step in assigning putative functions to this group of proteins, I used the respective protein sequences to perform a homology analysis (BLAST) against the human proteome. 30 out of the 63 have a homologue in the human proteome. The remaining 33 proteins do not have a known homologue and are considered as "plant-specific" (Table 3.5 shows a subset of the protein dataset). From the 30 proteins with assigned putative function, 21 of them are nuclear in human cells and comprised of ribosomal proteins, nuclear pore proteins and nucleolar proteins (see Appendix-III for full list of unknown proteins and their respective human homologs).

The fact that almost 40% of the proteins are characterised as "plant-specific" leads to two suggestions. Either these proteins are involved in plant-specific nuclear processes or these proteins do perform known functions but their sequence has diverged from the rest of the eukaryotes to such a degree that no close homologues in other eukaryotes can be detected.

AGI	Protein Acc. No	Protein name	E-value
AT1G07170	Q7RTV0	PHF5A_HUMAN	2.00E-061
AT1G10490	plant specific		
AT1G15420	Q15061	WDR43_HUMAN	4.00E-008
AT1G18850	plant specific		
AT1G30240	Q8IZL8	PELP1_HUMAN	3.00E-008
AT1G31660	Q13895	BYST_HUMAN	2.00E-090
AT1G32130	Q96ST2	IWS1_HUMAN	3.00E-026
AT1G63810	Q9H6R4	NOL6_HUMAN	E-113
AT1G69070	P78316	NOP14_HUMAN	5.00E-058
AT1G70770	Q6NUQ4	TM214_HUMAN	9.00E-013
AT2G15860	plant-specific		
AT2G18220	Q9Y3T9	NOC2L_HUMAN	6.00E-073
AT2G26680	plant specific		
AT3G10650	plant specific		
AT3G14120	P57740	NU107_HUMA 1e-23	
AT3G18790	Q9ULR0	ISY1_HUMAN	2.00E-057
AT3G22520	B7ZMJ9	TACC2_HUMAN	8.00E-004
AT3G49720	plant specific		
AT3G57940	plant specific		

Table 3.5 Homology analyses of proteins with unknown functions.

Arabidopsis proteins annotated as "unknown" from the TAIR8 database, were used for BLAST analysis of the human proteome to identify homologous sequences. Where no sequence homologue was identified, the respective *Arabidopsis* protein was annotated tentatively as "plant-specific". This table presents a subset of the proteins. A complete list of proteins can be found in Appendix-IV.

3.7.4 Physical properties of proteins in the K005 protein sample

A major determinant for the activity of many proteins is the pH of its immediate environment. The isoelectric point (p*I*) of a protein is the pH value at which the protein molecule has a net charge of zero. When the pH of a protein's environment is very close to its p*I* then the protein is least soluble, in some cases leading to reduction of its activity and to protein aggregation (Schwartz *et al*., 2001). p*I* value has been proposed to be a good indicator of protein subcellular localisation in *Drosophila* and mammalian cells (Schwartz, *et al*., 2001; Sutherland *et al*., 2001). In the mammalian nucleus, different subnuclear compartments appear to harbour proteins of a certain p*I* range, with the most basic proteins being the nucleolar and RNA-binding proteins, whereas chromatin-binding and nucleoplasmic proteins were only slightly basic (Sutherland *et al*., 2001). The above observations for mammalian cells suggest that there is dependence between the p*I* value of a protein and its subcellular localisation. Equivalent studies have not been undertaken in plants. Therefore, using the large nuclear protein dataset from K005 experiment, I checked whether there is a correlation between nuclear localisation of a protein and its corresponding p*I* value.

3.7.4.1 Determination of p*I* distribution of proteins from *Arabidopsis* nuclei

Since the nuclear$_{PU}$ dataset originates from nuclei-enriched preparations, the distribution of p*I* values should differ from the respective p*I* distributions of proteins that do not reside in the nucleus (Sutherland *et al*., 2001). This comparison could also serve as another test assessing for the under-representation of non-nuclear proteins in the nuclear$_{PU}$ dataset. Thus, I compared the nuclear$_{PU}$ dataset with proteins annotated as mitochondrial or chloroplast in the TAIR8 database (www.*arabidopsis*.org). Finally, I asked whether proteins from the published *Arabidopsis* nucleolar proteome exhibited a different p*I* value distribution compared to those from the whole nucleus.

When compared to the mitochondrial/chloroplast proteins, nuclear$_{PU}$ proteins showed an enrichment for very acidic proteins at the *pI* range up to 3.2 (Figure 3.15A), with 25% (163/657) and 11% (300/2689) of the total number of nuclear$_{PU}$ and mitochondrial/chloroplast proteins occupying this low end of p*I* range, respectively. On the other end of the spectrum, a considerable enrichment could be observed for the

nuclear$_{PU}$ proteins at p*I* range of 10.6 – 12.0 [13.5% (103/657) and 6.5% (173/2689) for the nuclear$_{PU}$ and mitochondrial/chloroplastic proteins, respectively]. Even though nuclear$_{PU}$ proteins were enriched at both ends of the pH spectrum, mitochondrial and chloroplast proteins were enriched between p*I* 7.0 and 9.0, reaching 33.5% (956/2689) of the total number of proteins, as opposed to only 18% (120/659) for the nuclear$_{PU}$ protein dataset.

p*I* distribution among the nuclear$_{PU}$ and nucleolar datasets exhibited the same trend (Figure 3.15B), even though proteins with basic p*I* values were preferentially enriched in the nucleolar dataset. Also, a sharp decrease could be seen for nucleolar proteins with p*I* values between 5 and 9.6 (less that 2% in each p*I* value group) that was interrupted by an increase in the number of proteins with p*I* values around 6.2. Localized enrichment at certain p*I* values for nucleolar proteins suggests that the nucleolus might be locally specialised. It will be interesting to determine if proteins sharing a particular p*I* are more or less likely to co-localise.

The above analysis suggests that proteins localised to specific cellular compartments, at least in the nucleus and nucleolus, tend to have p*I* values that cluster around distinct pH ranges and that these ranges do not overlap with the physiological pH of the respective organelle.

Figure 3.15. Comparison of p*I* value distribution among nuclear$_{PU}$, nucleolar and mitochondrial/chroloplastic protein datasets in *Arabidopsis*.

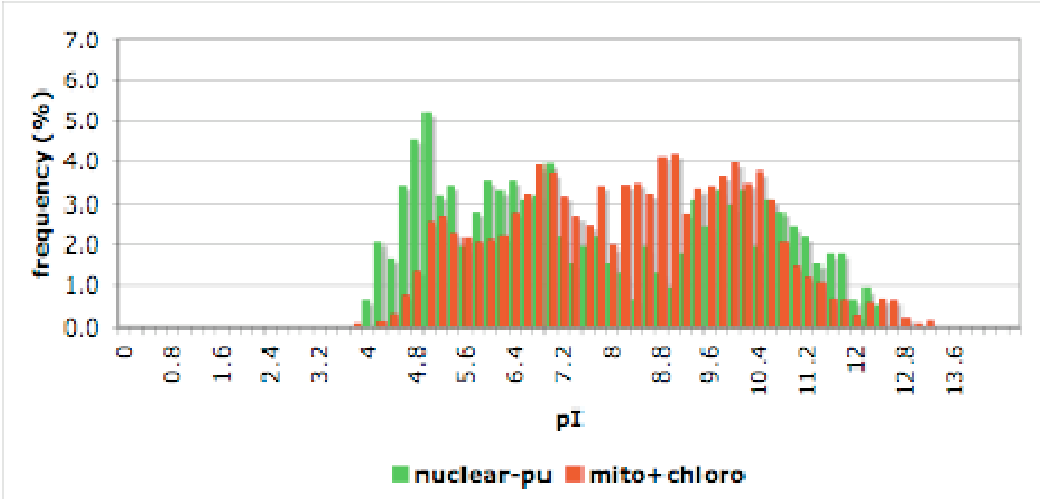
(A) Distribution of p*I* values between the nuclear$_{PU}$ and mitochondrial/chloroplast proteins

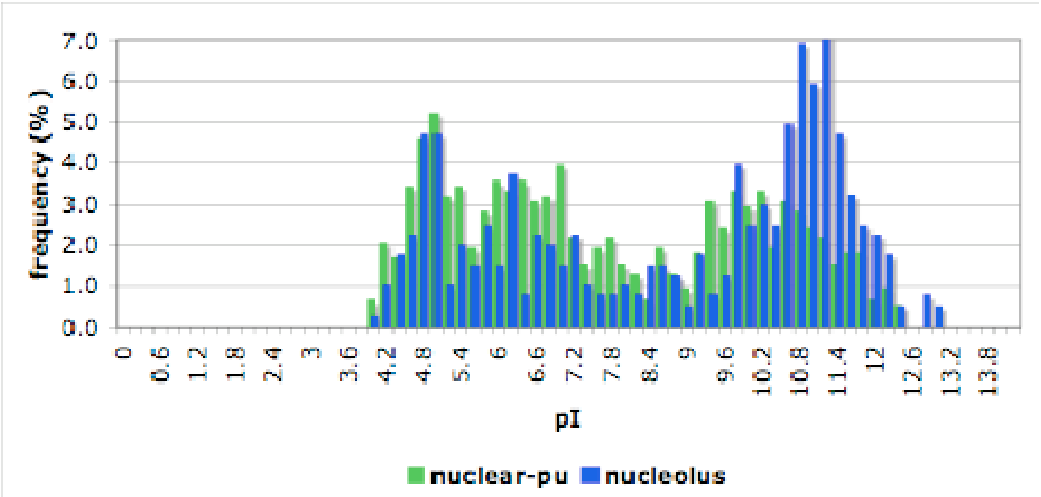
p*I* values of proteins identified in the K005 experiment were compared with p*I* values from a dataset of mitochondrial- and chloroplast-annotated proteins obtained from TAIR8 (www.*arabidopsis*.org)

(B) Distribution of p*I* values between the nuclear$_{PU}$ and nucleolar proteins

p*I* values of proteins identified in the K005 experiment were compared with p*I* values from a published nucleolar protein dataset (Pendle *et al*., 2005)

p*I* values were obtained using the "Bulk Data Retrieval" tool from TAIR (www.*arabidopsis*.org). Proteins in each dataset were grouped around a specific p*I* value and numbers within each group were converted into a percentage of the total number of proteins in each dataset.


frequency (%)
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
0
0.8
1.6
2.4
3.2
4
4.8
5.6
6.4
7.2
8
8.8
9.6
10.4
11.2
12
12.8
13.6
pI
nuclear-pu
mito+chloro


frequency (%)
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
0
0.6
1.2
1.8
2.4
3
3.6
4.2
4.8
5.4
6
6.6
7.2
7.8
8.4
9
9.6
10.2
10.8
11.4
12
12.6
13.2
13.8
pI
nuclear-pu
nucleolus

3.8 Discussion

In this chapter I have provided an initial description of the nuclear proteome of *Arabidopsis* cell cultures using subcellular fractionation and MS-based proteomics techniques. In the following sections I consider technical issues associated with nuclear isolation and technical reproducibility.

3.8.1 Protocol optimisation for nuclei isolation from proliferating and quiescent cell populations

Protocols on isolating nuclei from plant cell cultures have been described in the literature for *Arabidopsis*, *Medicago* (*Medicago transcatula*), chickpea (*Cucumis arietinum*) and rice (*Oryza sativa*) (Jones *et al*., 2009; Pendle *et al*., 2005; Tan *et al*., 2007). However, in all those studies, nuclei were isolated from suspension cells in exponential growth only. Here I report for the first time the isolation of nuclei from stationary phase cell cultures. This will allow further studies to be conducted on determining important changes that take place within the nucleus as cells pass from an active state of high cell division rates to an arrested state with minimum proliferation. Further improvement of nuclei isolation may be required for older cells.

3.8.2 Morphological alterations in cell and nuclei size associated with growth status of the cell culture

Nuclear size is dependent on the size and age of cells. Confluent (growth-arrested) HeLa cell cultures contain smaller nuclei than proliferating cells and this is associated with low metabolic activity and a reduction of nuclear pore sizes (Feldherr and Akin, 1990). Also, reduction of nucleus size has been observed in embryonic mouse brain cell monolayers due to contact-dependent growth inhibition (Zimmermann *et al*., 1976). The reduction of nuclear size of cells as they progress from proliferation to quiescence,

suggests that probably a similar "space-limiting" response influences the final nuclear size in *Arabidopsis*.

In plants, nucleolar size and organisation depends on the physiological status of the cell and on cell proliferation (Junera *et al*., 1995; Medina *et al*., 2000). In onion cells, nucleoli in proliferating meristematic cells are characterised by large size and abundant granular component. In contrast, those in non-meristematic cells are composed exclusively of a dense fibrillar component and they are almost 4-times smaller (Gonzalez-Camacho and Medina, 2005). Nuclear and nucleolar size of *Arabidopsis* suspension cells also varies in a similar manner making this suitable material for studying nuclear behaviour.

In yeast cells a constant karyoplasmic ratio is maintained that restrains the nuclear volume at 8% of the total cellular volume, even when cells are actively growing (Umen, 2005; Jorgesen *et al*., 2007; Neumman and Nushe, 2007). Controlled displacement of nuclei in multinucleated yeast cells by centrifugation showed that nuclei surrounded by larger cytoplasmic volumes gradually increased their size up to the upper threshold value (N/C at 0.08), (Neumann and Nushe, 2007). These experiments provide strong evidence that the size of the nucleus is controlled by unknown cytoplasmic factors, though not yet identified. In plants, the cell size is often related to the ploidy level of the nucleus, indicating that there may be signals from the nucleus to the rest of the cell to control the final cell size (Sugimoto-Shirasu and Roberts, 2003). However, no data have so far become available about cell size control under cell culture conditions.

The stable karyoplasmic ratio observed for suspension cells supports a link between cell size and nucleus size control, although the factors responsible remain unknown. Karyoplasmic ratio was first described more than a century ago by R. Hertwig (Wilson, 1925) and since then constant karyoplasmic ratio was demonstrated across nearly all organisms from bacteria (Weart *et al*., 2007) to animals (Jorgensen and Tyers, 2004) and plants (Melarango *et al*., 1993). Whether a cytoplasm-based mechanism for maintaining a constant karyoplasmic ratio exists in *Arabidopsis*, like in yeast, is also uncertain and further experiments are needed. If such mechanism does exists, then one might predict that mutants affected in nucleocytoplasmic transport, like a mutation in the exportin ortholog *HASTY* (Bollman *et al*., 2003), would also affect final nucleus size as cells progress from proliferation to quiescence.

3.8.3 Reproducibility issues concerning LC/MS/MS-based proteomics analysis

Reproducibility in proteomics experiments remains a serious concern (Aebersold, 2009). Recently, directed experiments that aimed at identifying the source of variation in proteomics experiments have analysed pre-defined protein samples multiple times, (Truck *et al*., 2007; Bell *et al*., 2009). However, the use of relatively simple protein mixtures does not mirror the complexity of the highly complex protein mixtures produced by organelle proteomes. Therefore, I will focus on issues related to run-to-run reproducibility problems in the same lab when using the same protocols and instrumentation. Additional variation arises from the differences in the biological samples that may or may not be related to the experimental conditions, not all of which can be controlled. Such variation can be combated by optimization and standardisation of sample preparation. In my case, optimization of the nuclei isolation protocol allowed me to finally obtain good quality nuclei, as judged by biochemical methods and MS-based analysis, but time did not permit standardisation.

Technical variation in the MS analysis arises at many points. Peptide mixture separation using liquid chromatography is one such source. HPLC columns often retain trace material from previous separations, the so-called “carryover” effect, resulting in false protein identification and false peptide ion intensities that will lead to incorrect quantification. Carryover is attributed mainly to hydrophobic peptides binding strongly to the stationary phase of the column or peptide aggregates sitting in the frits of the column amongst other things (Kirkland and Kim, 1995; Suelter and Deluca, 1983). Nevertheless carryover was not detected during my analysis since multiple washes of the column ensured that no residual material from previous runs was present. In case that multiple washes are not sufficient to eliminate the problem, long washes with isopropanol and trifluoroacetic acid (TFA) (Wang and Hanash, 2008) or with trifluoroethanol (TFE) (Mitulovic *et al*., 2009) can help. TFE is used for dissolving proteins prior to enzymatic digestion and the use of polyfluorinated alcohol is known to prolong the life of reversed phase columns (Bidlindmeyer and Wang, 2006).

At Section 3.5.1 I examined the similarities between the HPLC profiles of N2 and N6 samples and one of the observations was a difference at the elution times for common peptides between the two samples. Such tests are useful for assessing the run-to-run reproducibility of the HPLC system, since not optimal performance of HPLC can

directly affect estimations of ion intensities and, consequently, protein identification and quantitative information. Causes of the observed variability may include a shift in the elution gradient or an auto-sampler error or ageing buffers. Gradient shifts often occur when the equipment is adjusted (i.e replacement of capillary tubing) so can be easily eliminated with care. The auto-sampler needle – the needle that loads the tryptic mixture from the sample vial into the HPLC system – can introduce shifts due to air in the system, needle clogging or leaky valves (Dr. Gerhard Saalbach, JIC, personal communication) that result in variable elution times between consecutive MS runs. Although operator experience is an important factor in achieving reproducible experimental outputs, free commercial software (e.g. ChromEval) is available to assess HPLC performance and reproducibility (Sigmon *et al*., 2010).

3.8.4 Size of the *Arabidopsis* nuclear proteome

The known and predicted nuclear proteome in *Arabidopsis* comprises about 7,650 proteins (28% of the total protein-coding genes (http://gpcr.biocomp.unibo.it/esldb/)) and is compared to other eukaryotes in Table 3.6. Around 97% of the *Arabidopsis* annotations are based on homology or *in silico* assignment of nuclear localisation, highlighting the fundamental lack of experimental evidence for proteins residing in the *Arabidopsis* nucleus. To that end, the MS-based analysis of the *Arabidopsis* nuclear proteome presented here has increased the number of nuclear-localised proteins by at least 2-fold (see Table 1.2), providing a high-quality dataset of 654 proteins that could be used in a high-throughput screening assay using epitope-tagged proteins, for example GFP-tagged proteins (Koroleva *et al*., 2005), for *in vivo* validation.

However, this represents less than one tenth the expected number of protein identifications (Table 3.6). To increase this number, it may be necessary to undertake further biological fractionation of the organelle (for example as described by Pendle *et al*., 2005) to provide the MS with simpler mixtures. If further fractionation of the biological material is not technically feasible, then an increase in protein identification can be achieved at the level of peptide mixture pre-fractionation.. An increase in the number of fractions collected by the SCX column, combined with a shallower gradient,

would increase the number of peptides detected by the mass spectrometer and, thus, and increase the resolution power in protein identification.

Organism	Nuclear-localized proteins	Total number of protein-coding genes
Homo sapiens	12,360 (45%)[(a)]	26,383
Mus musculus	6,820 (18.7%)	36,536
Caenorhabditis elegans	4,733 (20%)	23,901
Saccharomyces cerevisiae	1,717 (35%)	4,883
Arabidopsis thaliana	7,650 (28%)	27,380

Table 3.6 Size of the nuclear proteome in different eukaryotes.

The actual number of nuclear-localized proteins shown in the second column represents the sum of proteins assigned a nuclear localisation by experimental validation or by homology-based analysis or by nuclear localization signal (NLS) prediction analysis, as analysed by the eSLDB database (http://gpcr.biocomp.unibo.it/esldb/).

[(a)] Numbers in the parentheses indicate the percentage of the nuclear proteome as a fraction of the total number of protein-coding genes in different eukaryotes.

3.8.5 Quality of the identified nuclear proteome

The K005 experiment was of relatively high quality as judged by microscopy, immunoblotting and the number of nuclear proteins identified. Also, the BiNGO tool indicated a significant enrichment for proteins with a nuclear-related function.

However, how to identify and eliminate contaminating (non-resident) proteins remains a problem for organellar proteomic analysis. Sometimes suspected contaminants are confirmed as actual residents and provide new insights into organelle function.

Examples include the exon-junction complex machinery in nucleoli (Andersen *et al*., 2005; Pendle *et al*., 2005) and the presence of cytokinesis-related proteins in ER (Skop *et al*., 2004).

My nuclear proteomics dataset contained proteins annotated as ER-localised, cytoplasmic or mitochondrial (Figure 3.14). Ideally these should be tested using an independent method such as GFP-based live cell imaging (Koroleva et al., 2005). In the meantime, the simplest explanation for the presence of such proteins is physical continuity between the nuclear outer membrane and these other compartments, particularly the ER, leading to contamination. Other explanations involve protein shuttling between compartments (Fricker *et al*., 1997) or localisation to multiple compartments. Indeed, several of the identified proteins localise in both nucleus and cytosol, [i.e. eIF2, ENOC, LOS2 and CDKA;1 (Koroleva *et al*., 2005; Prabhakar *et al*., 2009; Lee *et al*., 2002; Weingartner *et al*., 2001)]. In highly dynamic compartments, many proteins can be shared across several compartments. An example is the secretory system where over 500 proteins were simultaneously assigned to Golgi apparatus, ER, vacuolar membrane, plasma membrane or mitochondria and plastids (Dunkley *et al*., 2006). Moreover, elaborate statistical analysis allowed the authors to distinguish permanent residents of a particular organelle from contaminants and/or transient protein residents.

3.8.6 Functional analysis of the Arabidopsis nuclear proteome

The nuclear protein dataset extends the list of known plant nuclear proteins as identified by mass spectrometry (Bae *et al*., 2003, Repetto *et al*., 2008, Gallardo *et al*., 2007; Pandey *et al*., 2006). In the next section I will discuss some of the functional categories identified and how members of these categories fit to the current knowledge of nuclear biology.

3.8.6.1 Proteomics analysis identify key functional features of the nucleus

Proteins involved in ribosome biogenesis and translation

Identification of ribosomal proteins as the majority component in the "structural activity" category mirrors findings in other nuclear proteomics studies (Repetto *et al*., 2008; Pendle *et al*., 2005; Calikowski *et al*., 2003; Dellaire *et al*., 2003; Bae *et al*.,

2003). The only exception is a recent study on the identification of nuclear phosphoproteins from Arabidopsis cell cultures (Jones *et al*., 2009). Two importin-alpha protein transporters were also identified that are responsible for importing mature ribosomal protein into the nucleoplasm to produce mature 40S and 60S ribosomal subunits (Jakel and Gorlich, 1998). Activity of importins is facilitated by the presence of dynamin-like GTPases in human cells (King *et al*., 2004). Six dynamin-like proteins were identified in my nuclear preparations supporting the idea that the same mechanism may function in plants.

Eukaryotic translation initiation factors (eIFs), responsible for loading mature mRNAs onto the ribosome during the process of translation in the cytoplasm (Pestova *et al*, 2001), are also represented in the nuclear proteome. EIFs were found to localise in plant and mammalian nuclei (Bush *et al*., 2009; Repetto *et al*., 2008; Pendle *et al*., 2005; Khan and Komatsu, 2004; Calikowski *et al*., 2003; Andersen *et al*., 2002; Scherl *et al*., 2002) where they provide a proofreading mechanism for detecting aberrant pre-mRNAs and marking them for degradation via nonsense-mediated decay (NMD; Dreyfuss *et al*., 2002; Maquat, 2004).

Proteins involved in mRNA processing and transport

Many proteins of the mRNA processing machinery were identified (see Appendix-I for full list of proteins), including splicing factors (SR1 and SCL30), core components of spliceosome (U1A, U1, U2a) and components of the exon-junction complex (MAGOH, Y14; Dreyfuss *et al*., 2002; Le Hir *et al*., 2003). MAGOH and Y14 were also identified in *Medicago* nuclei (Repetto *et al*., 2008) and *Arabidopsis* nucleoli (Pendle *et al*., 2005). MAGOH colocalises with CDKC2 protein kinase in the nucleus when transiently expressed in *Arabidopsis* cell cultures (see Chapter 5) and CDKC2 also colocalises with and alters the distribution of splicing factors in *Arabidopsis* cell cultures (Chapter 5; Kitsios *et al*., 2008).

Nuclear proteins with unknown molecular functions

Proteins with no known function or "Unknown" were also identified. A subgroup of these was "plant-specific" (see Table 3.5). Additional information on such proteins

might reveal novel insights into the structure, organisation or function of the plant nucleus.

3.9 Future perspectives

Although 284 proteins were reproducibly identified across the three experiments (K003, K004 and K005) (Figure 3.12), they cannot be considered the result of biological replicates. These experiments represent protocol optimization procedures. Ideally, at least three biological replicates and three technical replicates for each biological sample are required for distinguishing between possible biological and technical variation (Prakash *et al*., 2007) and identify a reproducible set of proteins for experimental validation.

With a group of nuclear proteins confidently identified, these could be used to identify novel peptide motifs necessary for nuclear localisation and, thus, improve existing prediction tools such as Nucleo (Hawkins *et al*., 2007) and PredictNLS (Cokol *et al*., 2000). Such a novel protein set would also be invaluable for the generation of protein networks using tools such as IPA (Deighton *et al*., 2010), PPI Spider (Antonov *et al*., 2009) and SNOW (Minguez *et al*, 2009). Predicting direct or indirect interactions between proteins might allow us to design and test new hypotheses regarding nuclear function.

Chapter 4 - Quantitative proteomics analysis of the *Arabidopsis* nucleus

4.1 Introduction

4.1.1 Aims

The aim of this chapter is to compare the nuclear proteome at 2 defined stages in the growth cycle of *Arabidopsis* cells, namely day2 and day6 after sub-culture. At day2 the cultures are actively proliferating whereas at day6 the cultures have entered the stationary phase. An iTRAQ-based quantitative MS-analysis of nuclear extracts was used to compare relative levels of different proteins, aiming to shed light on the nuclear processes affected as cells move from proliferation to quiescence.

4.2 iTRAQ labelling of *Arabidopsis* nuclear extracts and pre-fractionation using LC

iTRAQ labelling for the K003, K004 and K005 samples was assessed manually, by identifying the proportion of identified peptides that carried the iTRAQ label (Table 4.1). The first two experiments had a very poor labelling whereas in K005, 98% of the peptides were labelled. Poor labelling in K003 and K004 experiments was due to the use of Tris (0.1M) buffer in one step during protein extraction from the purified nuclei. Tris contains a free primary amine that reacts easily with iTRAQ reagents leading to quenching and, thus, inefficient peptide labelling. Replacing Tris with HEPES buffer and being careful to use a protein precipitation method not containing ammonium ions (Section 2.10.1.2), resulted in almost 100% iTRAQ labelling of digested peptides from the nuclear extract of the K005 sample.

The peptide mixture was pre-fractionated on a PolySULFOETHYL A^{TM} strong cation exchange (SCX) column and selected fractions were run on the LTQ-Orbitrap XL™ mass spectrometer (see Section 2.12.3 for technical details). The SCX column chromatogram and applied salt gradient are shown in Figure 4.1. During the HPLC, weakly binding peptides are eluted first from the column, followed by a low-intensity 4 min peak. Immediately after this peak, a larger peak appears whose elution time window is 14 min and contains the majority of the eluted peptides. Following this is a

group of triple- or quadrupe-charged peptides. Finally, the high salt wash elutes any partially undigested material (Figure 4.1). Fractions from the 14-minute window, that contained most of the material (Figure 4.1A; double-headed arrow), were used for downstream MS-based analysis. Statistical parameters applied for protein identification are given in Section 2.12.4.2.

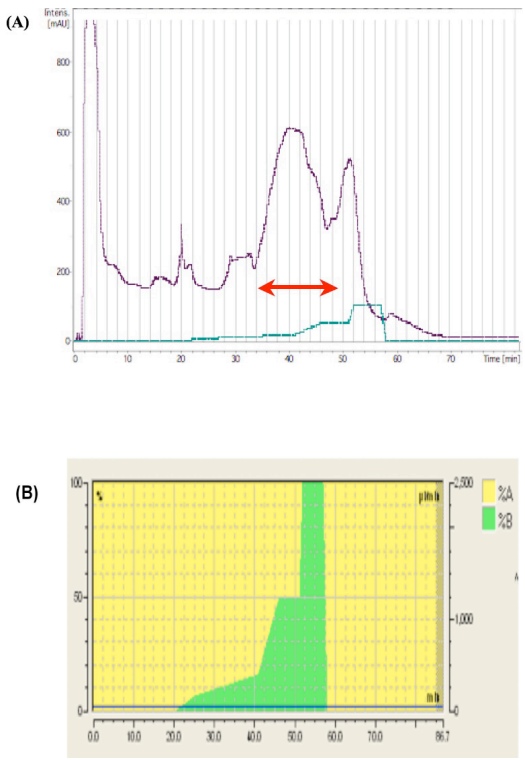


Figure 4.1 Fractionation of trypsin-digested nuclear proteins

Elution profile of K005 peptides over 85 min, under the indicated conditions, using a polysulfoethyl A SCX column.

(A) A chromatogram of absorbance at 214nm versus time of elution**.** Double-headed arrow indicates the 13 fractions that were selected and used for subsequent proteomics analysis

(B) The salt gradient used for elution is depicted. Flow rate was 50 µL/min and the salt gradient was generated using buffer A [10mM KH_2PO_4, 20% (v/v) acetonitrile, pH 2.7] and buffer B (buffer A plus 1M KCl).

	K003	K004	K005
Total no. of peptides	7957	9673	8970
iTRAQ-labelled peptides	1081	1230	8823
Percentage (%) of labelled peptides	13.59	12.72	98.36

Table 4.1 Efficiency of peptide labelling with iTRAQ reagents across all the proteomics experiments

4.3 Dynamics of the nuclear proteome

Before starting presenting my results on the quantitative analysis of nuclear extracts from proliferative and stationary cell populations, I would like to clarify the use of nomenclature about changes at protein levels between the two states. For describing my results I used the terms of "up-regulated" or "down-regulated" for proteins that increase or decrease at their abundance, respectively, as they move from one state to the other. The use of these terms is quite widespread in the proteomics community as can be seen in highly cited proteomics journals like the "Journal of Proteome Research" and "Molecular and Cellular Proteomics" (a search using the Endnote program retrieved 40 and 120 abstracts, respectively, containing the terms "up-regulated" or "down-regulated", referring to protein levels in quantitative proteomics studies). Change of the abundance of a protein detected during the MS analysis, reflects a regulatory mechanism functioning either at the level of gene expression or protein level (i.e. differential gene expression, differential protein stability, feedback mechanism). Therefore, the use of terms "up-regulated" or "down-regulated" cannot be considered as misleading, since they denote a mode of regulation, even though not identifiable by quantitative MS analysis.

Of the 645 proteins identified with high confidence (>=95% of confidence and at least two peptides assigned) by Scaffold Q+ software, 616 were labelled with the iTRAQ reagent and 589 had a ratio assigned by Scaffold Q+. The iTRAQ-115 and iTRAQ-117

reagents were used for protein quantification and iTRAQ-115 values were used as the reference for calculating the final protein ratio. Figure 4.2 shows the MS/MS spectrum of the parent peptide "GIEFVPLHVK" belonging to histone deacetylase 2A (HD2A). The lower left end of the same spectrum shows the two reporter ions at 115 and 117Da. The intensities of these reporter ions are used to calculate the Log_2 ratio of the specific spectrum for the respective peptide. Generally, Scaffold Q+ calculates the protein Log_2 ratio by first calculating the media Log_2 ratios of individual peptides, based on the reporter ion intensities of the assigned spectra, and then these media Log_2 ratios are used to obtain the median Log_2 ratio for the respective protein. Figure 4.3 gives an example for the calculation of Log_2 ratio of HD2A.

Since the more recent Scaffold Q+ version allows protein quantification within the interface, protein identification was repeated along with protein quantification for the K005 dataset. Comparison of the 589 quantified proteins with the 654 $nuclear_{PU}$ proteins previously identified with the old version of Scaffold (Table 5.2), showed that 541 proteins were in common between the two datasets and had a quantification ratio assigned (see Appendix- II for full list of proteins and their respective ratios). This group of 541 proteins was used for downstream analysis.

Variation in protein ratios, determined using iTRAQ reagents, in replicate experiments is accepted to be at the range of 20% (Gan *et al*., 2007). Thus, various studies have considered a fold change higher than 1.2 as statistically significant for designating a protein as "up-regulated" or "down-regulated" between two samples (Kolla *et al*., 2010; Martin *et al*., 2008; Griffiths *et al*., 2007). Although the current study lacks biological replicates, I applied a fold change of 1.2, to designate proteins as "up-" or "down-regulated". Clearly some designations will be incorrect but with this caveat, I carried on with my analysis, designating proteins as "up-" or "down-regulated" using the day2 dataset as a reference.

The general distribution of log_2-transformed protein ratios for the 541 quantified proteins is given in Figure 4.4. Relative to the day2 reference point, the amount of about 30% of the quantified proteins increased by 1.2 fold or greater in the quiescent cells, 37% did not change and 34% decreased in the quiescent cells (Figure 4.4).

Figure 4.2 Tandem MS (MS/MS) spectrum of the doubly charged GIEFVPLHVK parent ion belonging to the HD2A protein

The upper image shows the MS/MS spectrum profile and the b- and y-ions automatically assigned by the Scaffold Q+ software. On the top of the spectrum the inferred amino acids are given, whereas the amino-terminus and lysines of the peptide are labelled by the iTRAQ parent ion. At the bottom of the figure, a close-up of the shaded region of the spectrum is shown where the reporter ions at 115 and 117Da are detected. m/z: mass-to-charge ratio

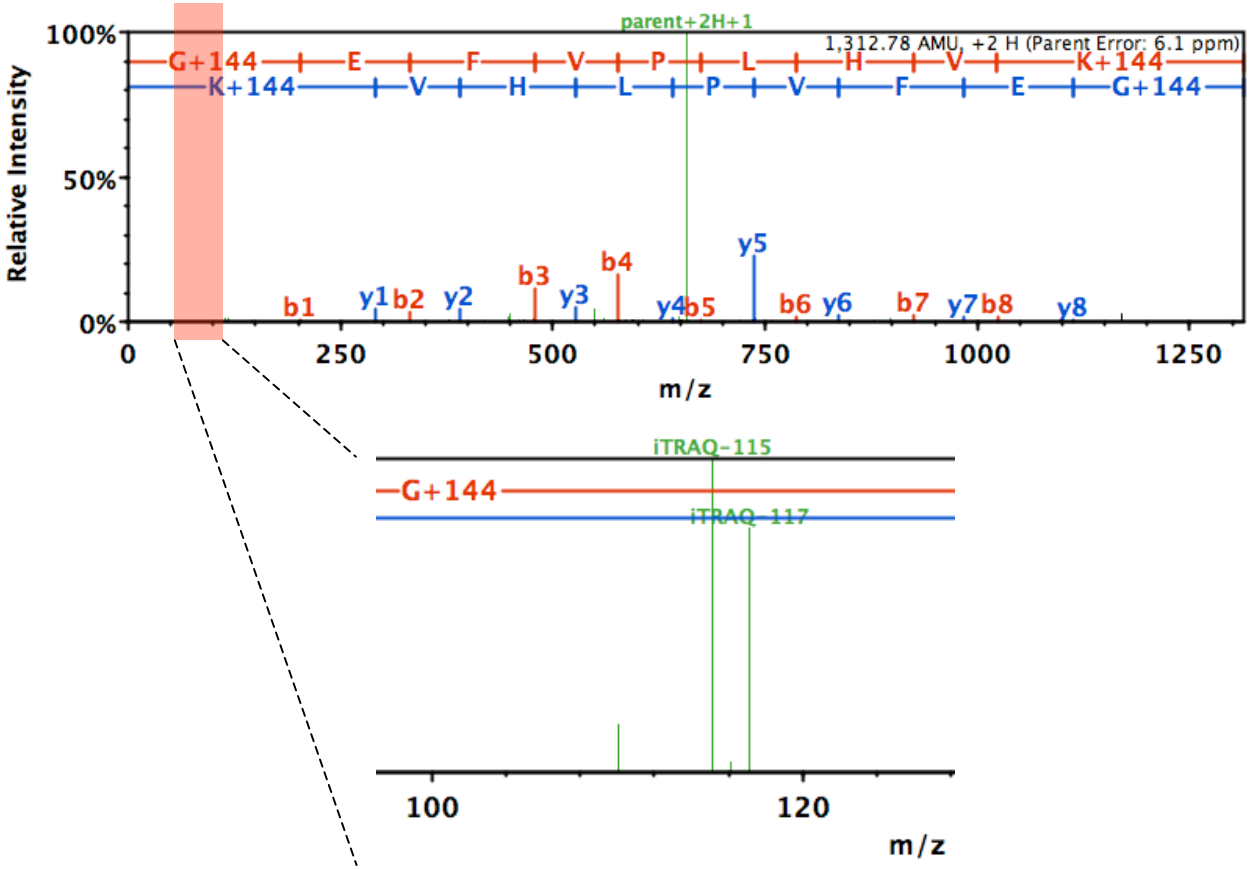

parent+2H+1
1,312.78 AMU, +2 H (Parent Error: 6.1 ppm)
G+144 E F V P L H V K+144
K+144 V H L P V F E G+144
Relative Intensity
100%
50%
0%
0
250
500
750
1000
1250
m/z
b1
y1
b2
y2
b3
y3
b4
y4
b5
y5
b6
y6
b7
y7
b8
y8
iTRAQ-115
G+144
iTRAQ-117
100
120
m/z

Figure 4.3 Example of histone deacetylase 2A (HD2A) quantification by the Scaffold Q+ sortware

(A) Protein display from Scaffold Q+ showing the position of identified peptides on the protein sequence. Green-coloured amino acids indicate modified residues, either due to the presence of the iTRAQ reagent at the amino terminus of the peptide or due to carbamidomethylation or methionine oxidation.

(B) Methodology applied by Scaffold Q+ to calculate the protein ratio of HD2A

First the peptide median Log2 ratio is calculated followed by the median Log2 ratio of the protein. Peptide Log2 ratios are represented by the median of the reporter ion intensities for all the assigned spectra. For HD2A, at the eight identified peptides there are 2, 6, 5, 6, 3, 2, 4 and 2 spectra assigned, respectively. Corresponding Log2 ratios for each of the spectra for a single peptide are shown in the second merged row of the table. The median Log2 ratio for each of the peptide is shown at the fourth row of the table. Finally, the Log2 ratio of HD2A between the N2 and N6 samples is calculated to be as -1.6 corresponding to a fold change of 0.33; corresponding to almost 3-times less HD2A in the N6 sample

(A)

AT3G44750.1 (100%) 26,371.8 Da
HD2A (HISTONE DEACETYLASE 2A); nucleic acid binding / zinc ion binding

M E F W G I E V K S G K P V T V T P E E G I L I H V S Q A S L G E C K N K K G E
F V P L H V K V G N Q N L V L G T L S T E N I P Q L F C D L V F D K E F E L S H
T W G K G S V Y F V G Y K T P N I E P Q G Y S E E E E E E E E E E D D S D G E D S E E
A V A K P K A K P A E V K P A V D D E E D E S D S D S D G M D E K A P V S A K K A K V A V T P Q K T D E
E E P T P K K P A S S K K R A N E T T P G S C K K T F N S G N A L E S H N K A K
K K K G G K A A N Q S P K S A S Q V S C
H A A A K *

(B)

Peptide	AANQSPK	ANETTPK	APVSAK	GEFVPLHVK	KGEFVPLHVK	SASQVSSCGSCK	TFNSGNALESHNK	VAVTPQK
Number of spectra	2	6	5	6	3	2	4	2
Spectra ratios	-1.37	-1.81	-1.94	-1.16	-1.58	-2.49	-2.17	-1.41
	-1.68	-1.80	-1.73	-1.13	-1.48	-2.47	-2.00	-1.39
		-1.75	-1.56	-1.10	-1.20		-1.93	
		-1.68	-1.17	-1.07			-1.72	
		-1.63	-0.27	-0.90				
		-1.61		-0.60				
Average ratio per peptide	-1.5	-1.7	-1.6	-1.1	-1.5	-2.5	-2.0	-1.4
Median ratio per protein	**-1.6**							

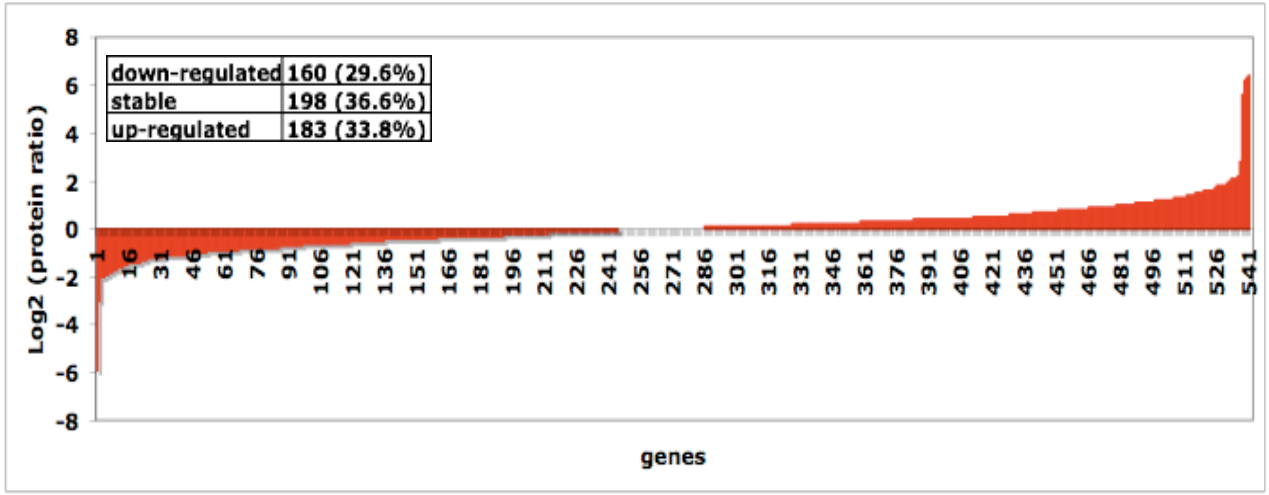


Figure 4.4 Protein ratio distribution of nuclear$_{PU}$ from sample K005.

Log_2 transformed protein ratios (from Scaffold Q+) plotted against proteins sorted in an ascending order of ratios. Numbers and percentages (as a portion of the total number) for each protein group are given in table.

4.3.1 Gene enrichment analysis of quantified nuclear$_{PU}$ proteins

BiNGO enrichment analysis for molecular function was performed in each of the three nuclear$_{PU}$ protein datasets (down-regulated, up-regulated and constant), and the resulting networks of GO categories are shown in Figure 4.5. "Binding" and "Structural molecule activity" are the two most over-represented categories in the down- and up-regulated proteins (Figure 4.5).

"Binding" GO category

The "binding" GO categories have different p-values assigned (different colours assigned to each circle), with that of down-regulated being lower than that of the up-regulated. A breakdown of up-regulated proteins in this category showed that a significant number of those related to nucleic acid binding are also related to translation factor activity, suggesting that the majority of these relate to RNA binding. The group of down-regulated proteins in the binding category includes proteins involved in DNA metabolic processes and chromatin binding as well as RRM-containing proteins (Figure 4.5A; See also Appendix-II). A reduction at the level of these proteins suggests an alteration of the status of processes related to the proliferative phase of the cell culture.

"Structural molecule activity" GO category

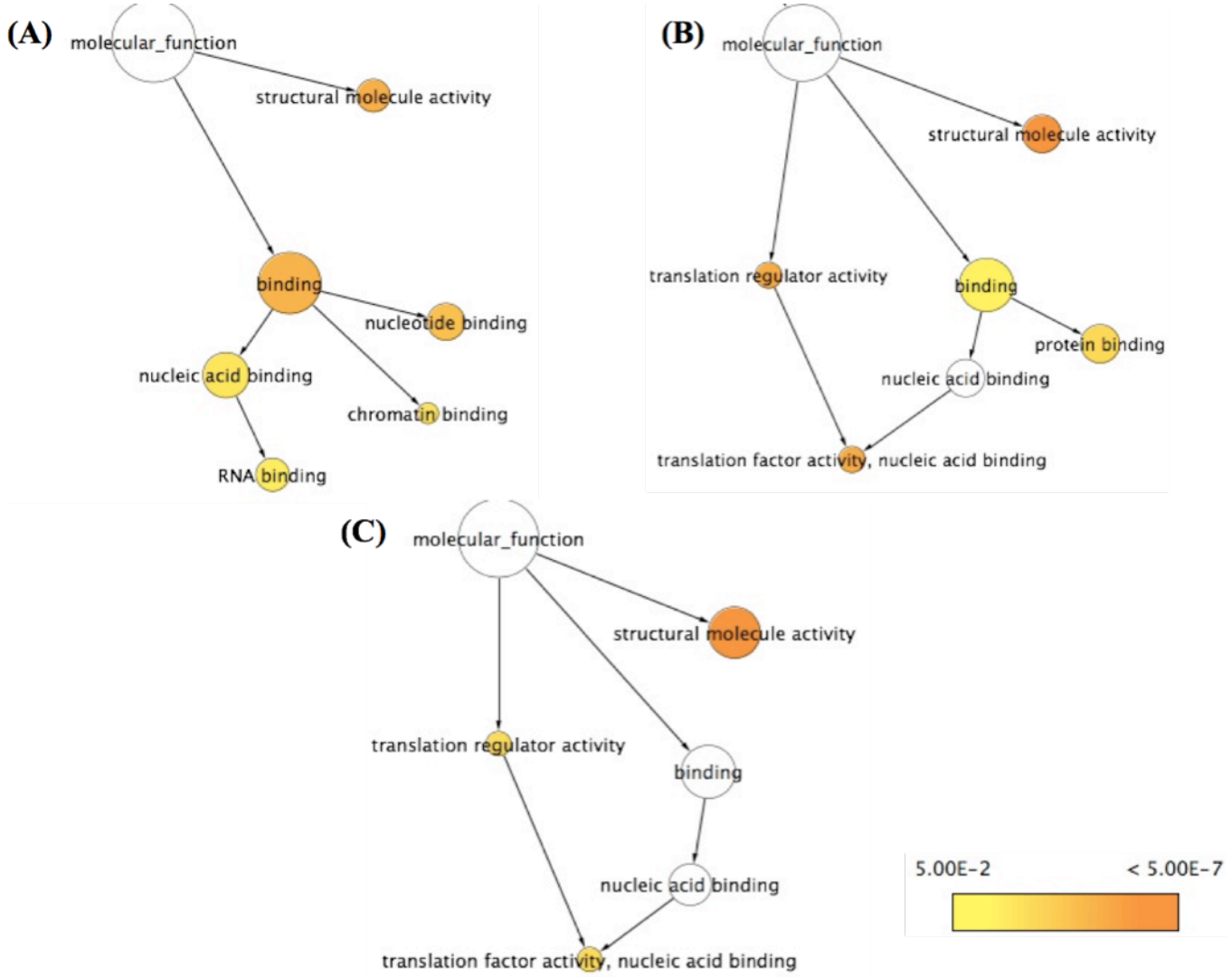
"Structural molecule activity" proteins formed enriched groups in all the protein groups (up-, down-regulated and constant), but this category was highly over-represented in the protein group that remained constant (Figure 4.5C), and comprised mainly of ribosomal proteins (Appendix-II). Considering the increased translational rates during the proliferation phase of a cell culture (Proud, 2002), constant levels for the majority of ribosomal proteins suggest that their activity is controlled post-translationally (see Chapter 7).

Figure 4.5 GO enrichment analysis for nuclear proteins

BiNGO tools were used for detecting over-represented GO groups based on molecular function and generating the respective networks. The size of each circle is proportional to the number of genes included.

(A) Network for the down-regulated proteins
(B) Network for up-regulated proteins and
(C) Network for the constant proteins.

Colour scale represents the statistical significance of each category with the ones with yellow colour being over-represented in proteins of a certain molecular function at the p value of 0.05 or lower, whereas the orange ones are over-represented at the significance level of 5×10^{-7} or higher.


(A)
molecular_function
structural molecule activity
binding
nucleotide binding
nucleic acid binding
chromatin binding
RNA binding
(B)
molecular_function
structural molecule activity
translation regulator activity
binding
protein binding
nucleic acid binding
translation factor activity, nucleic acid binding
(C)
molecular_function
structural molecule activity
translation regulator activity
binding
nucleic acid binding
translation factor activity, nucleic acid binding
5.00E-2
< 5.00E-7

4.3.2 Functional categorization of quantified nuclear$_{PU}$ proteins from the K005 experiment

Enrichment analysis does not provide a full spectrum of the functional distribution of proteins within a dataset. For this reason, I assigned a GO functional category to each of the proteins in the three datasets (down-regulated, up-regulated and constant), using the TAIR website tool "Bulk Data Retrieval" (www. a*rabidopsis*.org). Manual editing of the GO categories ensured that no overlap in members exists among the functional categories. A protein that belonged to more than categories was assigned to the Go category that most precisely represented its actual function. For example, a protein that was assigned to the "hydrolase activity" and "kinase activity" GO categories, which exhibit a parent-sibling relationship, was removed from the former GO category since the latter ones informs us about protein function in a more precise way.

As a confirmation of the enrichment analysis, DNA/RNA binding proteins represented the GO category with the highest percentage in the group of down-regulated proteins, and also had the highest percentage among the three datasets (Figure 4.6). Similar confirmation is obtained in the case of ribosomal proteins, as members of the "structural constituent of the ribosome" GO category. In the dataset of down-regulated proteins there is a group of proteins that form the "inferred function" GO category (Figure 4.6A). These are proteins whose function is not known but their subcellular localisation is known and, thus, their function could be inferred indirectly based on their subcellular residence. Also, the proportion of proteins involved in protein or ion binding is considerably higher in the group of up-regulated proteins, when compared to the other two datasets (Figure 4.6B). These protein GO category includes transporters and metal ion binding proteins that have been found to act in the cytoplasm or mitochondria. This suggests that these proteins are either contaminants or some of these proteins translocate to the nucleus in quiescent cells. *In vivo* protein localisation could discriminate between these two alternatives.

Figure 4.6 Percentage distributions of GO categories of nuclear$_{PU}$ proteins quantified by iTRAQ method

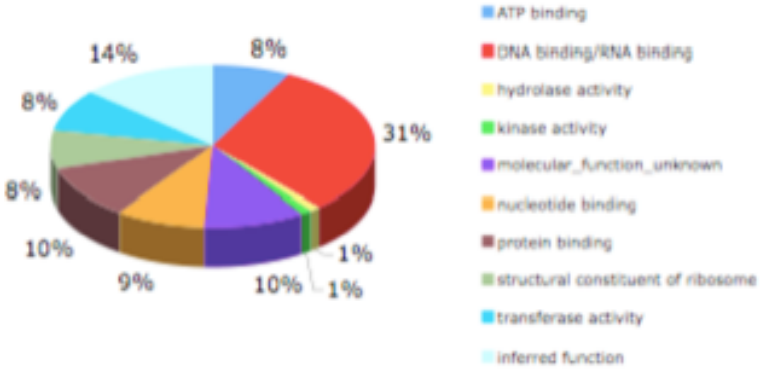
AGI numbers of proteins that were differentially expressed between the N2 and N6 samples were used as an input file in the TAIR website (www.*arabidopsis*.org) for obtaining all the GO annotations for each gene (see Appendix-II for protein members of each GO category).

(A) Pie chart showing percentage distribution of GO categories in the group of down-regulated proteins

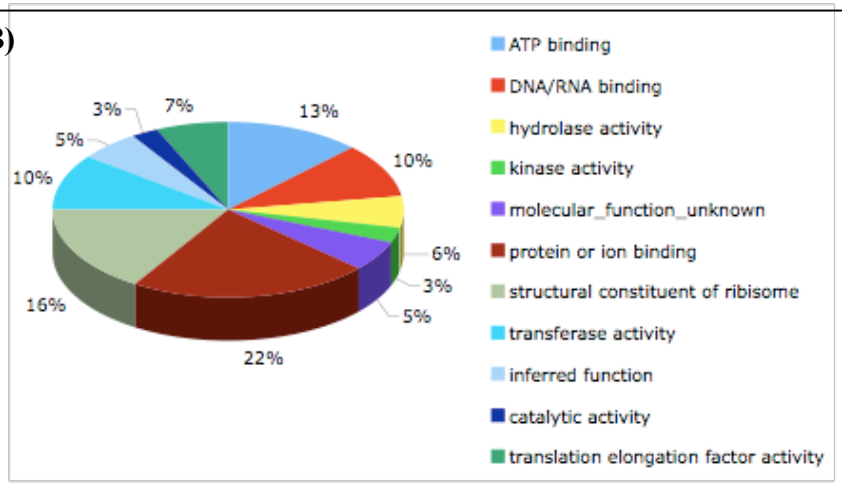
(B) Pie chart showing percentage distribution of GO categories in the group of up-regulated proteins

(C) Pie chart showing percentage distribution of GO categories in the group of constant proteins

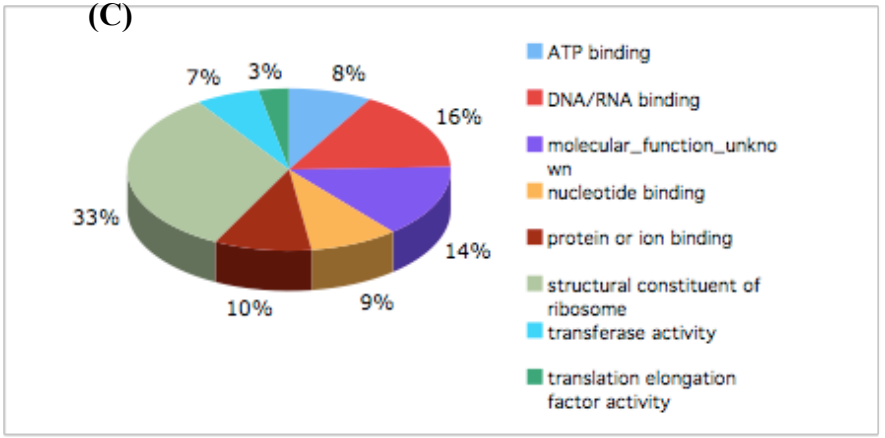
(A)


14%
8%
8%
31%
8%
1%
10%
9%
10%
1%
ATP binding
DNA binding/RNA binding
hydrolase activity
kinase activity
molecular_function_unknown
nucleotide binding
protein binding
structural constituent of ribosome
transferase activity
inferred function

(B)


3%
7%
13%
5%
10%
10%
6%
3%
16%
5%
22%
ATP binding
DNA/RNA binding
hydrolase activity
kinase activity
molecular_function_unknown
protein or ion binding
structural constituent of ribisome
transferase activity
inferred function
catalytic activity
translation elongation factor activity

(C)


7%
3%
8%
16%
33%
14%
10%
9%
ATP binding
DNA/RNA binding
molecular_function_unkno
wn
nucleotide binding
protein or ion binding
structural constituent of
ribosome
transferase activity
translation elongation
factor activity

4.3.3 RNA and DNA binding proteins quantified by iTRAQ

4.3.3.1 RNA-binding proteins tend to be down regulated in the quiescent cell nuclear proteome.

RNA processing proteins included members regulating both pre-rRNA and pre-mRNA cleavage and modifications. Two members of the small nucleolar RNAs (snoRNAs) family of *Arabidopsis* U3.2 and U3.6 (Brown *et al*., 2001), were down-regulated by 2-fold (log_2 ratios -1.1 and -0.9, respectively). snoRNAs are conserved across eukaryotes and are the RNA-based components of snoRNPs (sno ribonucleoprotein particles) that are responsible for processing pre-rRNAs to mature rRNA subunits (reviewed in Dieci *et al*., 2009). Additional protein components of snoRNPs were also found to be down-regulated such as NOL1/Nop2p-like, FtsJ-like, PARL1 and FIB2 that all mediate the methylation of ribose bases of rRNAs (de Beus *et al*., 1994; Ching *et al*., 2002; Barneche *et al*., 2000; Petricka and Nelson, 2007). The ratio of these proteins ranges from 1.7 for PAL1 to 4 for FIB2 (respective log_2 ratios were -0.9 and -2). Another protein involved in rRNA processing was the Block Of Proliferation1-like (BOP1-like), which was down-regulated in the quiescent cells with a log_2 ratio of -0.8. BOP1-like is responsible for the maturation of pre-rRNAs and its inactivation causes cell growth arrest (Strezoska *et al*., 2002). The pescadillo-like protein, involved in ribosome biogenesis, was also down-regulated with a fold change of two. In mammalian cells, pescadillo is important for nucleolus assembly and cell proliferation (Lerch-Gaggl *et al*., 2002).

Proteins regulating pre-mRNA processing were either down-regulated or remained constant in quiescent cells compared to proliferating ones. These include pre-mRNA splicing and mRNA export factors as well as proteins with putative RNA recognition motifs (RRM; Table 4.1B). Proteins of the former group are responsible for splice site recognition of pre-mRNAs and spliceosome assembly (Kalyna and Barta, 2004; Long and Caceres, 2009). Splicing components were either down-regulated or remained constant (Table 4.1B). Log_2 ratios for splicing factors ranged from -0.3 to -1, corresponding to fold change from 1.2 to 2 respectively, whereas a small number remained constant with log_2 ratios from 0.1 to -0.2.

The RRM domain is one of the most widespread protein domains in eukaryotic proteins and has been implicated in pre-mRNA and pre-rRNA processing via the formation of

protein-RNA or protein-protein complexes (Maris *et al*., 2005). Log_2 ratios of RRM-containing proteins ranged from -0.5 to -1.9, corresponding to a fold change between 1.41 to 3.7 (Table 4.1B). The Mago-nashi homolog (Magoh) protein was down-regulated in the stationary nuclei with a log_2 ratio of -0.3. Magoh is part of the EJC (Kataoka *et al*., 2001), interacts with Y14 RNA-binding protein in Droshophila (Zhao *et al*., 2000) and co-localises with *Arabidopsis* CDKC2 protein kinase (Chapter 3; Kitsios *et al*., 2008).

A small number of mRNA processing proteins were found to be up-regulated and included one RRM-containing protein, a poly(A) binding protein and a splicing factor. Log_2 ratios for these proteins ranged from 0.3 to 0.5 (Table 4.1B).

RNA binding proteins also included eukaryotic initiation factors (eIFs). Members of this protein family are responsible for recruiting mature mRNAs to the ribosome for subsequent translation (Browning, 2004). eIF2A and eIF4A-2 were up-regulated at log_2 ratios of 0.3 and 1 respectively. Other eIF factors, eiF4G, eIF4A, eIF3 and eIF1a, remained constant whereas none of eIFs was found to be down-regulated.

In summary, most of the RNA binding proteins where either down-regulated or remained constant. The increased abundance of RNA binding proteins in proliferation is in agreement with increased cellular activity manifested with high cell division rate.

4.3.3.2 DNA binding proteins show a diverse behaviour between proliferation and quiescence

Transcription factors that are highly expressed in proliferating tissues in plants, such as SERRATE (Clarke *et al*., 1999) and SPT16 (Lolas *et al*., 2009), were down-regulated in quiescent cells and the same holds true for the transcription initiation factor TAF4, which in *C. elegans* is sequestered to the cytoplasm for transcriptional suppression (Guven-Ozkan *et al*., 2008) (Table 4.1A). A large portion of constant proteins included those involved in chromatin binding and modification such as histones, histone deacetylases and high mobility group proteins. Histones showed diverse behaviour as the majority was up-regulated in proliferating nuclei, a small number down-regulated and the rest remained at similar levels whereas histone deacetylases were down-regulated in the stationary phase nuclei (Table 4.1A).

Members of the high mobility group (HMG) protein family are part of the chromatin remodelling complex, FACT, and bind transiently to non-condensed chromatin, facilitating anchorage of transcription machinery on their respective transcription sites (Duroux *et al*., 2004; Launholt *et al*., 2006). The observed down-regulation of most HMG proteins in quiescent nuclei is consistent with their involvement in promoting cell proliferation (Reeves *et al*., 2001).

Table 4.1. DNA or RNA binding proteins that are differentially expressed between the proliferative and stationary phases of the cell culture.

The first column of each table provides the AGI number, the second column gives $\log_2$ protein ratios as calculated from the respective iTRAQ-based peptide ratios (see also Figure 4.3), and the third gives the protein name or description. The reference state is the proliferative phase (N2). A complete list of the DNA/RNA binding proteins is given in Appendix-II

(A) DNA binding proteins that are down-regulated (i), up-regulated (ii) or remain constant (iii)
(B) RNA binding proteins that are down-regulated (i), up-regulated (ii) or remain constant (iii)

(A)

AGI	Log_2(N6/N2)	Description
AT5G22650	-3	HD2B
AT4G10710	-2	SPT16 (GLOBAL TRANSCRIPTION FACTOR C)
AT3G44750	-1.6	HD2A (HISTONE DEACETYLASE 2A)
AT2G17560	-1.5	HMGB4 (HIGH MOBILITY GROUP B 4)
AT3G28730	-1.4	ATHMGA2
AT1G61730	-1.2	DNA-binding storekeeper protein
AT1G74560	-1.1	NRP1 (NAP1-RELATED PROTEIN 1)
AT5G43130	-0.9	TAF4; transcription initiation factor
AT5G03740	-0.8	HD2C
AT4G17950	-0.7	DNA-binding family protein
AT4G08350	-0.7	KOW domain-containing transcription factor
AT3G10690	-0.6	DNA gyrase subunit A family protein
AT5G64420	-0.6	DNA polymerase V family
AT2G33845	-0.6	DNA-binding protein-related
AT4G26110	-0.6	NAP1;1; DNA binding
AT2G27100	-0.6	SE (SERRATE)
AT3G20670	-0.5	HTA13; DNA binding
AT3G53650	-0.4	histone H2B, putative
AT3G51880	-0.4	HMGB1
AT3G57660	-0.4	NRPA1
AT1G29940	-0.4	NRPA2 (nuclear RNA polymerase A 2)
AT4G21710	-0.4	NRPB2; DNA binding

(i)

AGI	Log_2 (N6/N2)	Description
AT2G28720	1.4	histone H2B, putative
AT1G07660	1.1	histone H4
AT2G30620	1	histone H1.2
AT1G06760	0.8	histone H1
AT1G09200	0.7	histone H3
AT5G54640	0.5	HTA1; DNA binding
AT4G05420	0.4	DDB1A (UV-damaged DNA-binding protein 1A)
AT4G21100	0.4	DDB1B
AT3G10270	0.4	DNA topoisomerase II, putative
AT3G45980	0.3	HTB9; DNA binding
AT5G11260	0.3	HY5; DNA binding
AT5G47210	0.3	nuclear RNA-binding protein, putative
AT4G02060	0.3	PRL (PROLIFERA); DNA binding

(ii)

AGI	Log_2 (N6/N2)	Description
AT1G48620	-0.2	HMGA5; DNA binding
AT2G19480	-0.2	NAP1;2 (NUCLEOSOME ASSEMBLY PROTEIN1;2)
AT1G09770	-0.1	ATCDC5; DNA binding
AT5G67630	-0.1	DNA helicase, putative
AT1G51060	-0.1	HTA10; DNA binding
AT4G27230	-0.1	HTA2 (HISTONE H2A)
AT1G09770	-0.1	ATCDC5; DNA binding
AT3G22320	0	ATRPABC24.3 (RNA POLYMERASE SUBUNIT); DNA binding
AT3G13940	0	DNA-directed RNA polymerase
AT1G14900	0	HMGA; DNA binding
AT2G38250	0.1	DNA-binding protein-related
AT1G52740	0.1	HTA9; DNA binding
AT4G32605	0.1	transcription factor
AT5G27670	0.2	HTA7; DNA binding

(iii)

(B)

AGI	Log_2(N6/N2)	Description
AT5G55660	-1.4	GTP binding / RNA binding
AT3G11964	-1.1	S1 RNA-binding domain-containing protein
AT1G09760	-1	U2A' (U2 small nuclear ribonucleoprotein A)
AT5G14520	-1	pescadillo-related
AT5G04600	-1	RRM-containing protein
AT3G16810	-0.9	APUM24; RNA binding
AT5G55670	-0.8	RRM-containing protein
AT1G60650	-0.7	glycine-rich RNA-binding protein, putative
AT2G16940	-0.7	RRM-containing protein
AT5G59950	-0.7	RNA and export factor-binding protein, putative
AT1G07360	-0.6	zinc finger (CCCH-type) family protein / RRM-containing protein
AT3G50670	-0.5	U1-70K (SPLICEOSOMAL PROTEIN U1A); RNA binding
AT4G00830	-0.5	RRM-containing protein
AT3G26560	-0.4	ATP-dependent RNA helicase, putative
AT5G60980	-0.4	NTF2 family protein / RRM-containing protein
AT1G02140	-0.3	HAP1/MAGO/MEE63
AT1G03140	-0.3	splicing factor Prp18 family protein
AT1G06220	-0.3	MEE5 translation elongation factor
AT2G20490	-0.3	NOP10; RNA binding
AT4G26630	-0.3	GTP binding / RNA binding
AT4G35800	-0.3	NRPB1

(i)

AGI	Log_2(N6/N2)	Description
AT1G67680	0.8	7S RNA binding
AT3G58510	0.5	RH11 RNA helicase, putative
AT2G42520	0.4	DEAD box RNA helicase, putative
AT1G49760	0.4	PAB8 (POLY(A) BINDING PROTEIN 8); RNA binding
AT2G23350	0.3	PAB4; RNA binding
AT5G40490	0.3	RRM-containing protein
AT3G18790	0.3	Isy1-like splicing factor

(ii)

AGI	Log_2(N6/N2)	Description
AT3G55460	-0.2	SCL30 (SC35-like splicing factor 30)
AT2G24590	-0.2	splicing factor, putative
AT3G55220	-0.2	splicing factor, putative
AT5G64270	-0.2	splicing factor, putative
AT1G60850	-0.1	ATRPAC42 (RNA polymerase I subunit 42)
AT2G33730	-0.1	DEAD box RNA helicase, putative
AT1G69250	-0.1	nuclear transport factor 2 /RRM-containing protein
AT1G66260	-0.1	RNA and export factor-binding protein, putative
AT2G47580	-0.1	U1A (SPLICEOSOMAL PROTEIN U1A); RNA binding
AT1G51510	-0.1	Y14; RNA binding
AT3G19760	0	eIF4A, putative
AT5G37720	0	RNA and export factor-binding protein, putative
AT4G18465	0	RNA helicase, putative
AT2G44710	0	RRM-containing protein
AT3G26420	0.1	ATRZ-1A; RNA binding
AT4G17520	0.1	nuclear RNA-binding protein, putative
AT1G02840	0.1	SR1 (splicing factor 2); RNA binding

(iii)

4.3.3.3 Proteins of unknown function

Quantification analysis in Scaffold Q+ identified proteins of unknown function that were of similar abundance in proliferative and quiescent cell cultures. In total, 53 proteins of unknown function were quantified. Nineteen of them were down-regulated in the N6 sample, 10 were up-regulated whereas 24 remained constant (Table 4.2). The presence of this number of unknown proteins provides an opportunity to identify novel components involved in cell proliferation or growth.

(A)

AGI	Log2(N6/N2)
AT1G69070	-1.8
AT2G43650	-1.6
AT1G31660	-1.4
AT5G57120	-1.4
AT1G63810	-1.3
AT3G57940	-1.2
AT2G26680	-1.1
AT5G66540	-1.1
AT1G15420	-1
AT3G23620	-1
AT4G28200	-0.9
AT5G48240	-0.9
AT3G15460	-0.6
AT3G16310	-0.6
AT1G23280	-0.4
AT1G32130	-0.4
AT3G57990	-0.4
AT3G10650	-0.3
AT5G48680	-0.3

(B)

AGI	Log2(N6/N2)
AT2G44060	0.3
AT3G56900	0.3
AT5G42570	0.4
AT1G73230	0.8
AT4G30010	1
AT3G48140	1.1
AT2G15860	1.2
AT1G55900	1.4
AT3G18790	2.2
AT5G64260	2.8

(C)

AGI	Log2(N6/N2)
AT1G60170	-0.2
AT2G03510	-0.1
AT3G10380	-0.1
AT3G12140	-0.1
AT3G14120	-0.1
AT5G65770	-0.1
AT1G07170	-0.1
AT1G15200	0
AT1G67230	0
AT3G49720	0
AT4G24820	0
AT5G12370	0
AT5G24710	0
AT5G42960	0
AT5G51200	0
AT1G70770	0.1
AT3G29075	0.1
AT4G32605	0.1
AT4G32910	0.1
AT5G02050	0.1
AT5G12410	0.1
AT5G40480	0.1
AT1G61120	0.2
AT5G62270	0.2

Table 4.2. Proteins of unknown function quantified using the iTRAQ method. (A) Down-regulated proteins, **(B)** up-regulated proteins and **(C)** constant proteins.

4.4 Correlation of quantitative changes at the mRNA and protein levels of nuclear-localised gene products

The combination of microarray analysis and proteomics can reveal mechanisms of gene and protein regulation by correlating transcript and protein levels on a global scale, thus providing an holistic insight into the responses of the biological system studied (Xia *et al*., 2006; Li *et al*., 2007). Based on the degree of correlation between the two types of datasets, one can ask to what extent changes at the transcriptome level are directly translated into changes at the proteome level.

A query of microarray data repositories identified a single microarray experiment, where the expression profile of genes in the same *Arabidopsis* cell culture was measured as cell cultures move from proliferation to quiescence (Menges *et al*., 2003). Samples analysed in the specific microarray experiment were taken at 1, 3, 5 and 7 days after subculture, which did not exactly correspond to the samples used for the K005 experiment. However, since there were few significant changes in mRNA levels from day_1 to day_3 and from day_5 to day_7 (Jim Murray, personal communication), Values from day_1 and day_3 and from day_5 and day_7 were averaged, used to calculate the ratio of day_(1,3) to day_(5,7) and compared to proteomics datasets from day_2 and day_6 nuclei. A graphical representation of pairwise comparisons for $\log_2$ transformed protein ratios for the K005 dataset and respective mRNA ratios are shown in Figure 4.7. The R^2-value is very low, indicating a poor correlation between changes in transcript levels and their corresponding proteins. The majority of genes are down-regulated in stationary cells, whereas there is a considerable variation in the abundance of the corresponding proteins in the nucleus.

Although the correlation between mRNA and protein levels is often low in such a diverse gene sample, stronger correlations can emerge when genes/proteins belonging to the same family or molecular function are compared (White *et al*., 2004; Washburn *et al*., 2003). However, even when I compared mRNA and protein levels for transcription factors or translation initiation factors only, no pattern emerged (data not shown). Thus, it appears that for the proteins studied, the mechanisms regulating their abundance appear to act at the post-transcriptional level. Ideally protein and mRNA should be prepared from the same material. Although the cell culture used for proteomics came from the Murray lab and should be genetically identical, the experiments were carried

out several years apart. Also, even though the culture conditions and media were based on the methods described, the effect of growing the material in different places is likely to also introduce differences.

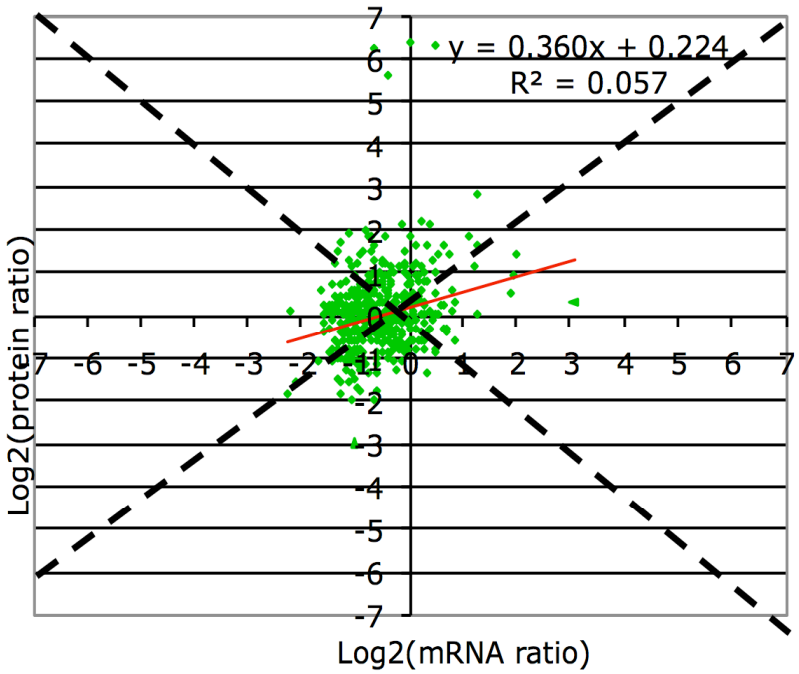


Figure 4.5. Correlation analysis between transcript and protein levels of genes from K005 sample.

Protein ratios from K005 sample were $\log_2$ transformed and compared with $\log_2$-transformed ratios from a published microarray experiment (see text for details about formatting microarray dataset). For both data the reference point was the proliferation state

The dotted lines indicate the axis where the R^2-value equals to 1 (or y=x).

4.5 Discussion

The aim of this chapter was to perform the first quantitative proteomics analysis of *Arabidopsis* nucleus in cell cultures, comparing nuclei from proliferating cell cultures with those from stationary cultures. Although further refinement and improvement of the obtained data would be desirable, this aim was largely achieved. In the following sections I will discuss the challenges in protein quantification using iTRAQ technology, and also incorporate the research findings in the current literature regarding cellular proliferation and quiescence in eukaryotes.

4.5.1 Quantitative proteomics analysis of *Arabidopsis* nuclei

Enrichment analysis as well as functional categorization of proteins quantified by iTRAQ showed that the DNA/RNA-binding category was the most abundant (Figures 4.5 and 4.6). In the following sections I will discuss how this finding complies with the published literature on cell proliferation and quiescence. In addition, I will discuss the alteration of the levels of specific nucleolar proteins and how these changes relate to nucleolus size reduction as cells move from proliferation to quiescence.

4.5.1.1 DNA-binding proteins and cell proliferation

Members of histone deacetylase-2 protein family (HD2A, HD2B and HD2C) belonged to the group of DNA-binding proteins with the most prominent decrease in the quiescent cell population, suggesting a role for this family in the transition of a cell culture from proliferation to stationary phase. Histone acetylation/deacetylation is an important mechanism of gene transcription regulation (Loidl, 2004), with the activity being mainly related to transcriptional suppression in eukaryotes. Nevertheless, HOS2 deacetylase was found to be necessary for gene transcription in yeast (Wang *et al*., 2002). Also, abolishment of histone deacetylase 1 (HD1) activity in mouse embryos resulted in proliferation defects due to up-regulation of CDK inhibitor proteins $p21^{WAF1/CIP1}$ and $p27^{KIP1}$ (Lagger *et al*., 2002). In terms of external stimuli, HD1 was

found to be upregulated upon hormone stimulation in mouse growth-arrested cells (Bartl *et al*., 1997).

The histone deacetylase-2 (HD2) family is plant-specific. Members of this family in *Arabidopsis* have been implicated in establishing leaf polarity independently of *ASSYMETRIC LEAVES1* (*AS1*) and *AS2* (HD2A, HD2B; Ueno *et al*., 2007), in regulating ABA responses (HD2C; Sridha and Wu, 2006). So far, no reports have been published on the role of HD2 members in proliferation or quiescence. The only report from *Arabidopsis* comes for the HD1 protein, a distant homolog of the HD2 members, whose polysome-associated mRNAs were repressed in sucroce-induced starved cell cultures (Nicolai *et al*., 2006). This repression was inversely correlated with the acetylation levels of histone H4 in the starved cells, thus providing a link between translational control and chromatin activity (Nicolai *et al*., 2006). The fact that HD1 was the only HD identified as differentially regulated in that study, is either due to the fact that this is indeed the only HD that is regulated during sucrose-induced starvation, something that is unlikely, or due to the use of whole cell extracts for the microarray analysis, which consequently produced low-resolution data; in contrast to my high-resolution nuclei-enriched fractions. Therefore, it is possible that down-regulation of the HD2 protein levels in quiescent cells is also essential for an increase in acetylated histones at the stationary phase of the cell culture, which in turn would affect the transcription of genes necessary for quiescence state. Assessment of the acetylated status of specific histones, by chromatin immunoprecipitation using antibodies against acetylated histones, would help us to assess whether our hypothesis is true or not. A good candidate for this experiment could be the histone H2A that appears to remain constant from proliferation to quiescence (Table 4.1). Additional DNA-binding proteins that act as negative regulators of cell quiescence and were up-regulated in my proteomics analysis included Spt16, NRP1 and TITAN7. These three proteins are involved in chromatin remodelling and chromosome dynamics with a positive regulation in cell proliferation (Duroux *et al*., 2004; Zhu *et al*., 2006; Tzafrir *et al*, 2002).

High mobility group (HMG) proteins form the second most abundant family of chromosomal proteins after histones (Johns, 1982). They are small and extremely dynamic protein molecules found across eukaryotes that bind randomly genomic DNA,

facilitating gene transcription and DNA replication (reviewed in Bianchi and Agresti, 2005). Up to date, no report is available on the levels of this group of proteins during proliferation. My results had shown that HMGB-type proteins are substantially upregulated during proliferation with only one remaining constant. Up-regulation of HMG protein family is in line with high transcription and DNA replication rates during proliferation, mainly due to high cell division rate.

4.5.1.2 mRNA processing proteins and translation initiation factors

The majority of splicing factors in the analysis remained at a constant level and only a small fraction was down-regulated (Table 4.1). The activity of many splicing factors are likely to be altered by processes acting post-translationally, such as reversible sequestration into speckles (Docquier *et al*., 2004), a process that is mediated by changes in protein phosphorylation in the case of RSp31 (Tillemans *et al*., 2006) and SR45 and SR34 (Ali and Reddy, 2006). Therefore, even splicing factors that appear to be present at a constant level are probably regulated at other levels. The role of protein kinases and transcriptional activity on the behaviour of splicing factors is presented in Chapters 5 and 6.

The presence of translation initiation factors in the nuclear proteome is intriguing. Their long accepted and primary function is based in the cytoplasm (Pestova *et al*, 2001) but translation was shown to be coupled to transcription in mammalian nuclei by detecting the incorporation of [^{3}H]-lysine in nascent transcripts (Iborra *et al*., 2001). Since then, many studies have found translation factors within the nuclei of diverse species (Bush *et al*., 2009; Andersen *et al*., 2002; Pendle *et al*., 2005; Strudwick and Borden, 2002).

Translational activity is minimal in quiescent cells (Proud, 2002). Since the majority of eIFs remain at constant levels suggest, as with splicing factors, regulation is likely to post-translational, probably involving protein phosphorylation. There is clear evidence for such regulation, but mostly for cytoplasmic eIF factors (reviewed by (Pierrat *et al*., 2007). However, a portion of mammalian eIF4G was found to localise in the nucleus in association with the cap-binding complex (McKendrick *et al*., 2001). At least for eIF4G, regulation of its activity through phosphorylation comes from studies in

mammalian cells, where serines 1108, 1148 and 1192 was found to be phosphorylated by the PI-3 kinase *in vivo* after serum stimulation (Raught *et al*., 2000).

4.5.1.3 Nucleolar proteins and cell proliferation

Ribosome biogenesis, a nucleolar process, is crucial for cell growth and proliferation (Warner, 1999; Lerch-Gaggl, 2002). Ribosome biogenesis can consume up to 80% of a cell's energy and declines dramatically as cells enter quiescence (Thomas, 2000). This reduction is mirrored by smaller nucleolar size (see Section 5.3; Figure 5.6) as well as reduced protein levels for certain nucleolar proteins. PARL1 is the predominant form of nucleolin, which is involved in rRNA processing and ribosome synthesis in *Arabidopsis* (Petricka and Nelson, 2007). Down-regulation of PARL1 in quiescent day6 cells, probably mirrors decreased ribosome biogenesis. Another nucleolin protein, NopA64, was down-regulated in differentiated onion (*Allium cepa*) tissues (de Carcer *et al*., 1997), again supporting the idea that the amount of these proteins is decreased in cells that have ceased to divide. Another nucleolar protein, FIB2, was downregulated almost 4-fold. FIB2 associates with box C/D small nucleolar RNAs (snoRNAs) directing 2'-O-ribose methylation of the rRNA (Barneche *et al*., 2000). Similar to nucleolin, onion FIB2-like protein is more abundant in meristematic root cells than in differentiated cells (Medina *et al*., 2000).

Therefore, the changes in these nucleolar proteins correspond with changes in nucleolar size and support the notion that nucleolar size and structure reflects activity (Raska *et al*., 2006; Gonzalez-Gamacho and Medina, 2005).

4.5.2 Integration of transcriptomics and proteomics datasets

Combined analysis of the transcriptome and proteome could provide new insights into the regulation of gene expression. However, the correlation between the changes at the transcript and protein level tends to be low (Nie *et al*., 2007) and I also observed a similar phenomenon (see Figure 4.5). This general lack of correlation, however, could be informative. One idea is that the quiescent cell needs to hold itself in a state of

readiness (with key systems already assembled, but inactive) to respond to more favourable growth conditions. In this scenario, functions where gene and protein levels are highly correlated may include key regulatory factors, whose abundance directly modulates the pathways in which they are involved. For example genes involved in osteoblast specification or cell adhesion- and protein folding-related genes of human breast epithelial cells (Conrads *et al*., 2005; Rogers *et al*., 2008) showed high correlation between mRNA and protein levels, supporting the idea that some important biological processes are tightly controlled at the transcript and protein level.

4.5.3 Sensitivity

In complex mixtures, the number of proteins that can be reliably identified and quantified becomes an important issue. Quantification of the more abundant nuclear proteins, where identifications are based on multiple peptides, was relatively straightforward. However, regulatory proteins such as kinases tend to be present in at low levels and were identified only from one or two peptides and only in some preparations (See Appendix-I). Many known nuclear kinases were not detected but this is not surprising considering that less than 10% of the expected nuclear proteome was identified (Section 3.8.3). Ideas for increasing the sample resolution are discussed in Section 3.8.3.

4.5.4 Future directions

Biological replication of the experiment will be necessary to identify a consistent set of deregulated proteins between cell proliferation and quiescence. Subsequent *in vivo* confirmation of protein localisation would add extra strength to my data, especially by using proteins defined as “unknown function” as well as proteins that are annotated as “cytoplasmic” from the TAIR annotation tools. The use of fluorescent tags for assessing protein localisation is a very robust method used in previous high-throughput studies (Koroleva *et al*., 2005; Pendle *et al*., 2005), whereas high-sensitivity western blot kits in combination with advanced imaging methods (ECL Plus/CCD camera; Dickinson and

Fowler, 2002) could be used to assess for the amounts of selected proteins in nuclear extracts from cells in proliferation and quiescence.

A good use of the BiNGO enrichment analysis tool will be to generate protein networks within each group of down-regulated, constant or up-regulated proteins. In this way functional relationships between proteins could emerge, either by using previously published data or by homology-based comparison with other eukaryotes. This analysis would open the way for pinning down the predicted interactions *in vivo* and allow us to hierarchically arrange proteins within the proper biological context.

Conclusively, data generated in this chapter, could be used as a platform for getting a deeper understanding on the transition from cell proliferation to quiescence in *Arabidopsis* cell cultures and, in a certain extent, identify functional differences and similarities with other eukaryotes.

Chapter 5 Function of Cyclin-dependent kinase C

5.1 Introduction

Pre-mRNA processing constitutes a very important nuclear function but little is known about its underlying regulatory mechanisms in *Arabidopsis* (see Section 1.8). Splicing factors mediate, along with other proteins, the removal of introns from the pre-mRNA transcript. The majority of these proteins in the quantitative proteomics analysis (Chapter 4) remained at a constant level and only a small fraction was down-regulated (Table 4.1). The activity of many splicing factors are likely to be altered by processes acting post-translationally, such as reversible sequestration into speckles (Docquier *et al*., 2004), a process that is mediated in some splicing factors by protein phosphorylation (Tillemans *et al*., 2006; Ali and Reddy, 2006; de la Fuente van Bentem *et al*., 2006). Speckled-like localisation of the CDKC2 kinase in *Arabidopsis* cell cultures (Kitsios, 2007) renders it as a potential regulatory factor of the splicing process by affecting the activity of splicing factors.

In the present Chapter, I describe the dynamics of GFP-CDKC fusions during the cell cycle and investigate the relationship between CDKC2 and different aspects of pre-mRNA processing. I tested the role of CDKC2 by expressing a dominant negative version of the kinase in cell suspensions and characterised its effect on localisation of splicing factors. During the course of the project, CDKC2 was found to interact (in the yeast-two-hybrid system) with Mago-nashi-homolog (Magoh) protein (Ali Pendle and Peter Shaw, unpublished data), a homolog of *Magoh-nashi* protein that is part of the exon-junction complex (EJC) and involved in mRNA export in mammalian cells (Dreyfuss *et al*., 2002; Le Hir *et al*., 2000). For these reasons, I undertook co-localisation studies of CDKC2 fused to GFP with members of exon-junction complex (EJC) fused to RFP. Finally, to assess the function of GFP-CDKC2 protein, I transformed Arabidopsis *cdkc2-2* mutant plants with 35S::YFP:CDKC2 construct. Results indicate that localisation of the fusion protein mirrors that seen in cell cultures and that it was able to completely complement the *cdkc2-1* mutant phenotype back to the wild type, indicating that the protein fusion in functional *in planta*.

5.2 Sub-cellular localization of 35S:GFP:CDKC2 fusion protein during the cell cycle

CDKC2 is a nuclear protein (Cui *et al*., 2007) that localises to sub-nuclear speckles (Kitsios *et al*., 2008). Until recently, no data were available in the literature for the distribution of CDKC2 during the cell cycle. The results presented below were recently published (Kitsios *et al*., 2008).

I generated a 35S::GFP:CDKC2 expression vector that was stably transformed into tobacco BY-2 cell cultures. Figure 5.1 shows the subcellular localization of the GFP:CDKC2 fusion protein during the cell cycle in BY-2 cells. During interphase, GFP:CDKC2 is localized in the nucleus, with localised concentrations in spots, that probably represent nuclear speckles (Section 1.13; Kitsios *et al*., 2008). As soon as the cell enters prophase, when chromatin condenses and nuclear envelope breaks down, GFP:CDKC2 acquires a more diffuse cytoplasmic localization profile that persists during metaphase, anaphase and telophase. In late telophase/cytokinesis, as the new daughter nuclei are beginning to form, GFP:CDKC2 protein accumulates in the nuclei, where it begins to acquire its interphase-specific localization profile described above (Figure 5.1).

Figure 5.1. Distribution of GFP:CDKC2 during the cell cycle.

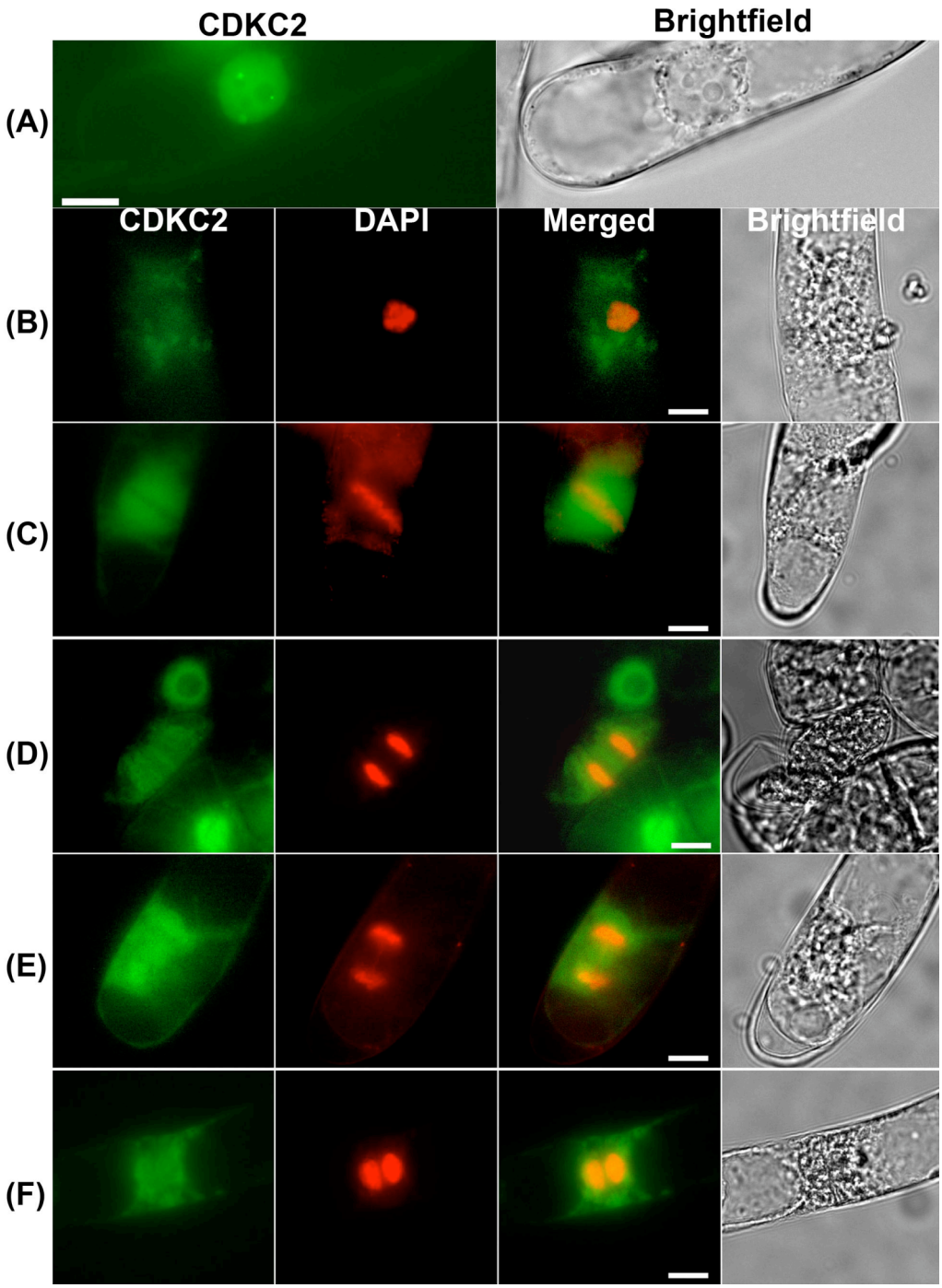
Tobacco BY-2 cell cultures stably expressing GFP:CDKC2 were observed to determine protein fusion localisation during the different stages of the cell cycle. GFP and DAPI are depicted in false colour, green and red respectively.

(A) Cell nucleus in interphase. The GFP and brightfield channels are shown.

(B) – (F) Protein localisation during mitosis. The first column depicts the GFP channel, the second column DAPI staining of the DNA, the third column a merged image of the first two and the fourth one a brightfield image of the cell.

(B) prophase
(C) metaphase
(D) anaphase
(E) telophase
(F) late telophase/cytokinesis

The images shown in these panels are the result of a maximum projection of a group of deconvoluted optical z-sections. The same applies to the rest of the images in this chapter, unless otherwise stated. Scale bar: 10 μm


CDKC2
Brightfield
(A)
CDKC2
DAPI
Merged
Brightfield
(B)
(C)
(D)
(E)
(F)

5.3 Is kinase activity necessary for normal CDKC2 localisation?

Activity of TFIIH and pTEF-b multiprotein transcription factors is inhibited in the presence of 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), a transcription elongation inhibitor (Pinhero *et al*., 2004; Yankulov *et al*., 1997). DRB exerts its inhibitory role by competing with the kinase subunits of TFIIH and pTEF-b for ATP and GTP molecules, which are required for CTD phosphorylation (Zandomeni *et al*., 1986; Trembley *et al*., 2003). CDKC2 has been suggested to be the kinase subunit of the plant pTEF-b factor (Fulop *et al*., 2005; Cui *et al*., 2007) and its subcellular distribution is altered dramatically in the presence of DRB (Kitsios *et al*., 2008). However, the specificity of DRB is still uncertain and to assess whether the kinase activity of CDKC2 is required for its localisation in the nucleus or any aspect of mitotic re-localisation, I generated a mutant variant of the protein that is inactive.

A conserved aspartic acid (D182) residue in the ATP-binding pocket, essential for mediating the transfer of γ-phosphate of ATP to the substrate (De Bondt *et al*., 1993), was altered to asparagine (N182) by *in vitro* mutagenesis (see Chapter 2; Section 2.5) and the mutation was confirmed by sequencing. Mutation of D182 in structurally related kinases from yeast and human cells led to kinase inactivation (Mendenhall *et al*., 1988; van den Heuvel and Harlow, 1993).

The resultant construct (35S:GFP:CKDC2_D182N; designated as GFP:CDKC2_DN hereafter) was either transiently or stably transformed into Col-0 and BY-2 cell cultures, respectively. Localization of the mutant fusion during interphase and in different stages of the cell cycle was recorded in tobacco BY-2 cell cultures stably expressing GFP:CDKC2_DN (Figure 5.2). During interphase, the mutant protein showed a less spotted and more diffuse nucleoplasmic localization as compared to the wild type kinase (Figure 5.2A). During mitosis, the localization profile of GFP:CDKC2_DN protein fusion was very similar to its wild type counterpart (Figure 5.2B-F). More specifically, in prophase the mutant kinase was released into the cytoplasm where it remained until telophase. In late telophase/cytokinesis, GFP:CDKC2_DN protein fusion re-established its interphase localization profile within the nucleus.

Figure 5.2. Distribution of GFP:CDKC2_DN during the cell cycle.

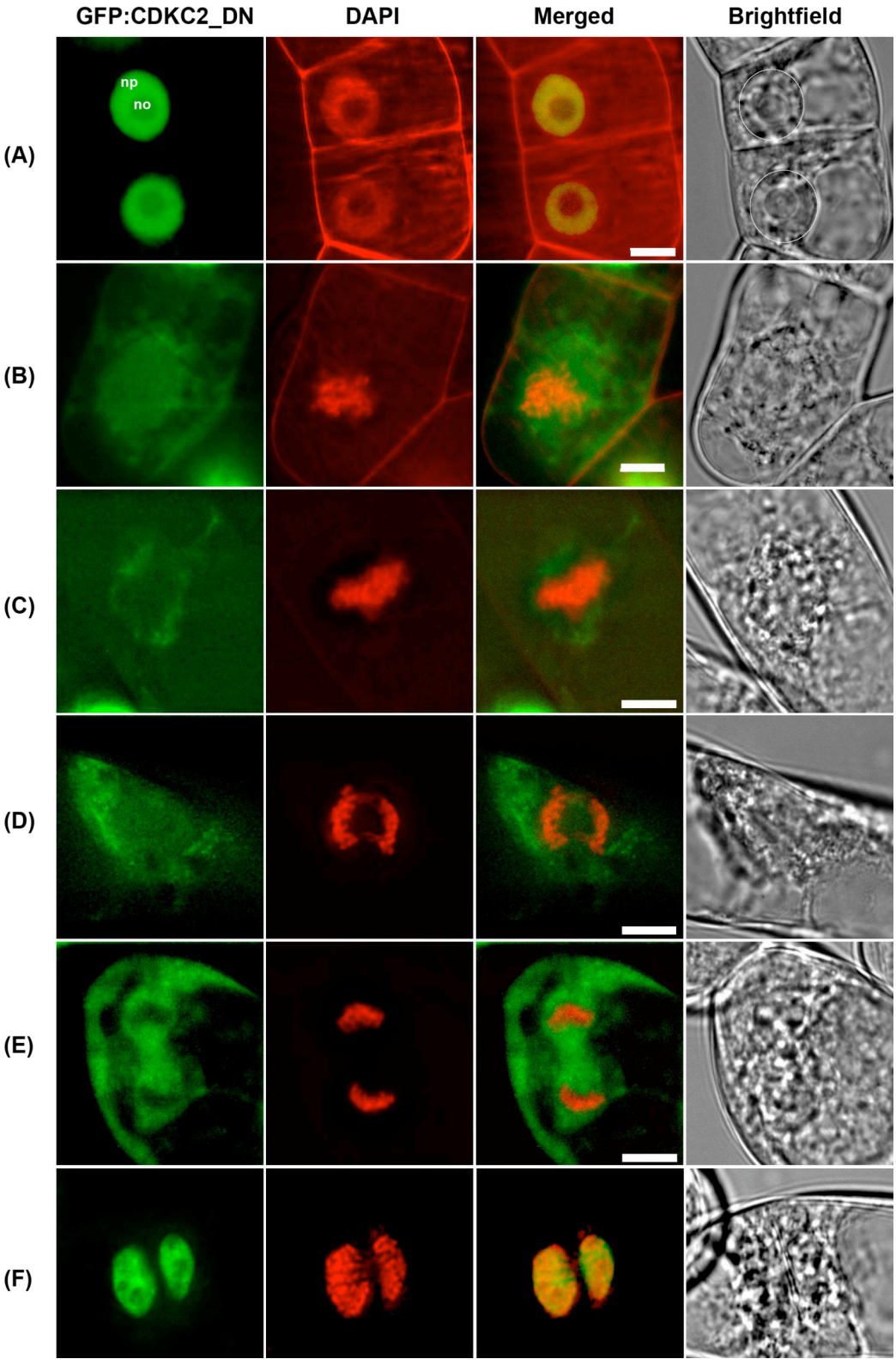
Tobacco BY-2 cell cultures stably expressing GFP:CDKC2_DN were observed to determine protein fusion localisation during the different stages of the cell cycle. The first column shows the GFP channel, the second the DAPI channel (false-coloured in red), the third a merged image of GFP and DAPI channels and the last column a brightfield image of the observed cell.

(A) GFP:CDKC2_DN localisation in interphase nucleus.

(B) – (F) Protein localisation during mitosis:

- **(B)** Prophase
- **(C)** metaphase
- **(D)** anaphase
- **(E)** telophase
- **(F)** late telophase/cytokinesis.

Scale bar: 10 μm . np: nucleoplasm, no: nucleolus. In brightfield image at the interphase nuclei are depicted with dashed circles.


GFP:CDKC2_DN
DAPI
Merged
Brightfield
np
no
(A)
(B)
(C)
(D)
(E)
(F)

Most interestingly, when Col-0 or BY-2 cells were observed under a microscope for prolonged periods (more than 40 min), I could see that the GFP:CDKC2_DN protein gradually translocated into the nucleolus. For the same period of time, the wild type protein was still localized in nuclear speckles and did not show signs of re-localization to the nucleolus. A typical example is shown in Figure 5.3A. In total, I repeated the experiment 4 times and each time I observed GFP:CDKC2_DN translocation into the nucleolus in 13 out of 15 cells. On the contrary, the GFP:CDKC2 protein fusion always localised in nuclear speckles. Thus, partitioning of CDKC2 between the nucleus and nucleolus in response to the stress imposed by observation depends on kinase activity.

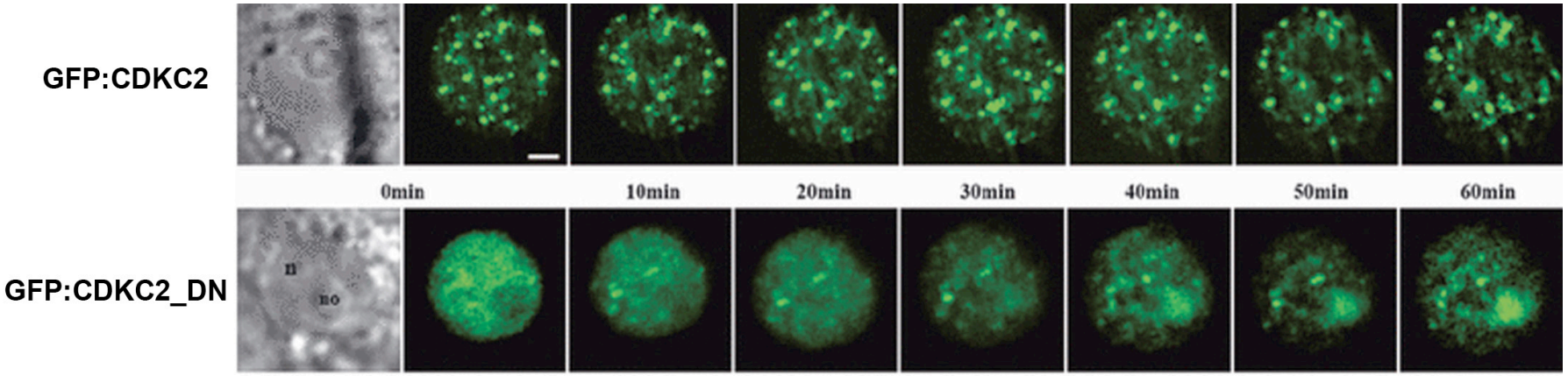
To test this hypothesis, cells expressing GFP:CDKC2 were treated with 100 μm roscovitine, a CDK-specific inhibitor that competes with ATP for binding at the ATP binding pocket (Hardcastle *et al*., 2002). After 2 hours in the presence of roscovitive, GFP:CDKC2 protein fusion translocated into the nucleolus or peri-nucleolar compartments (Figure 5.3B, II-IV), whereas when cells were treated with DMSO no changes in protein localisation were observed (Figure 5.3B, I). These data support the idea that CDKC2 kinase activity is necessary for retaining most of the protein in the nucleoplasm and that upon abolishment of this activity the protein is gradually translocated into the nucleolus.

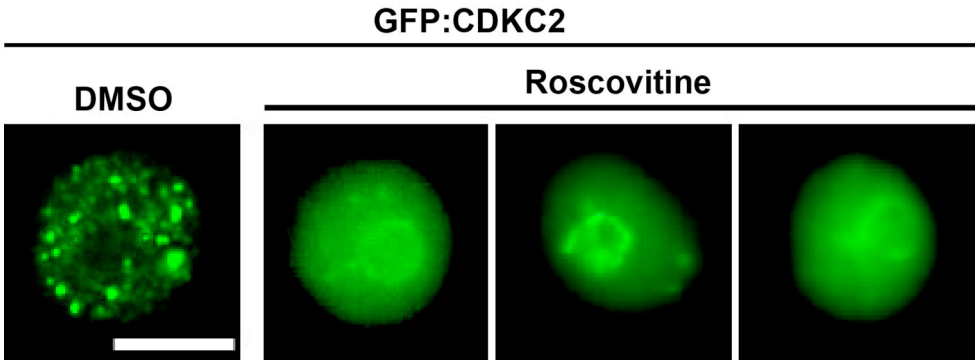
Figure 5.3 Partitioning of CDKC2 between the nucleus and nucleolus depends on stress and kinase activity.

(A) Arabidopsis Col-0 cells separately expressing GFP:CDKC2 and GFP:CDKC2_DN were left on the microscope slide and monitored every 10mins. Translocation of GFP:CDKC2_DN to nucleolus was apparent after 40min.

(B) Arabidopsis cell cultures expressing GFP:CDKC2 were treated with DMSO (I) or 100 μm roscovitine (II, III, IV) for 2 hours and protein fusion localisation was recorded. II, III and IV images show three different nuclei after the roscovitine treatment.

Images were acquired using a fluorescent microscope coupled to a CCD camera. Images represent maximum optical projections of selected z-sections. Scale bar = 10 μm . n: nucleus, no: nucleolus.


GFP:CDKC2
0min
10min
20min
30min
40min
50min
60min
GFP:CDKC2_DN
n
no


GFP:CDKC2
DMSO
Roscovitine

5.4 Co-expression analysis of GFP:CDKC2_DN with spliceosomal components SRp34 and Cyp64

GFP:CDKC2 co-localises with the splicing factor SRp34 and cyclophilin Cyp64, but their co-localisation profiles were different from those observed when each fusion protein is expressed alone (Kitsios, 2007; Kitsios *et al*., 2008). In Section 5.2, I showed that abolition of CDKC2 kinase activity altered the localisation profile of GFP:CDKC2. To test whether localisation profiles of SRp34 and Cyp64 are altered in the presence of the GFP:CDKC2_DN mutant, I transiently co-expressed protein fusions in Col-0 cultures. In Figure 5.4A, GFP:CDKC2 and SRp34:RFP protein fusion co-localise almost entirely in the nucleus showing a speckle-like co-localisation profile. When the kinase activity of GFP:CDKC2 is abolished, both protein fusions co-localise but this time their localisation profile is altered dramatically (Figure 5.4B). Both protein fusions co-localised in large speckles in the nucleoplasm and also in the nucleolus, mirroring the profiles observed when cells are treated with DRB (Kitsios *et al*., 2008). Repeats of the experiment gave the same result in 95-98% of the cells (n>50), with the only variation being that the number of large speckles varied from about 4 to 6 in different cells.

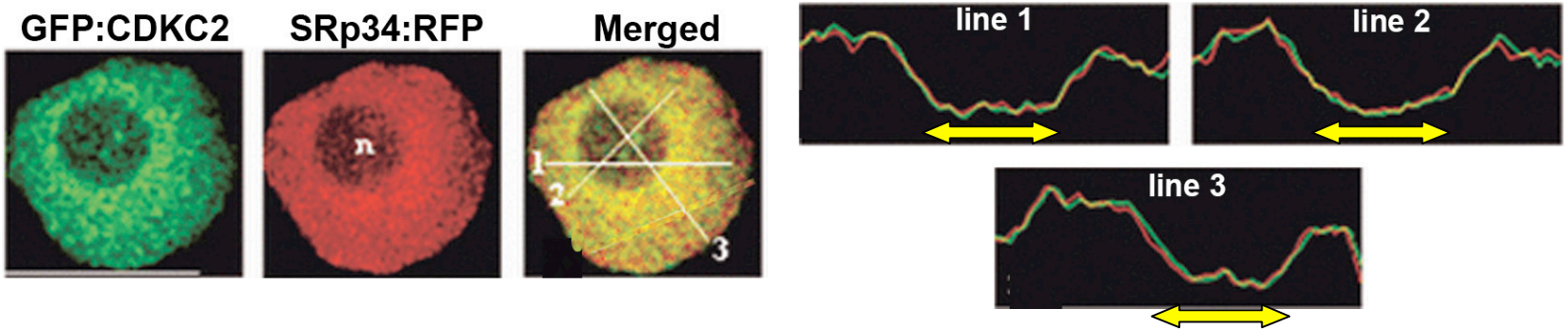
The Cyp64:RFP fusion co-localised to some degree with GFP:CDKC2 protein fusion (Figure 5.5). In the majority of cells examined, and after repeating the experiment, spots could be observed that were labelled by GFP:CDKC2 or Cyp64:RFP only (Figure 5.5; arrows). Interestingly, both GFP:CDKC2 and Cyp64:RFP colocalise to the nucleolus as well, whereas when expressed alone none of the protein fusions reside in the nucleolus (Figure 5.5B). This indicates that the two proteins may interact to affect each other's localisation. Even though both GFP:CDKC2_DN and Cyp64:RFP co-localised in the nucleoplasm, translocation of mutant kinase to the nucleolus was not mirrored by translocation of the Cyp64;RFP protein fusion protein (which remained in the nucleoplasm for the great majority of the cells (Figure 5.6A)). However, in a minority (8%; n=65 in total) of the cells observed, Cyp64;RFP showed nucleolar co-localisation with the kinase and a representative nucleus from this group of cells is shown in (Figure 5.6B).

Figure 5.4 Abolition of CDKC2 kinase activity alters the distribution of CDKC2 and SRp34 fusion proteins in Arabidopsis cell cultures.

(A) Representative nucleus with the localisation profiles of GFP:CDKC2 and SRp34:RFP fusion proteins.

(B) Representative nucleus co-expressing GFP:CDKC2 and SRp34:RFP fusion proteins.

Images in both composite panels are maximum projections of selected Z-sections through each nucleus. The nucleolus is circled in the brightfield image.
Present to the right of each composite panel in (A) and (B), are the RGB profiles of transects drawn through each nucleus using the merged images. Double-headed arrows indicate the RGB profile for the two protein fusions when a line crosses the nucleolus. Scale bars: 7 μm in (A) and 10 μm in (B).


GFP:CDKC2
SRp34:RFP
Merged
n
1
2
3
line 1
line 2
line 3

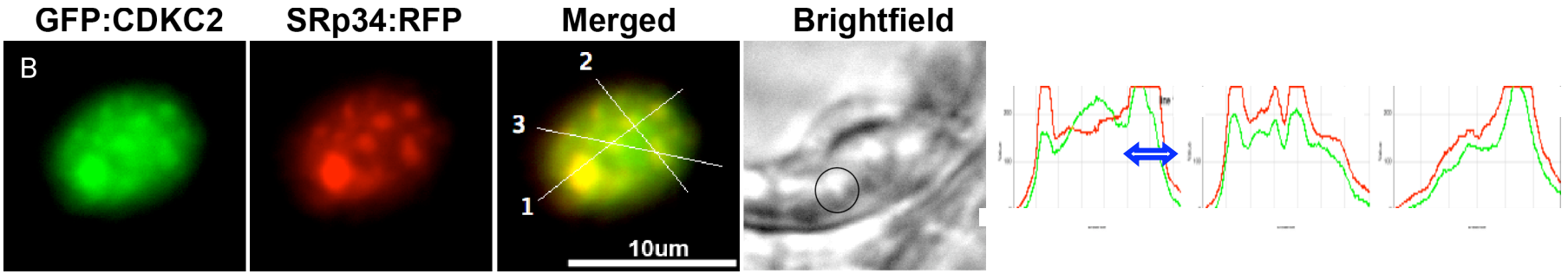

GFP:CDKC2
SRp34:RFP
Merged
Brightfield
B
2
3
1
10um

Figure 5.5 Co-expression analysis of GFP:CDKC2 and Cyp64:RFP in Arabidopsis nuclei.

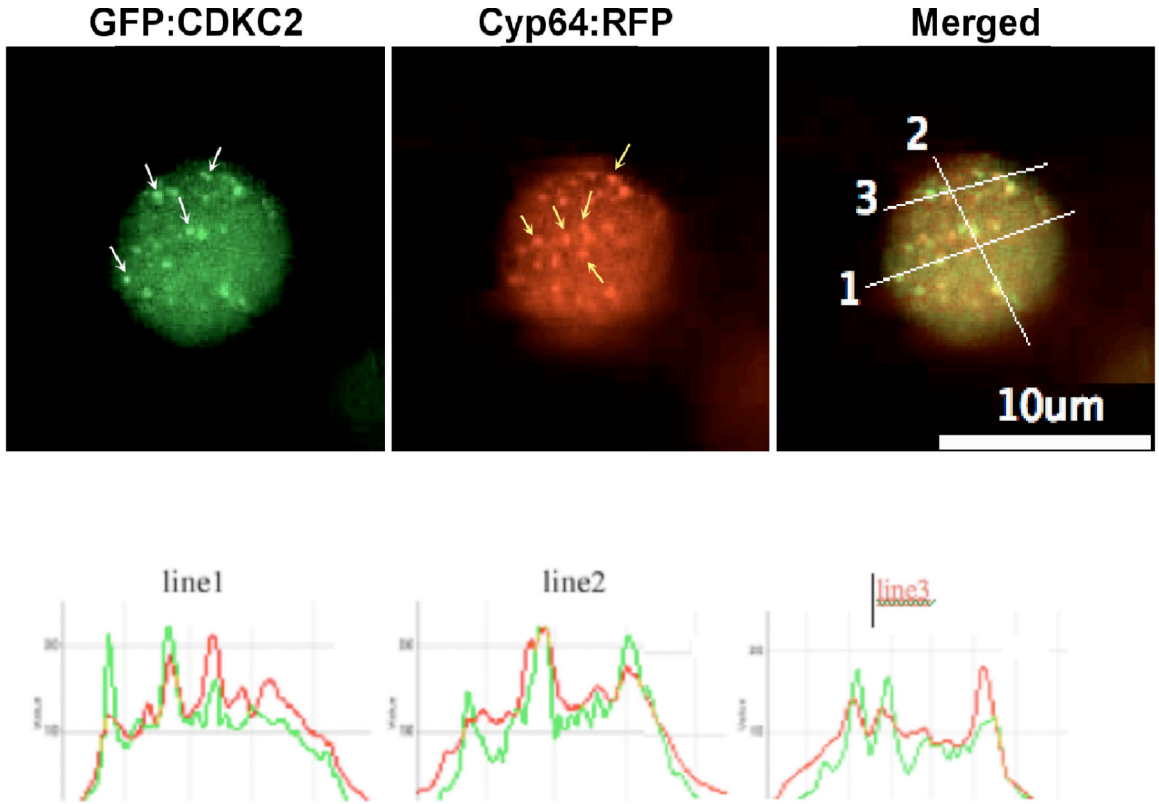
(A) Arabidopsis Col-0 cells co-transformed with constructs overexpressing fluorescent-tagged fusions of CDKC2 and Cyp64.

Maximum projection images of selected z-sections of a single nucleus are shown. On the left panel, the GFP channel is shown; on the centre the RFP channel is shown and on the right a merged image of two channels is shown**.** At the bottom of the panel, GFP and RFP profile of three transects drawn through the nucleus of the merged image are shown. White and yellow arrows show unique spots present in GFP:CDKC2 and Cyp64:RFP, respectively.

(B) Localisation profiles of GFP:CDKC2 and Cyp64:RFP protein fusions when expressed alone in cell cultures.

Scale bars: 10 μm

(A)


GFP:CDKC2
Cyp64:RFP
Merged
2
3
1
10um
line1
line2
line3

(B)

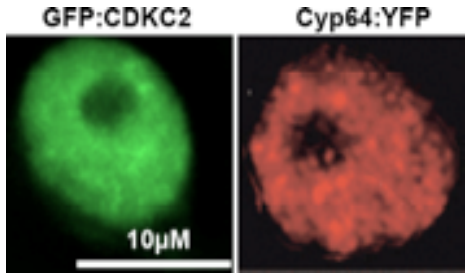
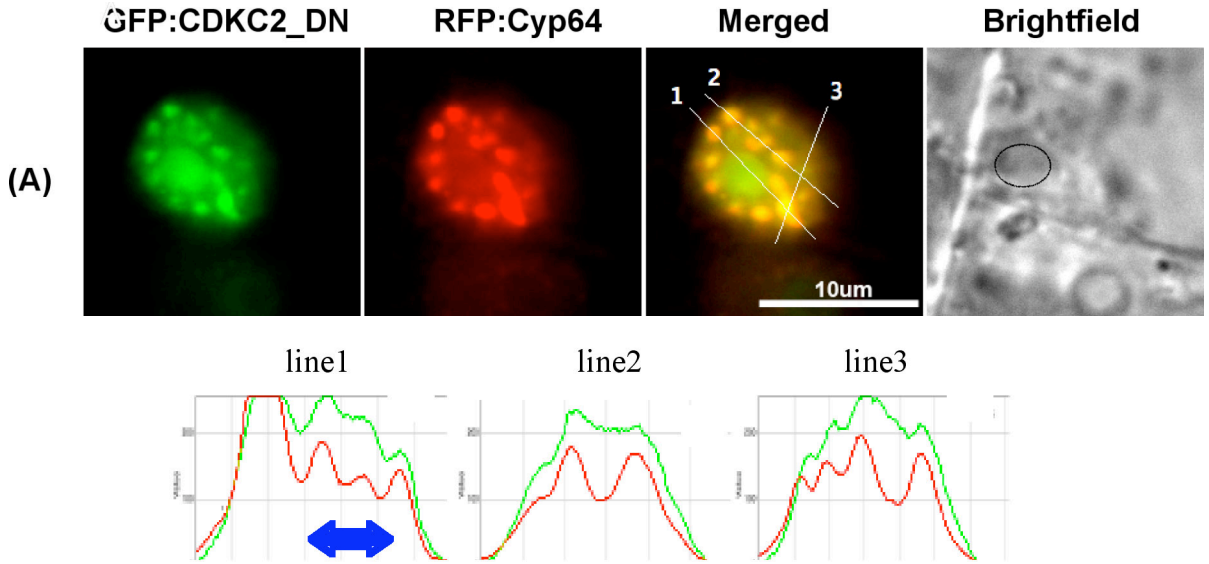
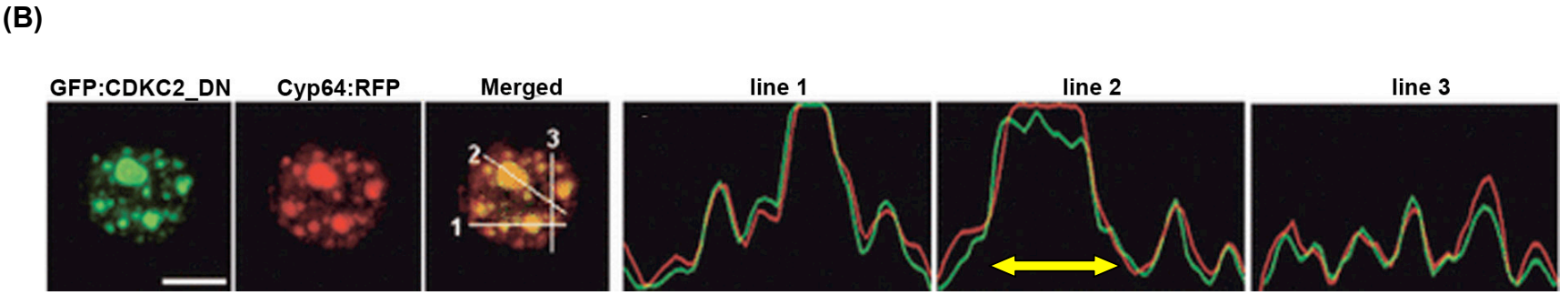

GFP:CDKC2
Cyp64:YFP
10μM

Figure 5.6. Co-expression analysis of GFP:CDKC2_DN and Cyp64:RFP in Arabidopsis nuclei.

(A) Majority pattern: GFP:CDKC_DN and Cyp64-RFP fusion proteins co-localise in the nucleoplasm within small and large speckles. On the right of the multiple panels, GFP and RFP channel profiles for three arbitrary transects across the merged image panel, are shown. The nucleolus is circled in the brightfield image.
(B) Minority pattern: As in (A) but this time the two fusion proteins co-localise in the nucleolus as well. On the right of the combined panel GFP and RFP profiles are shown, as described in (A). Beneath RGB profiles, double-headed arrowhead indicates when the respective transect passes through the nucleolus. Scale bars: (A) 10 μm and (B) 5 μm .

In both (A) and (B), a single plane passing through the nucleolus is used for assessment of protein co-localisation


(A)
GFP:CDKC2_DN
RFP:Cyp64
Merged
Brightfield
1
2
3
10um
line1
line2
line3


(B)
GFP:CDKC2_DN
Cyp64:RFP
Merged
line 1
line 2
line 3
1
2
3

5.5 CDKC2 co-localises with RNA Polymerase II in Arabidopsis cell cultures

CDKC2/CycT complex may represent the plant homolog of mammalian CDK9/CycT complex (Cui *et al*., 2007; Barocco *et al*., 2003), which is the catalytic subunit of the transcription factor pTEF-b and is responsible for phosphorylating the C-terminal domain of the large subunit of RNA PolII (Chapter 1). Even though it has been demonstrated, both *in vivo* and *in vitro*, that CDK9/CycT is responsible for this phosphorylation event, no localisation data are available showing the extent of co-localisation between CDKC2 and RNA PolII. A BY2 line stably expressing GFP:CDKC2 was used to test the degree of co-localisation between the two proteins. PolII localisation was determined by indirect immunofluorescence in fixed BY2 cells (Chapter 2; Section 2.9). The primary antibody used recognised the CTD of the RNA PolII (Abcam) and the secondary antibody was coupled to Alexa Fluor 545.

Extensive co-localisation between the two proteins in the nucleus was observed, whereas both were less abundant in the nucleolus (Figure 5.7). The strong colocalisation along with evidence of direct protein-protein interaction, supports the idea that CDKC2 and RNA PolII function in the same process. However, RNA PolII accumulated in small patches inside the nucleolus whereas accumulations of CDKC2 within the nucleolus were very much finer that may represent background signal. This reduced intensity of the GFP signal is probably the result of cell fixation.

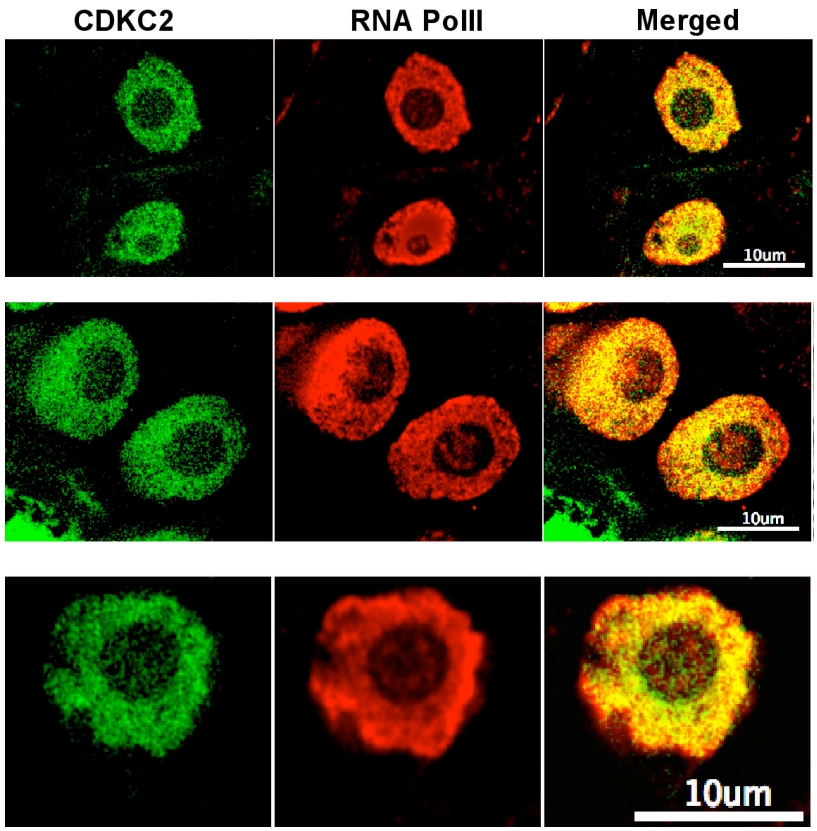


Figure 5.7. Co-localisation analysis between GFP:CDKC2 and RNA PolII.

Tobacco BY-2 cells expressing GFP:CDKC2 were fixed and permeabilised, before being incubated with a rabbit RNA PolII antibody. A rabbit Alexa Fluor 546 secondary antibody was used for detecting the bound primary antibody. Alexa Fluor 546 signal (RNA PolII) was visualised using a TRITC (Tetramethyl Rhodamine Iso-Thiocyanate) filter on a CCD microscope. The first column shows the GFP channel, the second column shows the TRITC channel and the third column shows the merged image between the two channels.

5.6 Investigation of GFP:CDKC2 co-localization with components of the exon-junction complex

Co-localisation of CDKC2 with spliceosomal components suggests an involvement of this kinase in pre-mRNA processing. Since EJC proteins are also part of this process, I assessed whether CDKC2 co-localises with selected EJC proteins. I obtained 4 constructs containing C-terminal RFP fusions of Arabidopsis EJC proteins [RNA-binding protein S1 (RNPS1), Magoh-nashi homolog (Magoh), ALY4 and elongation initiation factor 4A-3 (eIF4A-3), provided by Ali Pendle, Norwich; Pendle *et al*., 2005; Koroleva *et al*., 2009] and used them to transform Arabidopsis Col-0 cell cultures.

5.6.1 Co-expression of GFP:CDKC2 and RNPS1:RFP fusion proteins

Localisation of RNPS1:RFP protein fusion, when expressed alone, is nuclear and appears to have a "speckled" distribution (Figure 5.8A). This profile is in agreement with the reported localisation of an RNPS1-GFP protein fusion in Arabidopsis cell cultures (Pendle *et al*., 2005). Also, in some cells RNPS1 seemed to be more diffusely localised throughout the nucleoplasm, as shown in Figure 5.8A. When co-expressed with GFP:CDKC2, the proteins did not co-localise extensively (Figure 5.8B), except for a few spots inside the nucleus. Assessment of protein co-localisation along arbitrary transects is shown in Figure 5.8B. Also co-transformation of RNPS1-RFP fusion with the GFP:CDKC2_DN fusion did not lead to extensive co-localisation, although occasional distinct speckles that clearly contained both proteins were observed in the nucleoplasm (Figure 5.9). Mutant kinase was concentrated in fewer and larger spots and gradually translocated into the nucleolus, as observed previously (Figure 5.3).

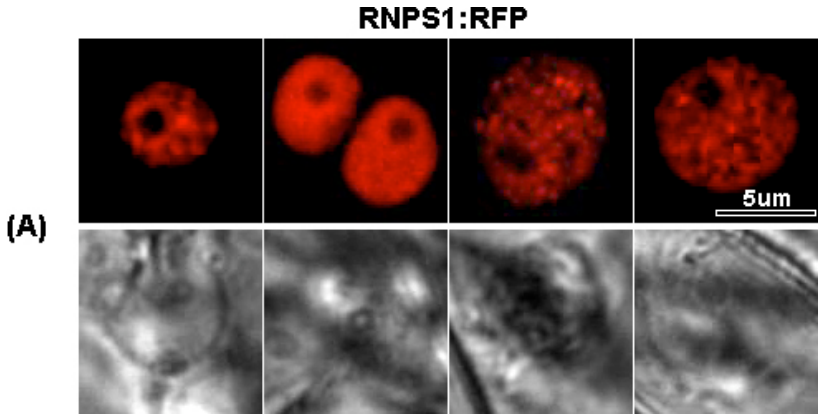
Figure 5.8 Co-expression analysis of GFP:CDKC2 and RNPS1:RFP fusion proteins in Arabidopsis nuclei.

(A) Col-0 cells transiently expressing RNPS1-RFP protein only. The upper row of panels depicts localisation profiles of the protein fusion in different Arabidopsis nuclei. The lower row shows corresponding brightfield images.

(B) Col-0 cells transiently co-expressing GFP:CDKC2 and RNPS1:RFP protein fusions. A representative nucleus is shown (inset, brightfield image of the nucleus). On the right of the panel, green and red channel profiles are shown, corresponding to arbitrary lines drawn along the merged image. Double-headed arrows show the RGB profile of a line when it passes through the nucleolus.

In both (A) and (B), a single optical plane passing through the nucleolus is showed.

Scale bars: 5 μm


RNPS1:RFP
(A)
5um

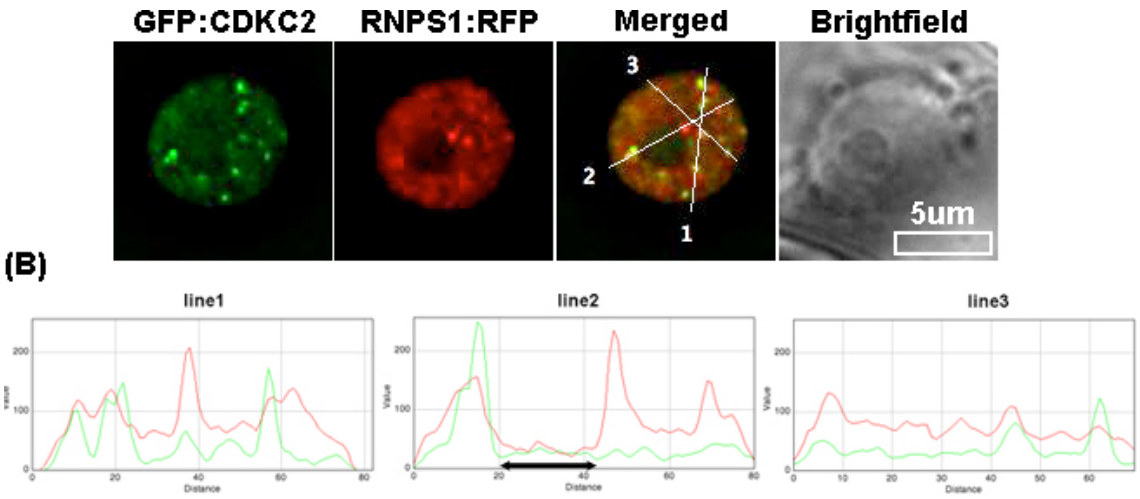
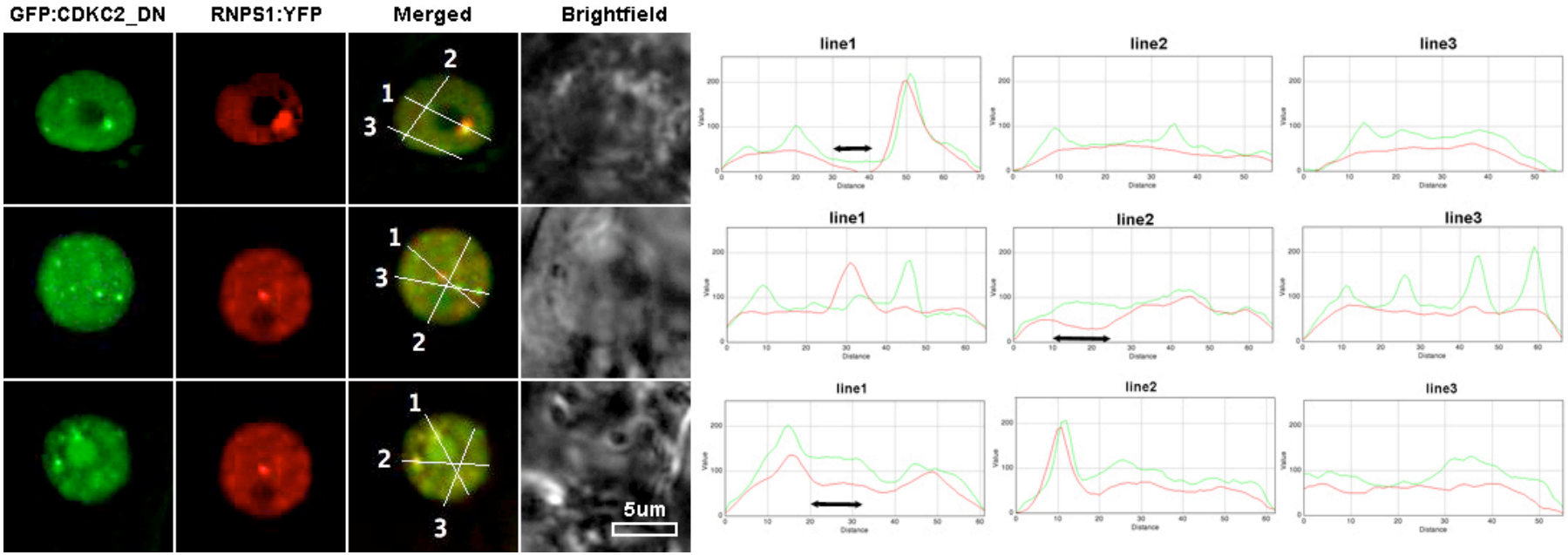

GFP:CDKC2
RNPS1:RFP
Merged
Brightfield
3
2
1
5um
(B)
line1
line2
line3
Distance
Value

Figure 5.9 Co-expression analysis of GFP:CDKC2_DN and RNPS1:RFP fusion proteins in Arabidopsis nuclei.

All three composite panels show localisation of fusion proteins in different nuclei and on their right are corresponding green and red channel profiles, showing the extent of co-localisation. For all the three different panels, a single optical plane passing through the nucleolus is used for assessment of protein co-localisation

Cell cultures were kept on the microscope slide, covered with a gas-impermeable slide, and observed at different time points. Images represent maximum projection of selected optical Z-sections of each nucleus. Images from the top panel were acquired at the beginning of the observation; images at the middle panel were collected at 30 min and images at the bottom panel after 50-60min of starting the observation.

Double-headed arrows indicate RGB profiles when the respective lines pass through the nucleolus. Scale bar: 5 μm


GFP:CDKC2_DN
RNPS1:YFP
Merged
Brightfield
1
2
3
5um
line1
line2
line3
Value
Distance

5.6.2 Co-expression analysis of GFP:CDKC2 and Magoh:RFP fusion proteins in Arabidopsis nuclei

Magoh:RFP was co-transformed with GFP:CDKC2 into Col-0 cell cultures. Both protein fusions showed extensive co-localisation both in the nucleoplasm and in nuclear speckles (Figure 5.10A). However, in some cells co-expressing the two protein fusions, Magoh did not show any obvious localisation in speckles but had a more diffuse nucleoplasmic localisation profile (Figure 5.10B). In this case the degree of co-localisation with GFP:CDKC2 was reduced.

Co-localisation between Magoh and CDKC2 was retained even after the abolishment of the kinase activity of the latter. The only difference was that the majority of Magoh:RFP did not translocate into the nucleolus over time, as for GFP:CDKC2_DN, but that which did co-localised with the inactive kinase in small spots (Figure 5.10; middle panel). In between 5 and 8% of the cells observed (three independent experiments) Magoh was diffusely distributed in the nucleoplasm without any localised concentration in speckles (Figure 5.11; bottom panel).

The above co-localisation data suggest that CDKC2 can, but does not always, co-localise with Magoh:RFP in speckles and in the rest of the nucleoplasm. Even though both protein fusions still colocalised, loss of CDKC2 kinase activity changed the degree of co-localisation due to gradual translocation of the inactive kinase into the nucleolus

Figure 5.10 Localisation profiles of GFP:CDKC2 with Magoh:RFP fusion proteins in *Arabidopsis* nuclei

Col-0 cells were transiently transformed with GFP:CDKC2 and Magoh:RFP. Representative nuclei are shown. Images were generated after projecting different slices into a single plane. Below its panel RGB profiles of 3 arbitrary lines, drawn through nucleus, are shown. For all the three different panels, a single optical plane passing through the nucleolus is used for assessment of protein co-localisation

(A) In the majority of cells the two fusion proteins co-localised in nuclear speckles.

(B) In some cells Magoh fusion protein had a diffuse nuclear localisation profile without an apparent co-localisation signal with GFP:CDKC2.

Double-headed black arrows indicate the RGB profiles of GFP and RFP when the respective line passes through the nucleolus.

Scale bar: 10 μm

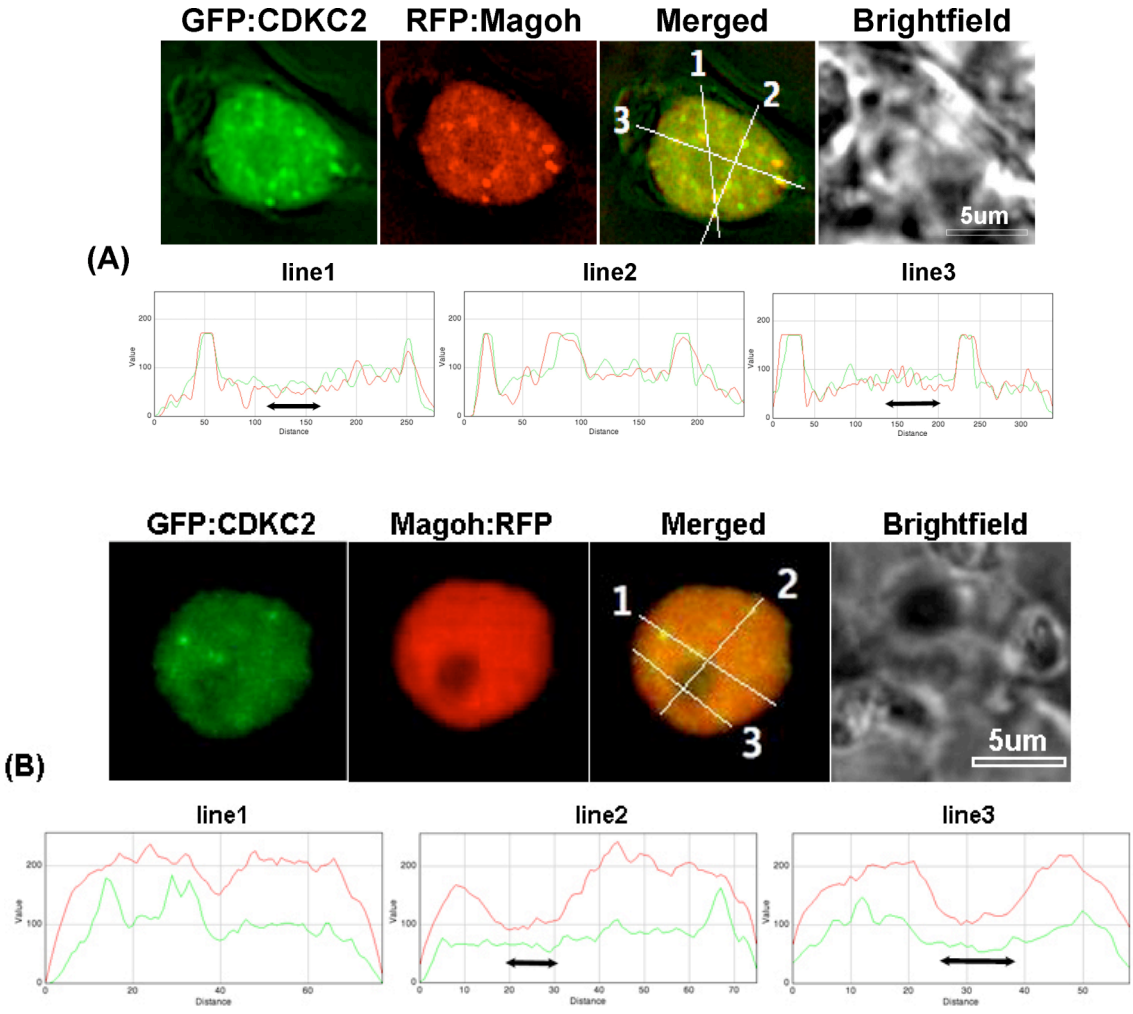
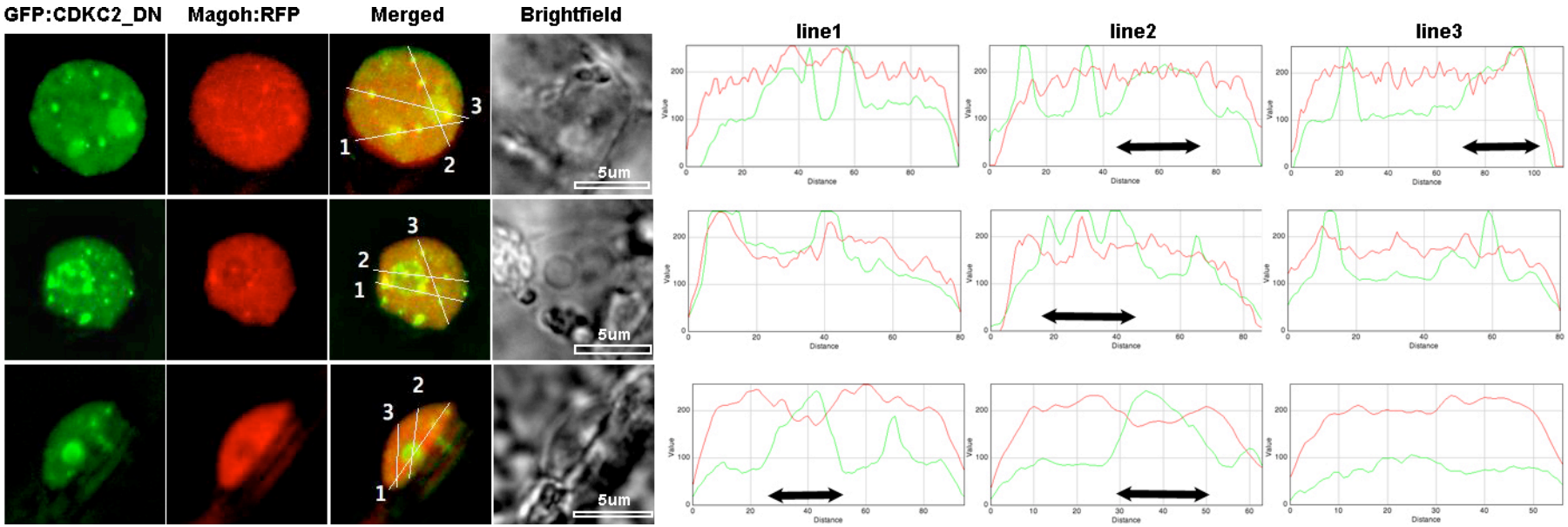

GFP:CDKC2
RFP:Magoh
Merged
Brightfield
1
2
3
5um
(A)
line1
line2
line3
Value
Distance
GFP:CDKC2
Magoh:RFP
Merged
Brightfield
1
2
3
5um
(B)
line1
line2
line3
Value
Distance

Figure 5.11 Co-expression analysis of GFP:CDKC2_DN and Magoh:RFP fusion proteins in *Arabidopsis* nuclei.

Col-0 cells were transiently transformed with GFP:CDKC2_DN and Magoh:RFP. Representative nuclei are shown from three different timepoints. The top panel show typical images from a nucleus taken immediately (within 5 minutes) after cells were transferred to a slide and covered with a gas-impermeable coverslip. Middle panels present a typical cell 20 min later and the bottom panel 1 hour later after. Arrowheads in the Magoh:RFP images indicate spots that co-localise with the ones of GFP:CDKC2_DN in the merged image. Transects through the images are shown on the right. Double-headed black arrows indicate the RGB profiles of GFP and RFP when the respective line passes through the nucleolus.
For all the three different panels, a single optical plane passing through the nucleolus is used for assessment of protein co-localisation

Scale bar: 5 μm


GFP:CDKC2_DN
Magoh:RFP
Merged
Brightfield
5um
line1
line2
line3
Value
Distance

5.6.3 Localisation analysis of GFP:CDKC2 and GFP:CDKC2_DN fusion proteins with Aly4:RFP

An important component of the mammalian EJC is the Aly4/REF protein. It interacts with the Y14 protein, another component of the EJC, and probably Aly/REF recruits the nuclear export receptor TAP/p15 (Le Hir *et al*., 2001). Aly/REF also interacts with the Y14 protein, a strong interactor of Magoh in *Droshopila melanogaster* (Zhao *et al*., 2000). A parallel interaction of Magoh with the TAP/p15 suggested that Aly4/REF, Magoh and Y14 are important for nuclear export of mature mRNAs (Shyu and Wilkinson, 2000). The *Arabidopsis* Aly4-GFP protein fusion concentrates mainly in the nucleolus but also shows a speckled-like distribution in the nucleoplasm (Pendle *et al*., 2005).

In total, 20 cells were documented in three independent experiments. When expressed alone, the Aly4 fusion localises mainly in the nucleolus and shows a generally diffuse nucleoplasmic localisation (Figure 5.12B). Nucleoplasmic distribution of Aly4:RFP did not mirror exactly published localisation profile when expressed as a GFP, suggesting that the identity of the fluorescent tag effects protein subcellular localisation.

When co-expressed with GFP:CDKC2, Aly4:RFP showed the same localisation profile with that observed when expressed alone (Figure 5.12B). Moreover the degree of co-localisation was low, as could be judged by the RGB profiles resulting from the merged images (Figure 5.12B, bottom). More than 20 nuclei were documented in three independent experiments, and localisation profiles for both fusion proteins were the same as in those experiments when each of protein fusions was expressed alone.

Minimum co-localisation was also observed when Aly4:RFP was co-expressed with GFP:CDKC2_DN in cell cultures (Figure 5.13). When cells were incubated on the slide under a coverslip, and imaged every 20 minutes, GFP:CDKC2_DN translocated into the nucleolus, exhibiting the same behaviour as when expressed alone (compare with Figure 5.3). On the other hand, localisation profile of Aly4:RFP did not appear to change within the same amount of time.

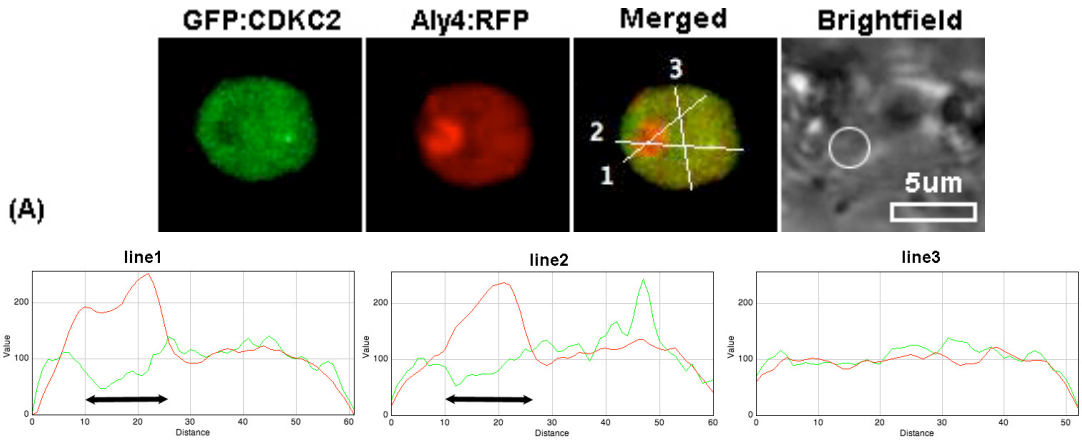
Figure 5.12 Co-expression analysis of GFP:CDKC2 and Aly4:RFP fusion proteins is *Arabidopsis* nuclei..

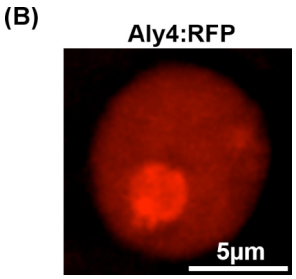
(A) *Arabidopsis* cells co-expressing GFP:CDKC2 and Aly44:RFP as indicated. Nuclear co-localisation profiles of a representative nucleus are shown. Below the multi –panel are the RGB prolies of three lines drawn along the merged image. Double-headed black arrows indicates the position of the nucleolus in the RGB profiles. The nucleolar position is circled in the brightfield image.

(B) Localisation profile of Aly44:RFP fusion protein when expressed alone in *Arabidopsis* cell cultures.

In both (A) and (B), a single optical plane passing through the nucleolus is used for protein localisation or assessment of protein co-localisation

Scale bars: 5 μm .


(A)
GFP:CDKC2
Aly4:RFP
Merged
Brightfield
3
2
1
5um
line1
line2
line3
Value
Distance


(B)
Aly4:RFP
5μm

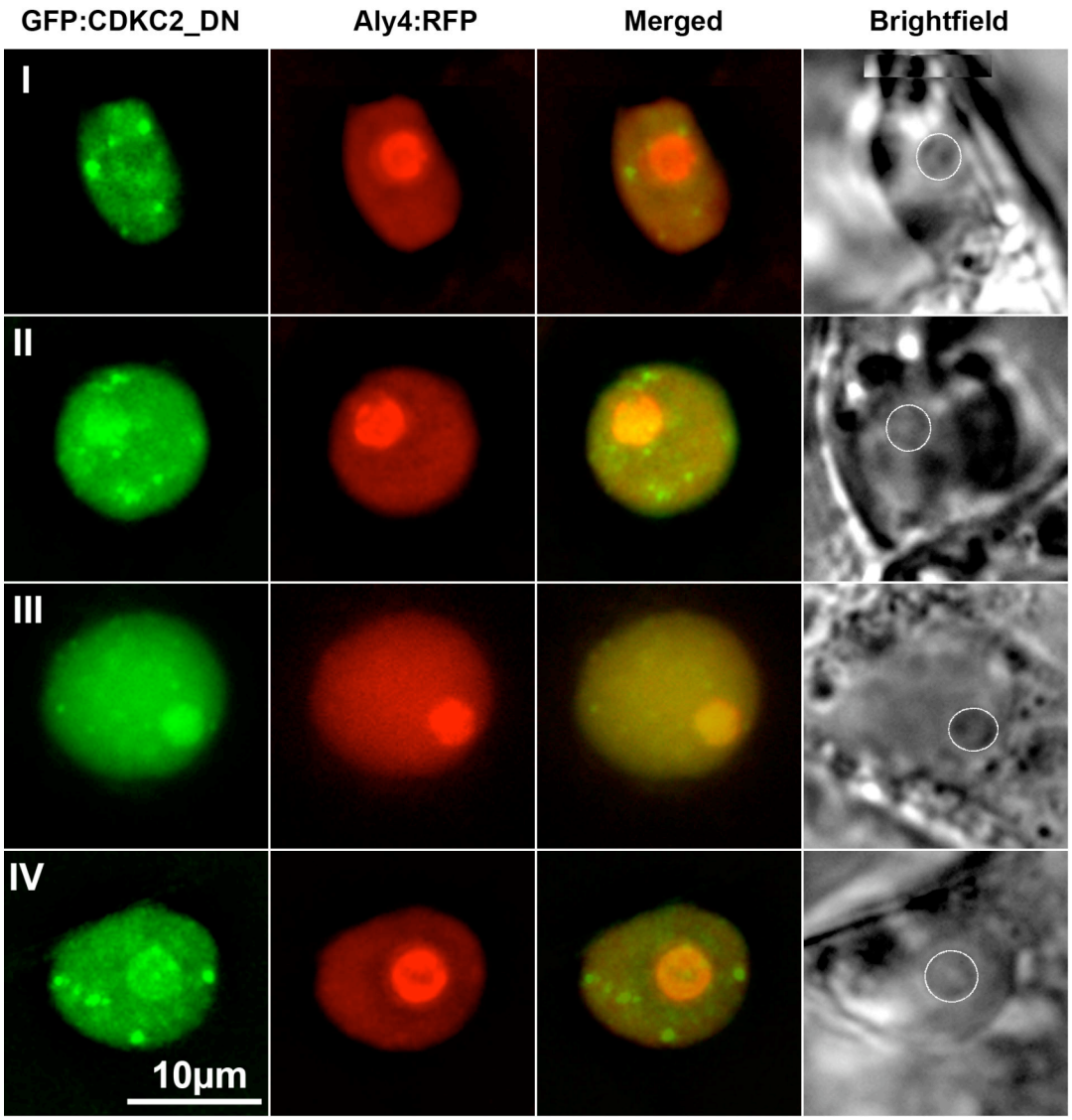


Figure 5.13 Co-localisation of Aly44:RFP with GFP:CDKC2_DN in *Arabidopsis* cell cultures.

Images of a single plane passing through the nucleolus from four different nuclei, taken at different times of incubation on the slide. One image was taken at the beginning of the experiment (I), 35 min later (II), 50 min later (III) and 1.5 hour (IV). Translocation of GFP:CDKC2_DN protein fusion into the nucleolus remains unaffected by the presence of Aly4:RFP. Nucleolus in brightfield images is circled. Scale bar: 10 µm .

5.6.4 Localisation analysis of GFP:CDKC2 and GFP:CDKC2_DN fusion proteins with eIF4A-3:RFP

Co-expression of wild type and mutant CDKC2 with eIF4A-3:RFP protein fusion showed overlapping localisation profiles, especially in terms of nuclear speckle distribution (Figure 5.14). When expressed alone, eIF4A-3 had a diffuse nuclear localisation pattern with a small amount of the protein found in the nucleolus (Figure 5.14A). When co-expressed with GFP:CDKC2, eIF4A-3:RFP maintained it a diffuse nuclear localisation profile, but now was localised into nuclear speckles as well (Figure 5.14B). Speckled distribution of eIF4A-3 matched that of GFP:CDKC2, showing extensive co-localisation. On the other hand, when co-expressed with the mutant GFP:CDKC2_DN, eIF4A-3:RFP re-acquires a diffuse nuclear localisation profile and does not accumulate to nuclear speckles (Figure 5.14C). Localisation of mutant CDKC2 in the presence of eIF4A-3 is similar to that when expressed alone; diffuse localisation in the nucleoplasm and concentration into a small number of relatively large spots (Figure 5.14C).

Figure 5.14. Co-expression analysis of GFP:CDKC2 and GFP:CDKC2_DN fusion proteins with eIF4A-3:RFP.

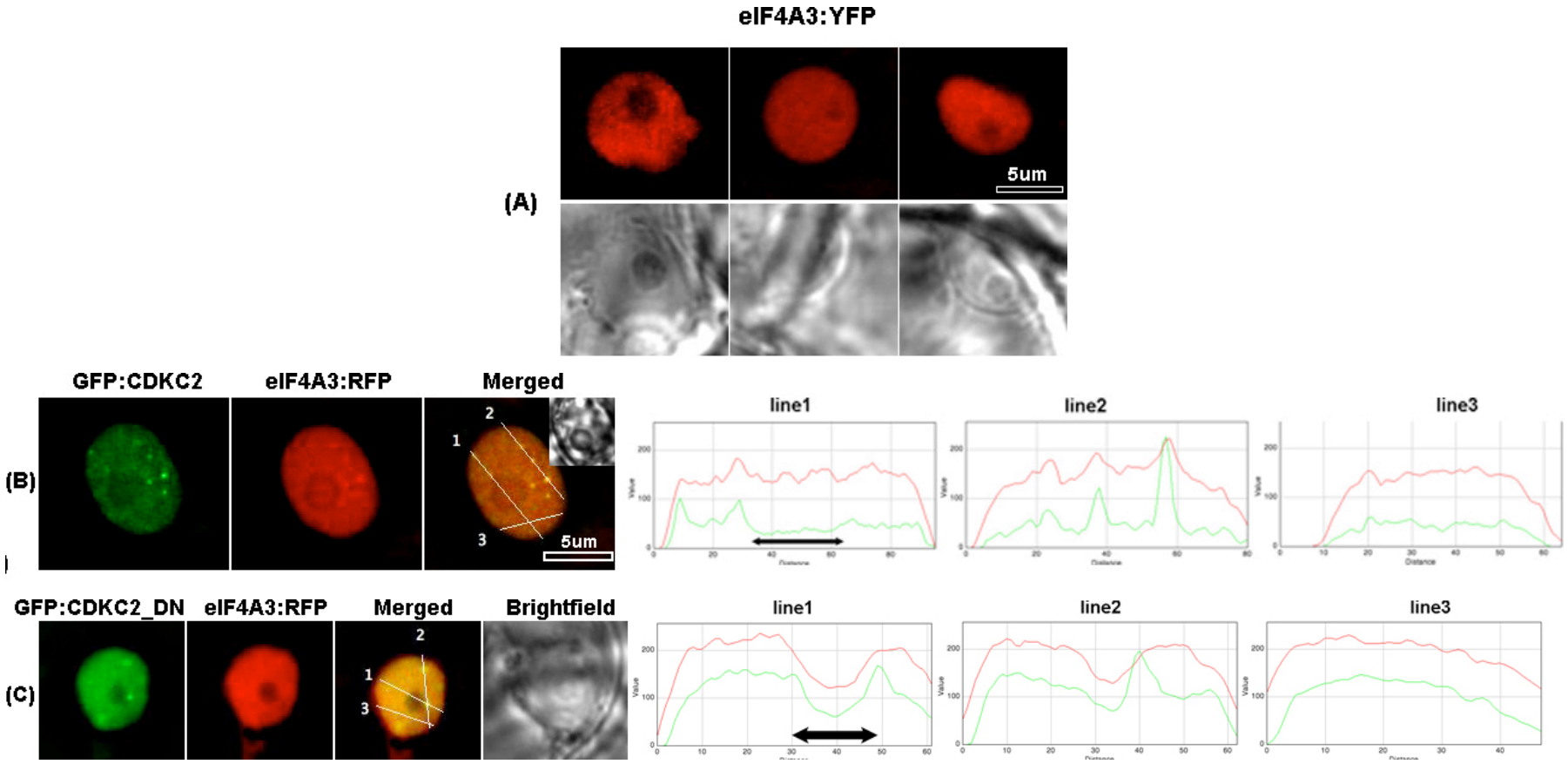
(A) Localisation profile of eIF4A-3 fusion protein in *Arabidopsis* nuclei with corresponding brightfield images at the lower row of composite panel.

(B) Co-localisation analysis of GFP:CDKC2 and eIF4A-3:RFP proteins in a representative nucleus.

(C) Co-localisation analysis of GFP:CDKC2_DN and eIF4A-3:RFP in a representative nucleus.

In all the panels a single optical plane passing through the nucleolus is used for protein localisation or assessment of protein co-localisation

Next to panels in (B) and (C) are the green and red channel profiles of the arbitrary lines drawn across the merged image of each panel. Double-headed black arrows indicate the part of RGB profile when line crosses the nucleolus. Scale bars: 5 μm


eIF4A3:YFP
(A)
5um
GFP:CDKC2
eIF4A3:RFP
Merged
(B)
1
2
3
5um
line1
line2
line3
GFP:CDKC2_DN
eIF4A3:RFP
Merged
Brightfield
(C)
1
2
3
line1
line2
line3

5.7 *In planta* analysis of GFP:CDKC2 protein fusion localisation

Determination of protein localisation in cell cultures provides us with important information on the subcellular distribution of a protein, particularly during the cell cycle. However, cell cultures grow outside the context of a multicellular organism, where cells acquire different developmental fates, leading to the formation of distinct tissues (Lau *et al*., 2010; Lodha *et al*., 2008). Proteins that acquire different localisation profiles in different tissues may reflect the role of the tissue they reside in. In this aspect, the suggested role of CDKC2 in transcription and/or splicing (Cui *et al*., 2007; Kitsios *et al*., 2007) may affect its localisation in the nucleus of cells that have different rates of transcription/splicing, i.e. meristematic versus differentiated tissues. Also direct observation of the localisation of a protein *in planta* can reveal clues to novel roles that would not be apparent in a cell culture system.

In this section I present data for CDKC2 and CDKC2_DN protein fusion localisation in plants. Both proteins were either expressed transiently as GFP fusions in *Arabidopsis* Col-0 cotyledons or stably over-expressed in Col-0 plants as YFP fusions. Figure 5.15 shows representative cell nuclei from *Arabidopsis* cotyledons transiently expressing GFP:CDKC2. The protein fusion shows nuclear localisation and concentration of the signal in nuclear speckles. This localisation profile matches the profile observed in cell cultures, suggesting that the protein distribution seen in cell cultures is essentially the same as the one observed in differentiated tissues. When GFP:CDKC2_DN was expressed in *Arabidopsis* cotyledons, the localisation profile also was similar to that observed in cell cultures. The fusion protein was diffusely distributed in the nucleoplasm and, in some cells, was also present in nuclear spots (Figure 5.16A).

Since there was extensive co-localisation between CDKC2 and Magoh proteins in cell cultures, I transiently expressed both proteins in cotyledons and observed the degree of co-localisation. In Figure 5.16B, both proteins accumulated in the nucleus but only CDKC2 localised in nuclear speckles. On the other hand, Magoh:RFP fusion appeared to have a diffused nuclear localisation without showing any signs of co-localisation with the CDKC2 protein fusion.

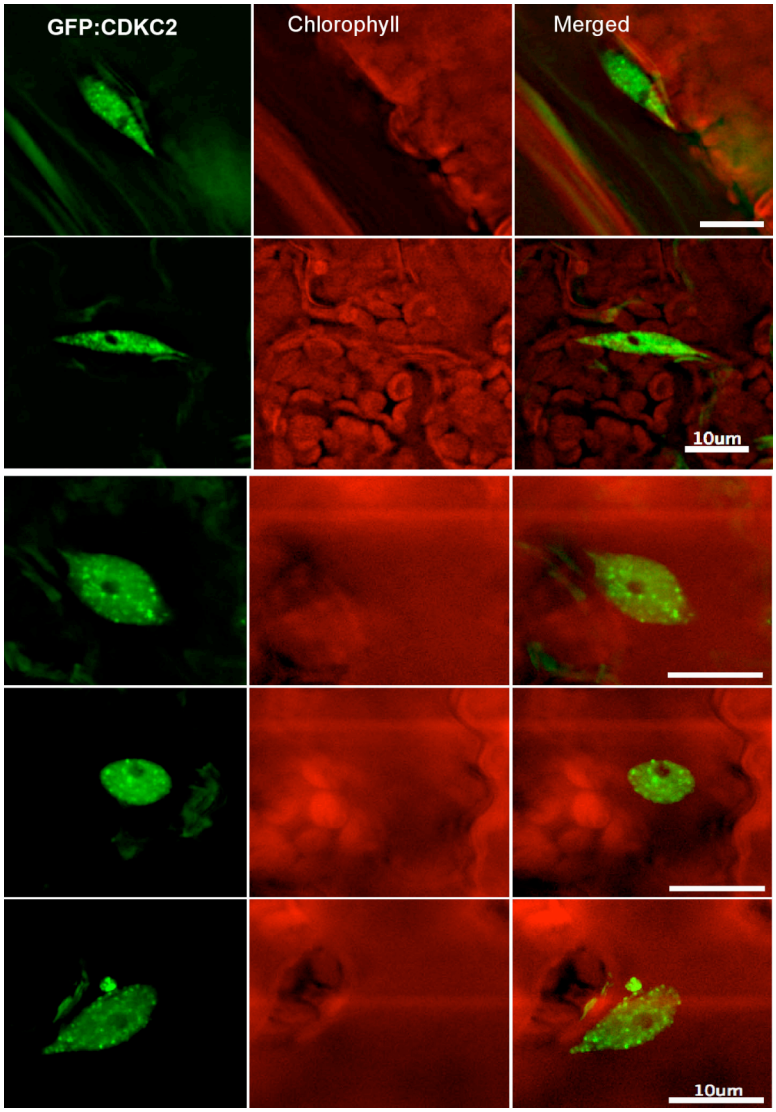


Figure 5.15 Transient expression of GFP:CDKC2 protein fusion in *Arabidopsis* cotyledons.

The composite panel illustrates representative nuclei from *Arabidopsis* cotyledon cells expressing the GFP:CDKC2 fusion. The first column shows the GFP channel, the second shows the TRITC channel and the third is a merged image of both channels. Scale bar: 10 µm

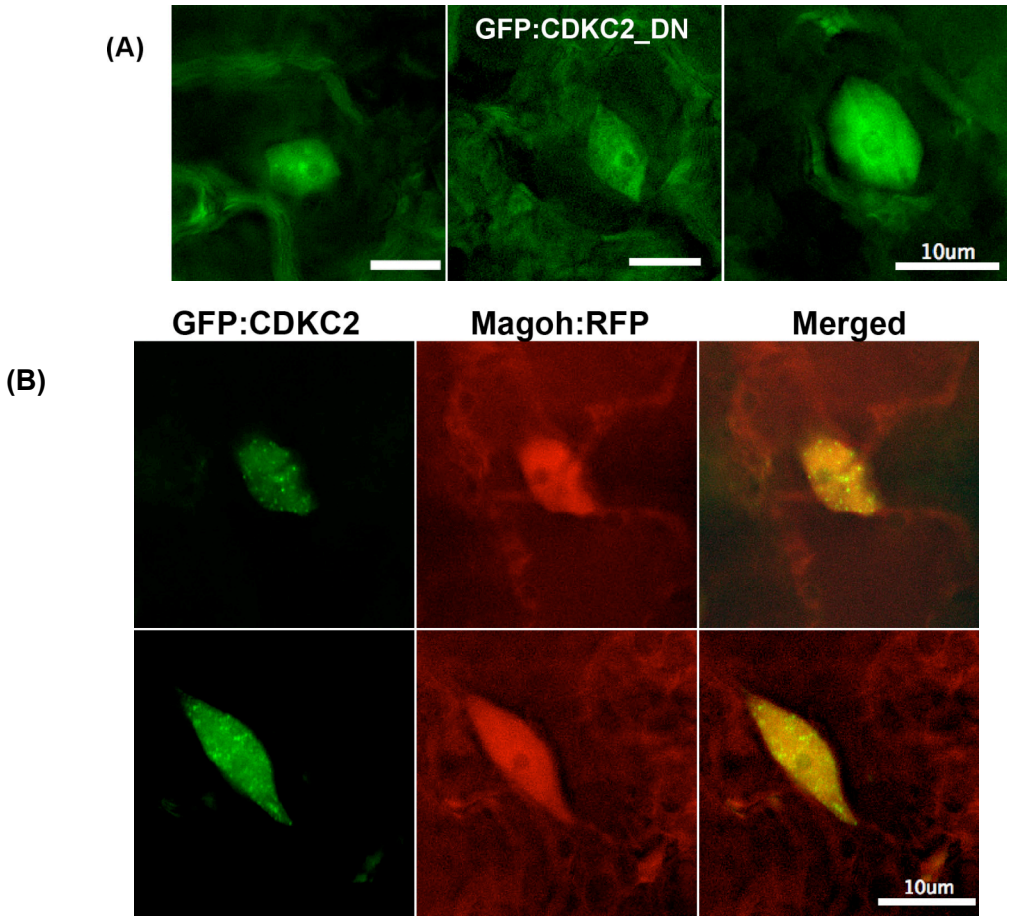


Figure 5.16 Transient expression of GFP:CDKC2_DN and co-expression of GFP:CDKC2 with Magoh:RFP in *Arabidopsis* 4-days-old cotyledons.

(A) Three representative cells are shown where the GFP:CDKC2_DN fusion localises in the nucleus.

(B) Two representative cells co-expressing GFP:CDKC2 and Magoh:RFP protein fusions. The first column of the composite panel shows the GFP channel, the second column the RFP channel and the third shows a merged imaged of the two channels. Scale bars: 10 μm .

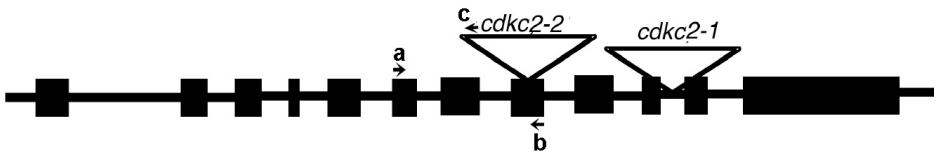
The above data show that GFP:CDKC2 localisation is similar in both undifferentiated cell cultures and in differentiated plant tissues. However, the discrepancy in co-localisation of CDKC2 and Magoh fusion proteins between cell cultures and cotyledons suggests that colocalisation is sensitive to physiological status of the cells.

5.8 Phenotypic analysis of *cdkc2* T-DNA insertion mutants and gene complementation

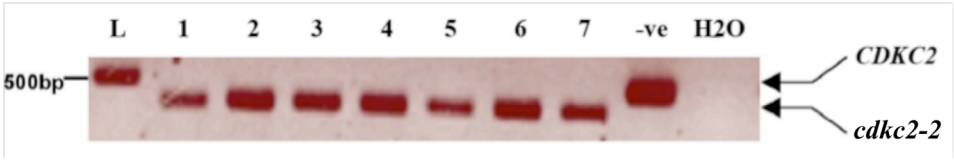
Insertion mutants in the *CDKC2* gene produced plants that were late flowering and had increased CaMV resistance (Cui *et al*., 2007). To test whether the CDKC2 protein fusion is functional, I obtained homozygous SALK T-DNA insertion mutants (SALK_029546 and SALK_149280), aiming to perform genetic complementation of the mutant plants. Predicted genotype related data was obtained from the T-DNA express website (http://signal.salk.edu/cgi-bin/tdnaexpress), SALK_149280 line (*cdkc2-1*) has an insertion in the 10th intron and the SALK_029546 line (*cdkc2-2*) has an insertion in the 8th exon (Figure 5.17A). The *cdkc2-1* line did not express the kanamycin resistance trait expected for this T-DNA, probably due to gene silencing (http://signal.salk.edu/tdna.FAQs.html), whereas the *cdkc2-2* expressed the resistance marker normally. I used PCR to confirm the presence of the T-DNA in *cdkc2-2* and confirm the homozygous status of the selected plants (Figure 5.17B). Primer pairs used for screening the hygromycin-resistant plants are given at Figure 5.17C.

I selected 3 PCR-positive plants (1, 2 and 3 at Figure 5.17) and used RT-PCR to check for the presence of *CDKC2* mRNA transcript. As shown in Figure 5.18 no detectable *CDKC2* mRNA was present in the leaves of homozygous plants whereas *CDKC2* transcript is present in Col-0.

(A)



(B)



(C)

cdkc2-2-for: 5'-ACTGCTGACTGGGCTTCACT-3'

cdkc2-2-rev: 5'-TCAGTTTTCCCAGGCAAGAT-3'

Lba1: 5' –ATGGTTCACGTAGTGGGCCATC-3'

Figure 5.17 Genotyping of T-DNA mutant plants.

(A) Position of T-DNA insertions for the *cdkc2-1* and *cdkc2-2* mutant lines. Black boxes represent exons and lines represent introns. a: *cdkc2-2*-for, b: *cdkc2-2*-rev, c: Lba1.

(B) Seven T-DNA mutant plants showing kanamycin resistance were screened to confirm that they were homozygous for the T-DNA insertion. The predicted size of the *CDKC2* PCR product is 486 bp and for the insertion line, 380 bp. L: DNA ladder; 1-7: number of each plant from *cdkc2-2* line; –ve: Col-0 geneomic DNA. For amplification of the genomic *CDKC2* I used the *cdkc2-2*-for/*cdkc2-2*-rev primer pair and for the detection of the T-DNA presence (*cdkc2-2*) I used the *cdkc2-2*-for/Lba1 primer pair

(C) Primers used to screen for the genotype of selected hygromycin-resistant plants. *cdkc2-2*-for/*cdkc2-2*-rev pair was used for the detecting the wild type gene and the *cdkc2-2*-for/Lba1 pair was used for detecting the T-DNA insertion.

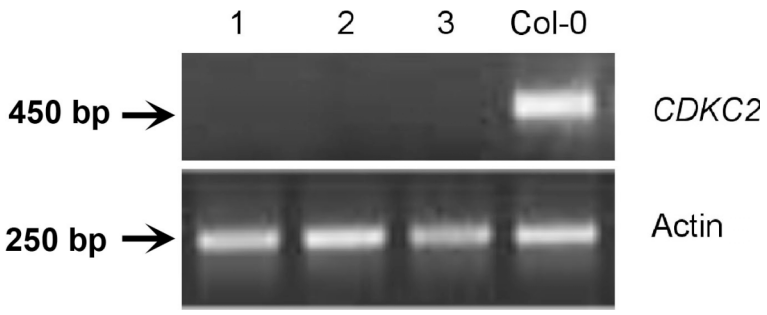


Figure 5.18 RT-PCR on selected homozygous T-DNA mutant plants.

Three *cdkc2-2* plants (1 – 3) were tested for the presence of *CDKC2* transcripts in leaves. RNA was extracted using Qiagen's RNA extraction kit and cDNA synthesis was performed using the Omniscript RT Kit from Qiagen. Actin was used as a control for using equal amounts of cDNA for the PCR. The sizes of expected PCR products are shown to the left of the gels.

I then undertook a phenotypic analysis. Towards the end of my PhD we received *cdkc2-2*/CDKC1_RNAi seeds (hereafter called *c2*/C1RNAi) from Dr. Chen at Purdue University and these were included in the analysis. Figure 5.18 shows representative plants for the three different genotypes at 8-days-old and 4-weeks-old. Plant stature of *cdkc*2-*2* mutant lines was not markedly different from that of the Col-0 plants throughout their vegetative growth, but the *c2*/C1RNAi plants were slightly smaller. Leaves of both mutant plants appeared more circularised than Col-0 (Figure 5.19 and Figure 5.20). The increase in leaf circularisation became more obvious when leaves from the three genotypes are laid in order, from the youngest to the oldest (Figure 5.20A). This altered leaf phenotype appeared to result from the decrease of the angle between the leaf lamina and the petiole. To examine whether the angle did change, I measured leaf-petiole angles from 6 mature leaves at the same developmental stage (6th

leaf), depicted in Figure 5.20A, in 7 plants of each genotype. The result was that the lamina-petiole angle in Col-0 was larger than the respective angles in *cdkc2-2* and *c2*/C1RNAi plants, at the p = 0.01 level of significance, whereas differences between the *cdkc2* and *c2*/C1RNAi plants were not statistically significant (Figure 5.13B). Thus, knocking down *CDKC2* gene affected leaf morphology, whereas this phenotype is not enhanced when both *CDKC* genes are inactivated in plants.

To verify the late flowering phenotype of *CDKC2* mutant plants, I recorded flowering times for mutant and wild plants. Indeed, *cdkc2-2* mutant plants flowered later than Col-0, with *c2*/C1RNAi flowering even later, revealing an additive effect of *CDKC1* knockdown on this physiological character. Averaging 10 plants, Col-0 flowered at 9.6 leaves, *cdkc2-2* at 19.6 leaves and *c2*/C1RNAi at 24.8 leaves (Table 5.1).

	Col-0	***cdkc2-2***	***c2*/C1RNAi**
Average flowering time (n=10)	9.6±0.52	19.6±0.66	24.8±0.98

Table 5.1 Average number of rosette leaves to flowering for Col-0, *cdkc2-2* and *c2*/C1RNAi plants.

Rosette leaves from 10 individual plants, for each of the genotypes, were counted at bolting time.

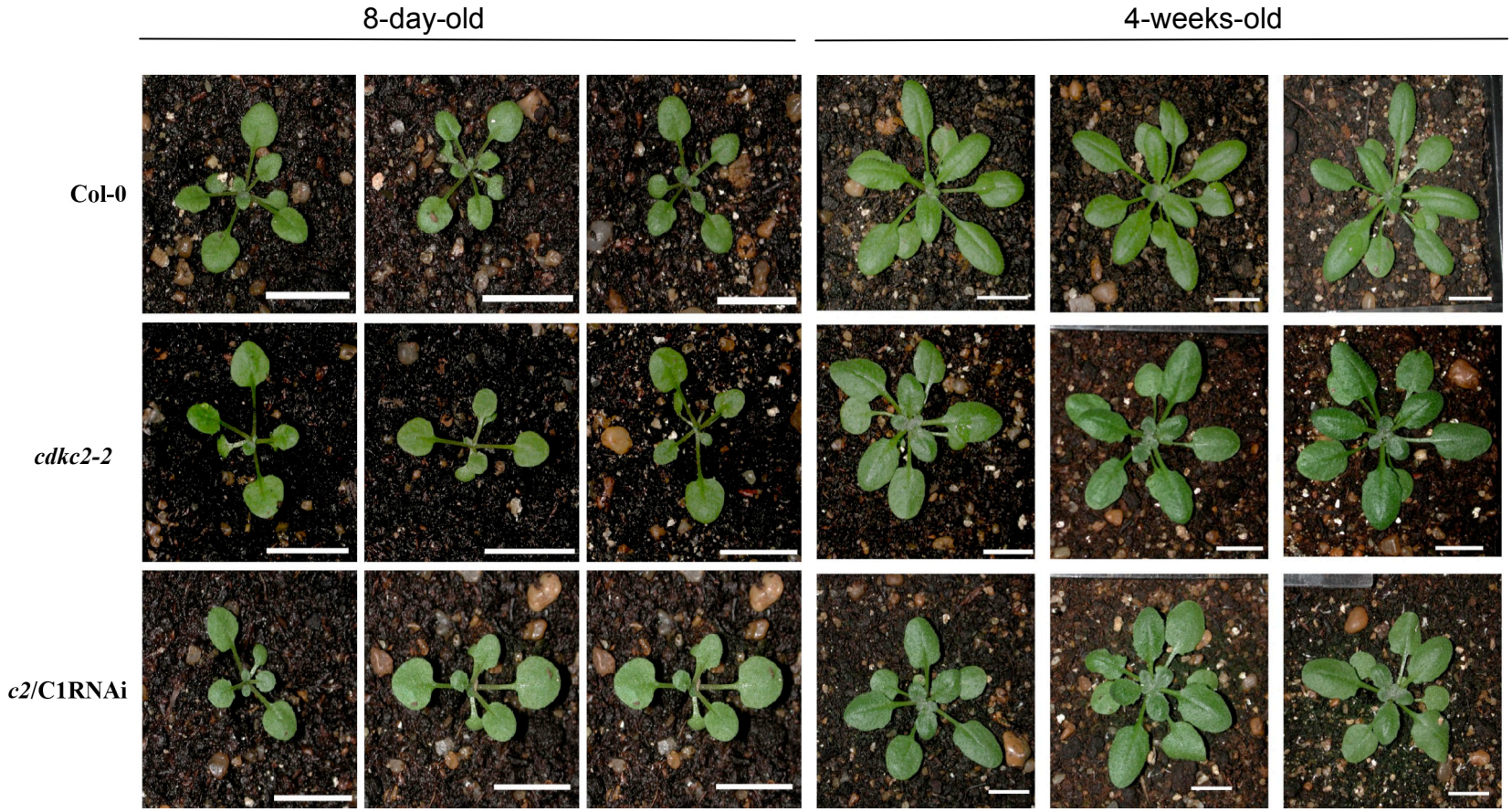


Figure 5.19 Phenotypes of Col-0, *cdkc2-2* and *c2*/C1RNAi plants.

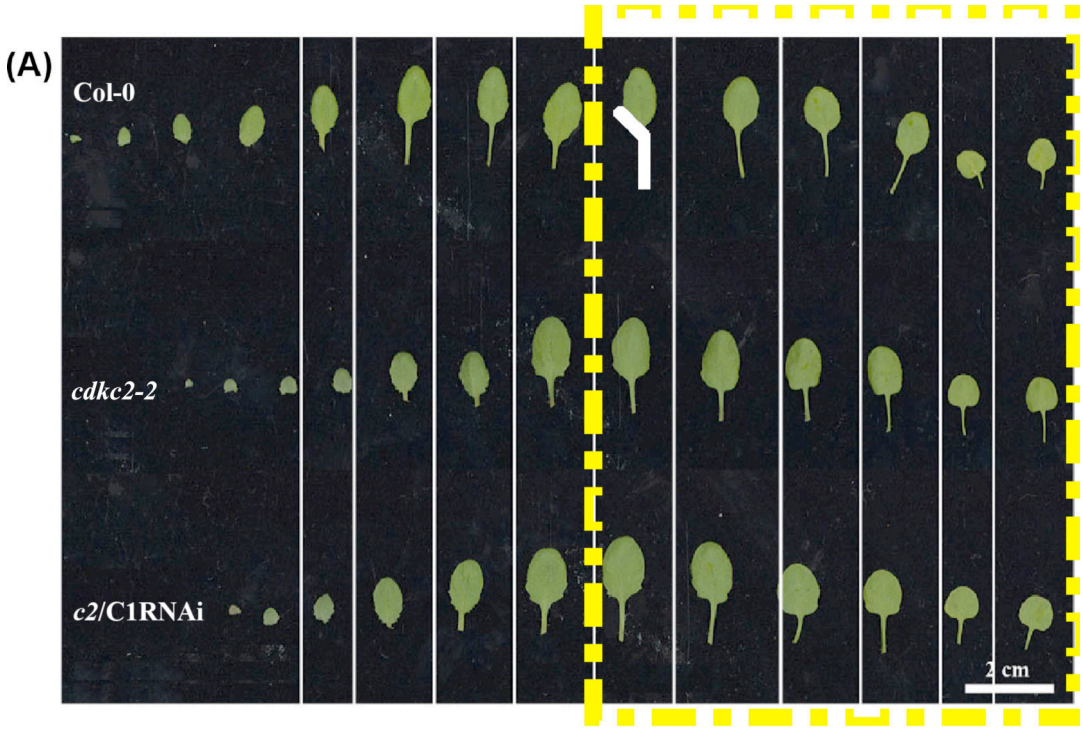
Three representative plants for each genotype are shown at 8 days (first 3 panels to the left) and 4 weeks (3 columns to the right) after germination. Scale bar = 1cm

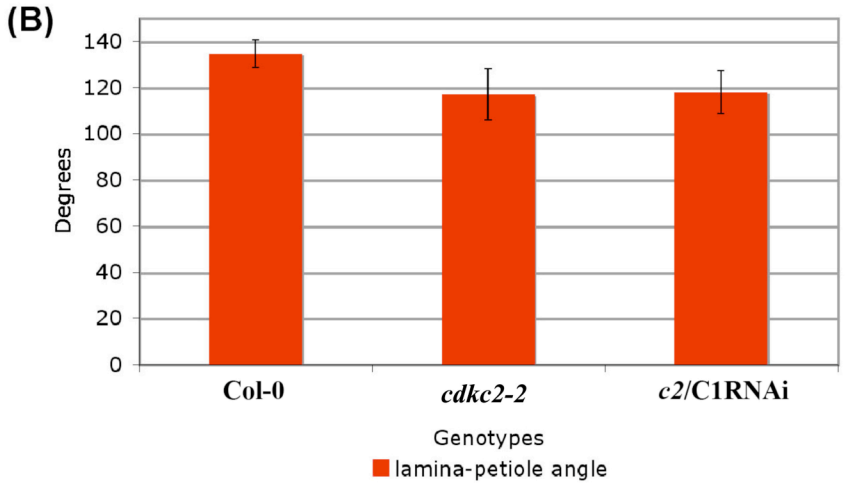
Figure 5.20 Morphology of rosette leaves of Col-0, *cdkc2-2* and *c2*/C1_RNAi plants and phenotypic analysis.

(A) Rosette leaves were collected from the three genotypes when the Col-0 plants had bolted. The oldest (seed) leaves appear on the right of the image.

(B) Phenotypic measurement of the angle created by leaf lamina and petiole [illustrated by the white bent line in (A)] in each of the three genotypes. Six plants were used from each genotype and the character was measured in the first six (oldest) leaves indicated by the yellow dashed square in (A)]. Measurements for each individual plant were averaged. Bars represented the average value of the averages from the 7 plants used for each genotype.

Scale bar = 2cm.


(A)
Col-0
cdkc2-2
c2/C1RNAi
2 cm


(B)
140
120
100
80
60
40
20
0
Degrees
Col-0
cdkc2-2
c2/C1RNAi
Genotypes
lamina-petiole angle

Scanning electron microscopy (SEM) analysis (carried out by Dr. Kim Findlay) was employed to obtain more precise information on the leaf morphology of CDKC2 mutant plants. For the analysis, we used the 5th leaf of Col-0, *cdkc2-2* and *c2*/C1RNAi plants at 12 days after leaf emergence. No apparent differences were obvious between the three genotypes (Figure 5.21) except that *c2*/C1RNAi leaves had abnormal trichomes. Most of the trichomes on *c2*/C1RNAi plants had between 4 – 5 branches instead of 3 branches present in *cdkc2-2* and Col-0.

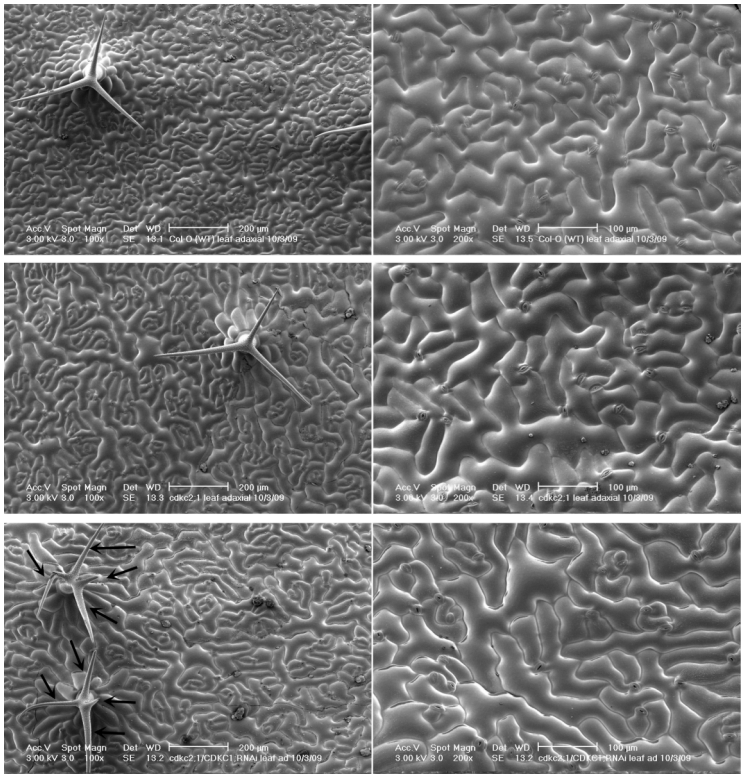


Figure 5.21 SEM analysis of the adaxial surface of leaves from Col-0, *cdkc22* and *c2*/C1RNAi plants.

On the top of the panel, two representative images are shown of the adaxial side of the 5th leaf in Col-0. In the middle of the panel, same type of images but for *cdkc2-2* and at the bottom of the panel is a leaf from a *c2*/C1RNAi plant. Samples were processed and analysed using SEM by Dr. Kim Findlay (JIC Microscopy Facility). Black arrows indicate the different trichome branches in the double mutant.

5.9 The effect of transcription and kinase inhibitors on the root growth of Col-0 and CDKC mutant plants

Transcription and kinase inhibitors both altered the subcellular distribution of CDKC2 protein in *Arabidopsis* cell cultures (Figure 5.3B; Kitsios *et al*., 2008), suggesting that they both affect CDKC function. Knocking down one or both members of the CDKC family has adverse effects on plant growth, so such lines should be more sensitive to drugs that perturb CDKC function. Therefore I tested the response of Col-0, *cdkc2-2* and *c2*/C1RNAi plants to different concentrations of DRB and roscovitine.

Figure 5.22 presents growth data from *Arabidopsis* roots in the presence of different DRB concentrations (25mM, 50mM, 100mM and 250mM). The control is also included where roots were growing in the presence of DMSO; solvent of both DRB and roscovitine. When growing in DMSO no differences with statistical significance were observed during the eight days of the experiment (Figure 5.22A). After transferring germinated seedlings of the three genotypes to media containing the respective drug concentration, root measurements were taken every two days. At day2, root growth retardation was apparent both for Col-0 and mutant phenotypes, when compared with the control sample, but no significant differences were observed between lines. However, by day4 differences between the mutant genotypes and Col-0 were observed at 25mM and 50mM of DRB. Both *cdkc2-2* and *c2*/C1RNAi showed a relative inhibition of root growth that was significant at p-value of 0.05 and 0.01 respectively, for both concentrations (Figure 5.22B). These differences between Col-0 and the mutants persisted also at day6 and day8 and became even stronger, as the differences between Col-0 and the mutant phenotypes were significant even at p-value of 0.001 (Figure 5.22B). The same trend was also observed at day8, with the addition that differences in root growth, between Col-0 and mutants, were now observed at 100mM of DRB as well at p-value of 0.05 and 0.01 for *cdkc2-2* and *c2*/C1RNAi, respectively (Figure 5.22A,B). The above data indicate that CDKC mutants are more sensitive to DRB than the wild type plants.

The use of CDK-specific inhibitor, roscovitine, had a similar effect on the root growth of the three genotypes but the degree of inhibition was stronger than that of DRB (Figure 5.23B). After two days of growth in the presence of roscovitine, root growth retardation was not statistically different between the Col-0 and the two mutant

genotypes (Figure 5.23B). As in the case of DRB, mutant plants exhibited greater reduction of root growth than Col-0 at day4 of the experiment, at 0.05mM of roscovitine (Figure 5.23A, B). At day6 and day8, root growth inhibition in mutant plants was statistically different from Col-0 p-value of 0.01, at 0.05mM and 0.5mM of roscovitine (Figure 5.23B). The difference between DRB and roscovitine is that roscovitine is a much stronger inhibitor of root growth since it exerts its action at much lower concentrations than DRB, presumably due to the wider spectrum of CDK targets affected. However, concentrations of roscovitine higher than 0.5mM appeared to be toxic for wild type and mutant plants.

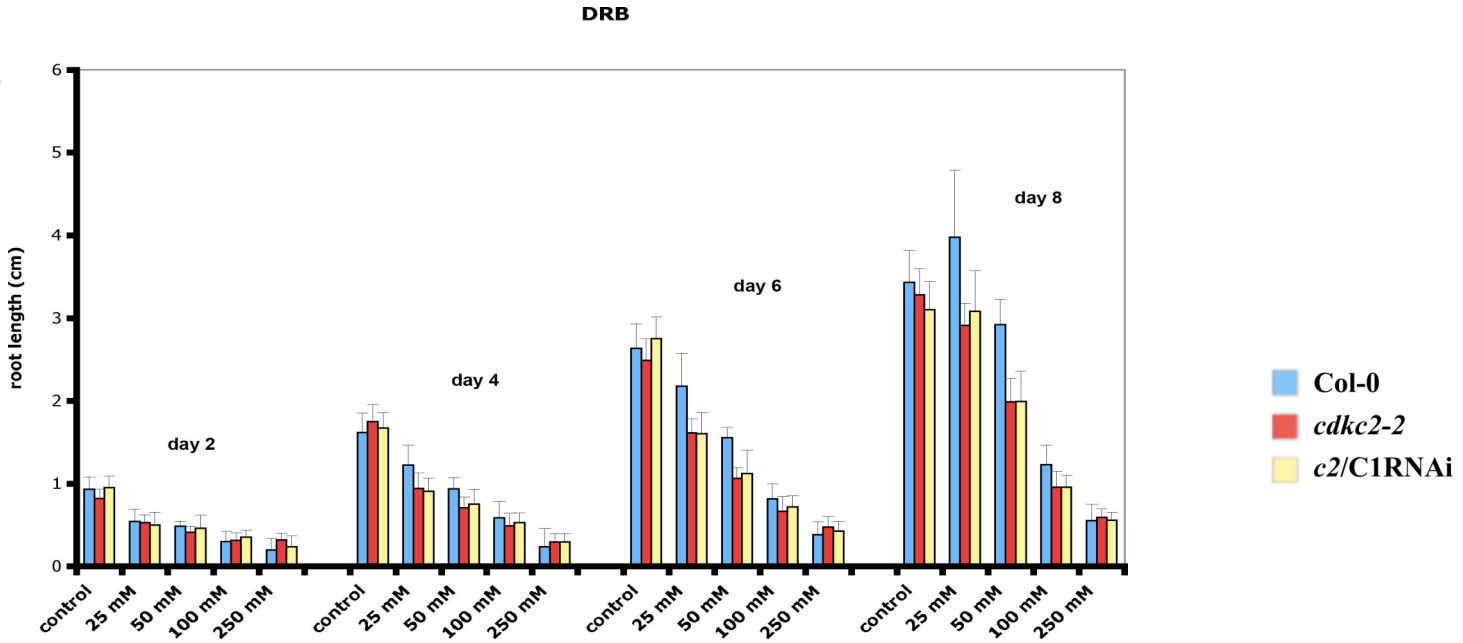
Figure 5.22 Effects of the transcriptional inhibitor, DRB, on root growth of Col-0, *cdkc2-2* and *c2*/C1RNAi seedlings.

Col-0, *cdkc2-2* and *c2*/C1RNAi plants germinated on Paul's media in the absence of DRB. Immediately after germination, 10 seedlings from each genotype were transferred onto Paul´s media containing either 200 µl DMSO (control) or the indicated concentrations of DRB (25, 50, 100 and 250 mM). Four measurements of root growth were recorded in 2-days intervals. All root measurements were normalised against the average root length of each genotype when seedling were transferred in the drug, so that at the day of transfer (day0) root length for each genotype had the value "0".

(A) Root length (y-axis) versus different concentrations of DRB at day2, day4, day6 and day8, after transferring germinated seedlings on medium containing the inhibitor.

(B) Table showing the significance values for experiments shown in (A) at which the root growth of *cdkc2-2* or *c2*/C1RNAi plants differs significantly from the root growth of Col-0. Asterisks indicate the different levels of significance: *, ** and *** correspond to p-values of 0.05, 0.01 and 0.001, respectively

(A)



(B)

		day2		day4		day6		day8	
		cdkc2-2	*c2*/C1RNAi	*cdkc2-2*	*c2*/C1RNAi	*cdkc2-2*	*c2*/C1RNAi	*cdkc2-2*	*c2*/C1RNAi
DRB	25mM			**	**	**	**	**	**
	50mM			**	*	***	***	***	***
	100mM							*	**
	250mM								

Figure 5.23 Effects of the CDK-specific inhibitor, roscovitine, on the root growth of Col-0, *cdkc2-2* and *c2*/C1RNAi seedlings.

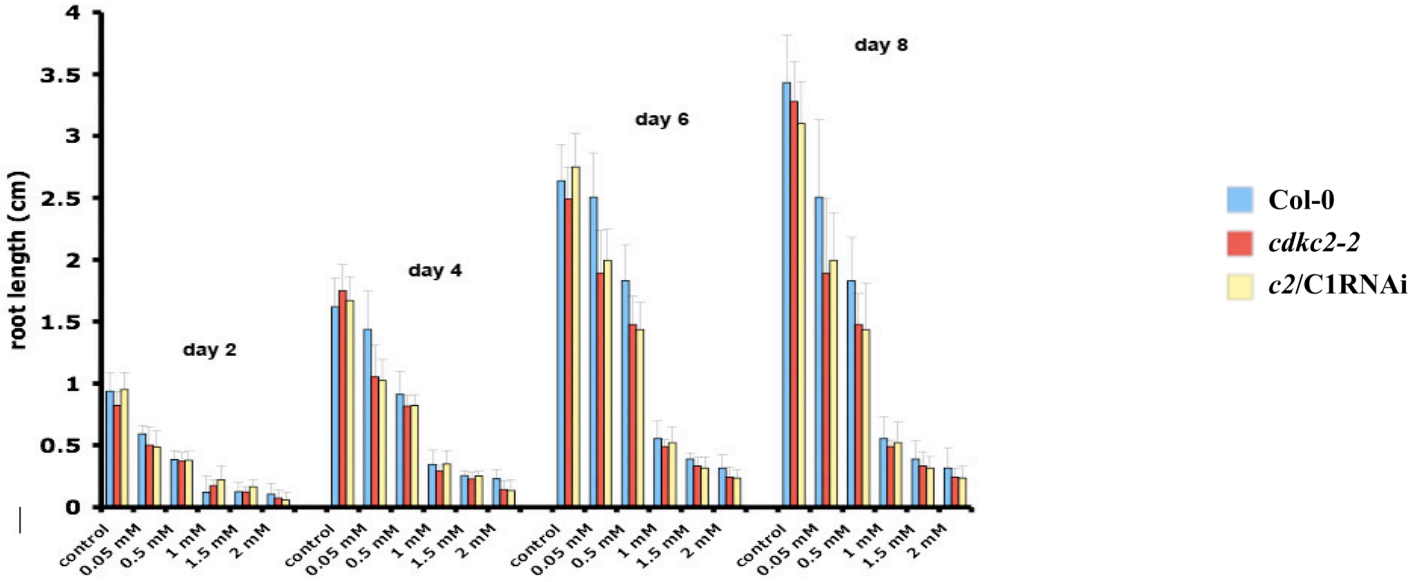
(Seedling handling, treatments and normalisation applied were the same as those described in Figure 5.22

Roscovitine concentrations: 0.05, 0.5, 1, 1.5 and 2mM)

(A) Root length (y-axis) versus different concentrations of roscovitine at day2, day4, day6 and day8, after transferring germinated seedlings on medium containing the drug inhibitor.

(B) Table showing the significance values for experiments shown in (A) at which the root growth of *cdkc2-2* or *c2*/C1RNAi plants differs significantly from the root growth of Col-0. Asterisks indicate the different levels of significance: *, ** and *** correspond to p-values of 0.05, 0.01 and 0.001, respectively.

(A)



(B)

		day2		day4		day6		day8	
		cdkc2-2	***c2*/C1RNAi**	***cdkc2-2***	***c2*/C1RNAi**	***cdkc2-2***	***c2*/C1RNAi**	***cdkc2-2***	***c2*/C1RNAi**
Roscovitine	0.05mM			*	**	**	**	**	***
	0.5mM					*	**	***	***
	1mM							**	**
	1.5mM							*	*
	2mM			*	*			***	***

5.10 Is the FP-CDKC fusion protein functional?

The subcellular localisation of the CDKC2 fusion protein in *Arabidopsis* cell cultures suggested a role of this kinase in regulating the activity and redistribution of spliceosomal components and suggested a possible involvement of the kinase with components of the EJC. However, the validity of this largely depends on whether or not the fusion protein is actually functional. To test the functionality, I transformed *cdkc2-2* plants, with the 35S:YFP:CDKC2 construct in the pEarleygate104 vector (see Figure 2.3 for map of the vector and hereafter called YFP:CDKC2). At this point I have to clarify a counterintuitive notion regarding the expression of *CDKC2* coding sequence under the 35S promoter. When the *GUS* gene was under the control of the 35S promoter, GUS activity was reduced by 4-fold in the *cdkc2-2* mutant line (Cui *et al*., 2007), suggesting the CDKC2 is important for the activity of the specific promoter. The use of 35S promoter to drive the expression of YFP:CDKC2 fusion in the *cdkc2-2* plants would be expected to lead to large suppression of the YFP:CDKC2 transcript. However, since the activity of the 35S promoter is not completely abolished (Cui *et al*., 2007), the small production of the YFP:CDKC2 protein fusion would lead to an increase to the activity of the promoter, which in turn would produce more fusion protein that would increase even more the activity of the promoter and so on. Thus, eventually the activity of the 35S promoter would be completely restored and enable the over-expression of the YFP:CDKC2 protein fusion.

YFP:CDKC2 protein fusion was generated using Gateway® cloning and transformed into plants as described in Section 2.8.1. T1 seeds were collected, sown and selected for BASTA resistance. Eight basta-resistant T1 plants were recovered. In the next generation, T2 plants were sprayed with basta and the segregation of herbicide resistance noted[1]. Several lines segregated at 3:1 [1]ratio (resistant : sensitive), indicating that the construct had integrated at (probably) only one site. From the resistant T2 plants I germinated the T3 seed on hygromycin. When 100% of T3 seeds were resistant to

[1] [1]Due to time limitations, prior to spraying the T2 plants, I transplanted in the greenhouse more seedlings from each T2 line, in order to have plants available for analysis from the hygromycin-resistant homozygous T2 lines. I call these extra plants as “backup” plants

hygromycin, the parent T2 plant was considered homozygous. Then, "backup" homozygous T2 plants (hereafter called YFP:CDKC2) were used for analysis.

I tested whether the drug resistant YFP:CDKC2 plants were late flowering or not, compared to Col-0 and the original mutant line. From the eight independent YFP:CDKC2 lines containing, six lines flowered at the same time as Col-0, as judged by the number of rosette leaves produced before bolting (Table 5.3). The other two lines had a number of leaves either close to that of the *cdkc2-2* mutant or an intermediate number between the mutant and wild type plants. This phenotype may be the result of reduced expression levels of the YFP:CDKC2 protein fusion. A comparison of leaf morphology among Col-0, *cdkc2-2* and YFP:CDKC2 plants showed that the complemented plants had regained a wild type leaf shape (Figure 5.24A). This was confirmed when more plants from a complemented line were checked (Figure 5.24C). Complementation was also observed in flowers. In *cdkc2-2*, petals are less wide than wild type petals and also show partial folding at the top of the petal. This phenotype is totally abolished in the YFP:CDKC2 lines and flowers resembled the wild type (Figure 5.24B).

Genotypes	Col-0	*cdkc2-2*	YFP:CDKC2
Rosette leaf no. (bolting)	5.6 ± 0.63	13 ± 0.4	6.3 ± 0.48
			6.6 ± 0.6
			5.6 ± 0.6
			6.1 ± 0.6
			6.2 ± 0.63
			6.5 ± 0.7
			9.4 ± 0.7
			11.2 ± 0.6

Table 5.3 Number of rosette leaves at bolting for Col-0, *cdkc2-2* and YFP:CDKC2 plants.

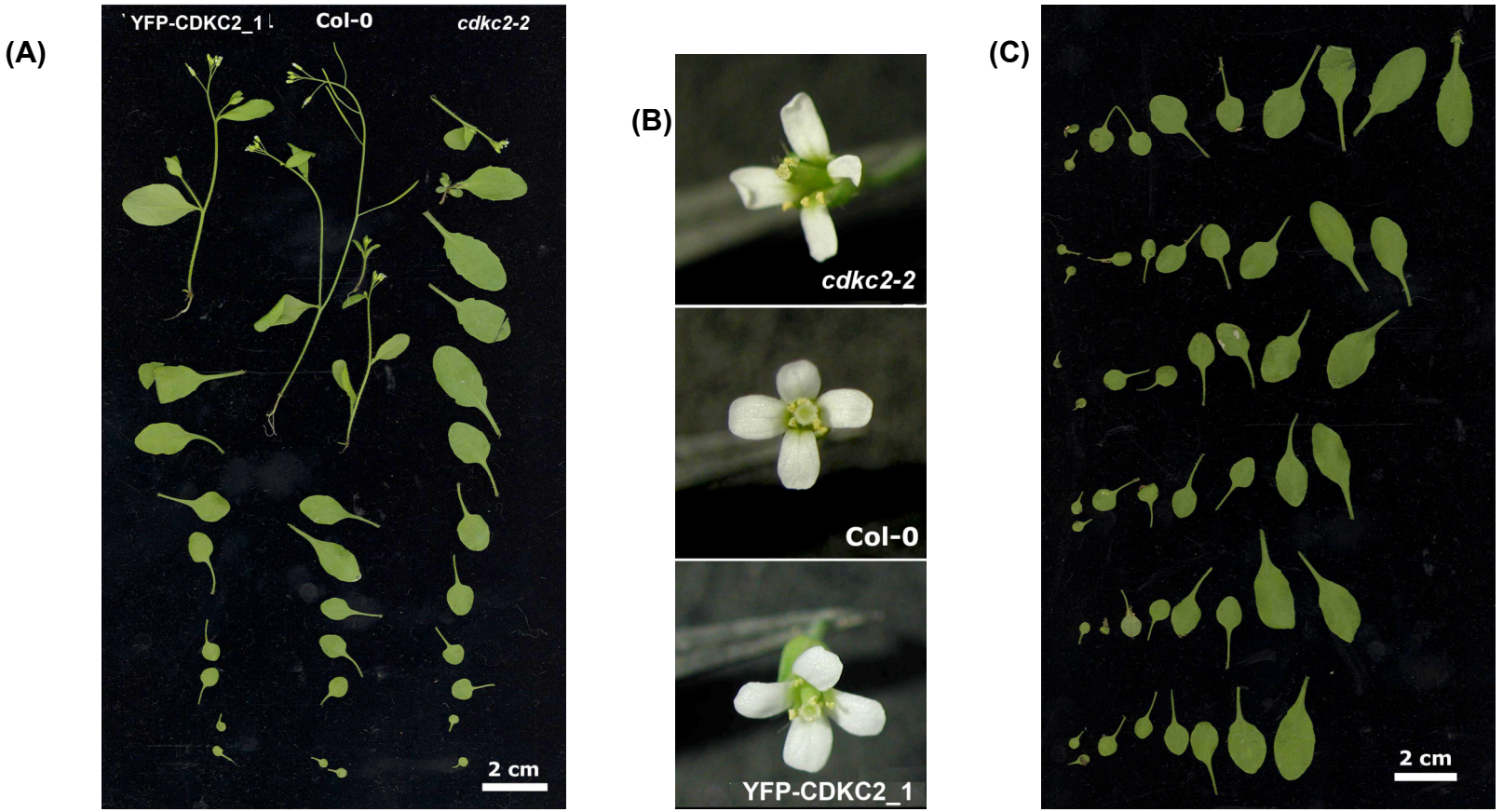
Rosette leaves were collected from the three genotypes and counted. One line was used in the case of Col-0 and *cdkc2-2* and 8 lines for YFP-CDKC2. Ten plants from each line had their leaves counted. The *cdkc2-2* line is homozygous and YFP-CDKC2 lines are T2 generation, all homozygous for the transgene.

Figure 5.24 Restoration of wild type phenotype in *cdkc2-2* plants by genetic complementation.

(A) "Leafograms" of a representative homozygous T2 homozygous YFP-CDKC2 plant (left) Col-0 (centre) and *cdkc2-2* (right). Cotyledons shown at the bottom, then leaves in developmental order and floral blot.

(B) Representative flowers from three different genotypes: *cdkc2-2*_(top), Col0 (middle) and YFP-CDKC2_1 (bottom).

(C) Rosette leaves from six different homozygous T3 plants from the complemented T2 homozygous line YFP-CDKC2_5.


(A)
YFP-CDKC2_1
Col-0
cdkc2-2
2 cm
(B)
cdkc2-2
Col-0
YFP-CDKC2_1
(C)
2 cm

To verify whether the wild type phenotype was due to the presence of the YFP:CDKC transgene, the T2 homozygous lines were genotyped using attB primers that are specific for the Gateway casette of the pEarleygate vectors (PCR was performed by Dr. Tao Zheng after I left the lab). From the 8 lines genotyped, all were positive for the transgene whereas no PCR product was detected in Col-0 plants. This result further confirms that the reversion of *cdkc2-2* plants back to wild type phenotype is due to the presence of the transgene.

5.11 Localisation of YFP:CDKC2 protein fusion in *Arabidopsis* roots

Stable transgenic plants expressing YFP:CDKC2 allowed me to follow protein localisation of the protein fusion *in planta.* In *Arabidopsis* root tips, the protein fusion localises in the nucleus as shown in Figure 5.24. Contrary to cell cultures transiently expressing the GFP:CDKC2 protein fusion, localisation of GFP:CDKC2 in root tips did not appear to concentrate in speckles but acquired a more diffuse nucleoplasmic distribution. This difference in the localisation could be due to different transcriptional status between the cell cultures (observed at day3 after subculture) and the meristematic cells. When I observed YFP:CDKC2 localization in the differentiated zone of the root I could detect localised concentration of the protein in speckles, as well as a more diffuse nucleoplasmic localisation signal. Similarly to the differences seen with the cell cultures, different distribution of CDKC2 in the differentiated root tissue may reflect lower rates of transcription in the latter (Fang *et al*., 2004; Docquier *et al*., 2004), leading to increased accumulation of the kinase in speckles.

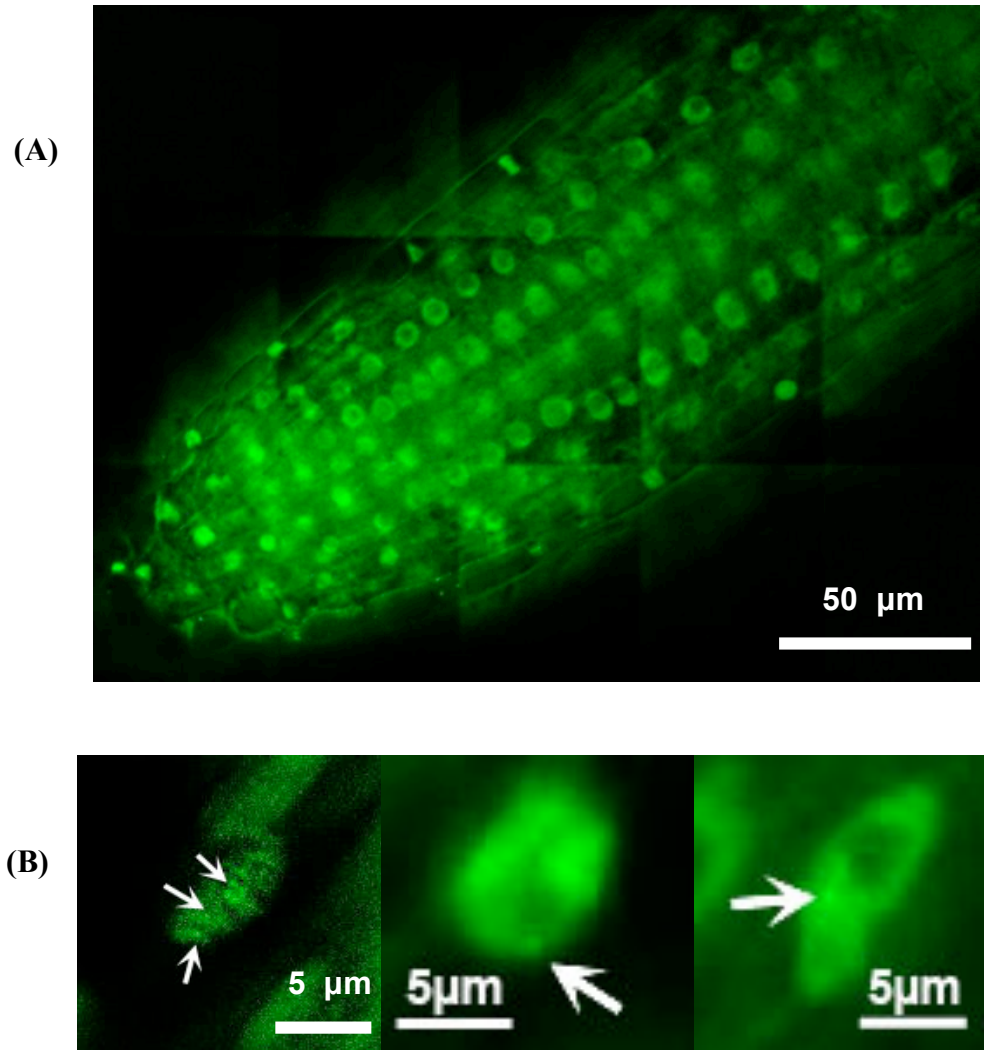


Figure 5.25 YFP:CDKC2 localisation in *Arabidopsis* plants.

Complemented plants stably expressing YFP:CDKC2 were visualised using a fluorescent microscope and the localisation of the protein fusion in root tip (A) and in the differentiated zone of the root (B) was monitored. Images represented a single optical Z-section of the root. Arrows in (B) indicate the concentrated localisation of YFP:CDKC2 fusion in speckles. Scale bars: 50 µm (A) and 5 µm (B).

5.12 Discussion

5.12.1 Colocalisation of CDKC2 with splicing factors

Co-localization of CDKC2 with splicing factors and their re-distribution upon treatment with a CDK-specific inhibitor suggests that reversible phosphorylation might regulate spliceosomal protein distribution (Kitsios *et al*., 2008). Use of the dominant negative mutant of CDKC2, CDKC2_DN, suggests that CDKC2 (and not other CDK-related kinases) is responsible for the observed distribution. In vitro kinase assays will be important to discriminate between direct and indirect targets. Re-distribution of mammalian spliceosomal proteins occurs in the presence of kinase inhibitors and, coupled to the presence of the RS motif in their sequence, pointed to the importance of serine phosphorylation on the nuclear distribution of these proteins (Misteli and Spector, 1997; Misteli, 2000). Indeed in mammalian cells, movement of the splicing factors in and out of the speckles is controlled by reversible phosphorylation of serine residue in the RS motif by SR-protein kinase 1 (SRPK1) and CDC2-like protein kinase 1 (CLK1) (Gui *et al*., 1994; Colwill *et al*., 1996).

At least one direct target is already known. CDKC2, along with its cyclin partner CYCT1;3 (Barrroco *et al*., 2003), can phosphorylate the CTD of RNA PolII *in vitro* (Cui *et al*., 2007). CDKC2 co-localises with the CTD of RNA PollII, supporting this idea. However, it is not known whether SRp34 and Cyp64 are CDKC substrates since the results reported here could also be explained by indirect effects of CDKC, acting, for instance on the CTD of PolII. Future experiments should aim to test if SRp34 and Cyp64 are indeed substrates of CDKC2 and whether CDKC2 has a direct effect on pre-mRNA splicing. The clones and genetic strains described in this chapter should allow these questions to be addressed. For example, deep sequence analysis of nuclear transcripts in mutant and wild type plants will allow us to assess whether loss of CDKC leads either to mis-splicing or to the accumulation of unspliced transcripts. In vitro kinase and protein-protein interaction assays could be used to access various spliceosomal proteins as substrates.

5.12.2 CDKC2 as a regulator of nucleo-nucleolar shuttling in response to stress

Shuttling between the nucleoplasm and nucleolus has been observed for many plant proteins (Pendle *et al*., 2005; Tillemans *et al*., 2005, 2006; Kitsios *et al*., 2008). The relative balance can depend on the physiological status of the cell for example, and the predominant localisation pattern can change due to hypoxic stress during live imaging conditions (Koroleva *et al*., 2009; Tillemans *et al*., 2006) or after drug-induced inhibition of specific cellular processes (Docquier *et al*., 2004; Tillemans *et al*., 2005, 2006; Kitsios *et al*., 2008). Sequestration of a protein either in the nucleoplasm or the nucleolus could be a mechanism for partitioning a protein away from its substrate under non-favourable conditions (Sansam *et al*., 2003; Shaw and Doonan, 2005). CDKC2 is localized mainly in the nucleoplasm but there is a low level in the nucleolus even under normal conditions. Its accumulation in the nucleolus, either as a kinase-dead version or in the presence of roscovitine, suggests that kinase activity is required to maintain these normally low levels. The same behaviour has been observed for a number of splicing factors such as SRp34, SRp31, ASF/SF2 and RSZp22 (Kitsios *et al*., 2008; Tillemans *et al*., 2005; Shav-Tal *et al*., 2005). Nucleolar accumulation of CDKC2_DN could be seen only after hypoxic stress, suggesting that the overexpression of the mutant kinase *per se* in *Arabidopsis* cells is insufficient to produce a roscovitine-like phenotype. ATP depletion due to hypoxic stress (Hochachka *et al*., 1996) is also required and this has also been connected to translocation of splicing factors to the nucleolus and perinucleolar compartments (Tillemans *et al*., 2005, 2006).

It would be useful to verify whether ATP depletion actually occurs under these conditions. The role of ATP depletion could be assessed using sodium azide, a respiratory inhibitor that has been shown to induce a nucleolar translocation of the EJC component eIF4A-III (Koroleva *et al*., 2009).

5.12.3 Co-localisation of CDKC2 with components of EJC

This is the first report where a plant CDK protein has been found to colocalise with an EJC protein component. CDK11, the mammalian homolog of *Arabidopsis* CDKG1, is the only CDK found to associate with a component of the EJC and splicing factor

RNPS1 (Hir *et al*., 2000; Mayeda *et al*., 1999; Loyer *et al*., 1998). GFP-CDKC2 only partially co-localised with the *Arabidopsis* RNPS1 but showed extensive co-localisation with the Magoh protein fusion in cell cultures. Phosphorylation of the CTD tail of RNA PolII by CDKC2 and its co-localisation with splicing factors and Magoh protein, points towards identifying CDKC2 as a possible convergent point in the regulation of transcription, splicing and EJC activity. Reciprocal pull downs in cell cultures co-expressing GFP:CDKC2 and Magoh:RFP fusions protein could be used to test whether these two proteins interact directly.

Regardless of whether the interaction is direct or not, dramatic redistribution of Magoh:RFP protein fusion occurs when co-expressed with CDKC2_DN (Figure 5.11). In *Drosophila melanogaster,* there is a cytoplasmic fraction of Magoh in addition to the predominant nuclear localisation, which is involved in the export of spliced mRNPs through binding to Y14, a component of EJC, and export factor TAP (Kataoka *et al*., 2001). To this end, it would be interesting to examine in the future whether CDKC2 activity affects the nucleocytoplasmic shuttling of Magoh protein and/or mRNA export. The latter could be monitored cytologically using fluorescein-labelled oligodT oligomer, which labels cellular mRNA through binding to its polyA tail (Gong *et al*., 2005).

5.12.4 YFP:CDKC2 is functional and complements *cdkc2* T-DNA insertion mutants

Functional complementation of *cdkc2-2* mutant phenotype to wild type by transformation with YFP:CDKC2 demonstrates that the protein fusion is functional. A previous low-resolution study indicated that CDKC2 fusion proteins locate to the nucleus (Cui *et al*., 2007). My results provide a high-resolution localisation profile of CDKC2 in plants, and show that protein fusion resides primarily in the nucleoplasm and accumulates into speckles in some plant cells. This dual localisation profile is similar to localisation of proteins involved in the splicing process, like splicing factors RSZp22, RSp31, RSZ33 and SRp34 (Tillemans *et al*., 2005; Tillemans *et al*., 2006). Other spliceosomal components that do not participate in splicing also concentrate into nuclear speckles, which are believed to be sites of storage/assembly of splicing

machinery (Fang *et al*., 2004). The diffuse nucleoplasmic fraction of CDKC2 possibly participates in active splicing, whereas the "speckled" fraction is maintained as a reserve ready to be recruited for the splicing of newly synthesised pre-mRNAs. Similarly to YFP:CDKC (Figure 5.25), splicing factors exhibit characteristic localisation profiles in meristematic as opposed to differentiated cells, possibly mirroring transcriptional and/or splicing activity of the respective cell types (Fang *et al*., 2004; Docquier *et al*., 2004).

5.12.5 Phenotypic analysis of *cdkc2-2* T-DNA insertion mutants

While undertaking the complementation experiment, it became apparent that the *cdkc2-2* T-DNA mutant had additional phenotypic defects not reported previously, including changes in the lamina-petiole angle, and perturbation of petal and trichome development. In addition, I also observed hypersensitivity of both *cdkc2-2* and *c2*/C1RNAi plants to DRB and roscovitine (Figures 5.22 and 5.23), supporting the idea that CDKC family members are involved in regulating gene transcription in *Arabidopsis* (Barocco *et al*., 2003). Alteration of specific plant developmental processes in the *cdkc2-2* and *c2*/C1RNAi mutants, suggest that CDKC1 and CDKC2 act by targeting specific pathways, either by phosphorylating specific proteins or by regulating the activity of RNA PolII in a pathway-specific manner. For example, tissue-specific enhanced expression of TFIID components is important for germline development and spermatogenesis in Drosophila and mammals (Hochheimer and Tjian, 2003; Hiller *et al*., 2004), illustrating an efficient strategy employed by organisms to regulate gene expression.

5.12.6 Future directions

To understand further the molecular basis of the phenotypes reported here and in (Cui *et al*., 2007), it will be important to determine how loss of CDKC affects gene expression. The most direct route will be to compare transcripts from mutant and wild type plants, using approaches such as microarray analysis. This should reveal which pathways are affected. Analysis of mRNA transcripts might reveal if splicing, rather than

transcription, is the primary process affected. Alternative approaches employing genetic analysis are also possible. Moreover suppressor and enhancer screens, will allow us to place CDKC within genetic pathways for given traits. For example, studying the effect that the depletion of CDKC proteins has on trichome morphogenesis could be aided by the use of already defined mutants (Zhang *et al*., 2005; Luo and Oppenheimer, 1999; Ilgenfritz *et al*., 2003).

Chapter 6 – In vitro identification of CDKC substrates

Subcellular localisation of CDKC2 protein and its co-localization with other proteins suggested novel roles in splicing and mRNA export. To directly test that any of these proteins could be a kinase substrate, I established an *in vitro* kinase assay for CDKC and used it to test whether one of the candidates, Magoh, is a substrate of the kinase.

6.1 Production of recombinant CDKC2 and CycT1;3 protein fusions in BL21 cells

For the *in vitro* analysis, I generated recombinant kinase and cyclin proteins. All of the recombinant proteins were tagged with GST, apart from CycT1;3, which was tagged with His. The following constructs were generated: GST-CDKC2 and GST-CDKC2_D182N encoding for the kinase-dead CDKC2 variant. cDNAs corresponding to the wild type and mutant kinase, were introduced into the pGEX4T-3 vector (GE Healthcare) that produces N-terminus GST protein fusions. CycT1;3 was provided by Dr. Zhixiang Chen (Purdue University) as an N-terminus His fusion in pET32a vector (Addgene) and a construct harbouring two repeats of the CTD peptide, a known substrate of CDKC2 (Cui *et al*., 2007), in the pGEX4T-3 vector. Even though authors claimed the presence of only two repeats of the CTD peptide, sequencing of the construct showed that there were 16 repeats of the YSPTSPS consensus sequence and 10 repeats of the partial CTD sequence YSPTSP (see Appendix-IV for sequencing result obtained using the pGEX-forward and pGEX-reverse primer pair and respective chromatograms). Respective plasmids were then transformed into BL21 cells as described in Section 2.4.1.2.Protein over-expression and solubility of the recombinant protein depend on different variables, including concentration of the inducing agent (IPTG) and the temperature and duration of induction, and require optimisation.

For CDKC2 and CDKC2_D182N, I used two independent colonies for each construct. Two different concentrations of IPTG were used (0.5mM and 1mM) for protein induction and cultures were incubated at 37^{o}C for 3 hours. Figure 6.1A-C shows protein extracts from cells expressing the three variants of CDKC2 protein fusions. Protein expression levels were very good for all three variants, whereas no differences could be observed for different colonies or IPTG concentration for all the kinase

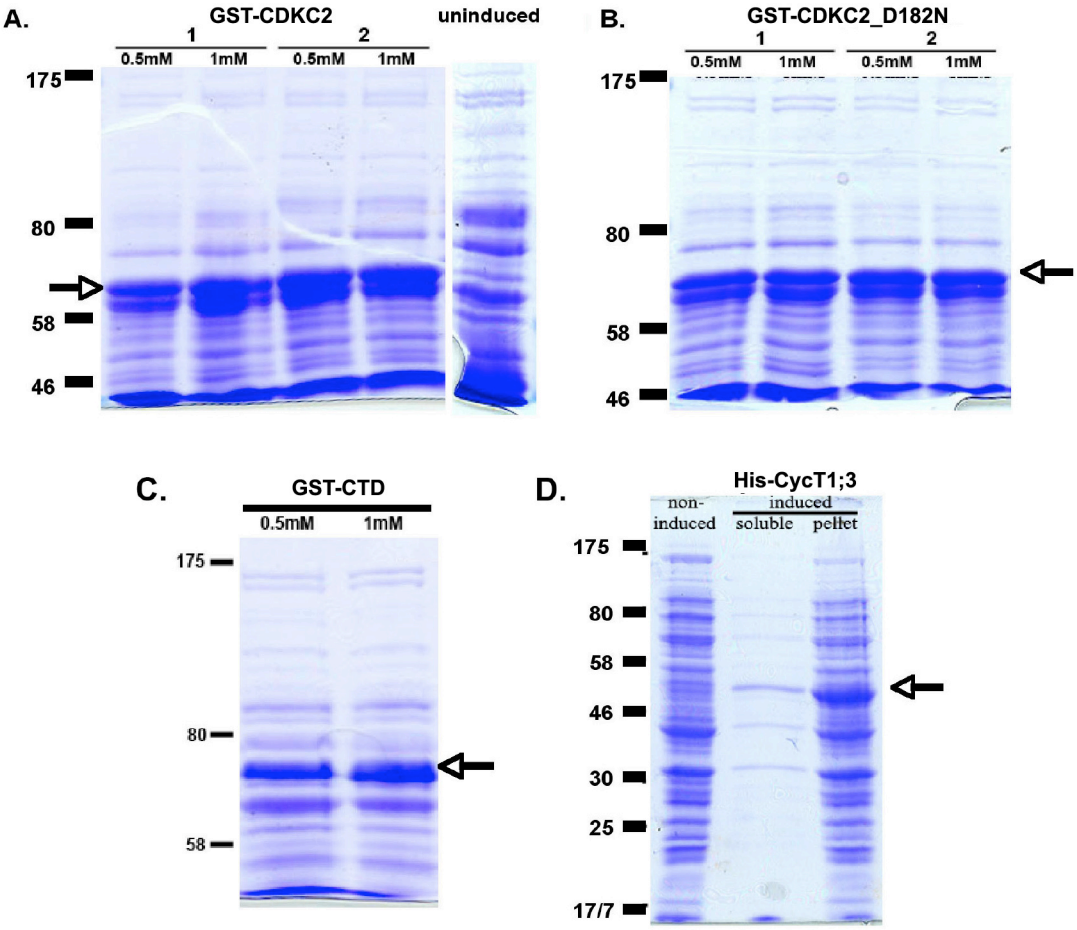
Figure 6.1. Recombinant protein expression in BL21 *E. coli* cell cultures.

(A-B) Protein extracts containing recombinant GST-CDKC2 (A) and GST-CDKC2_D182N (B). On the right of (A) there is a protein extract from uninduced control for comparison that applies to both images.

(C) Protein extracts containing recombinant GST-CTD peptide, using 0.5mM or 1mM IPTG,

(D) Soluble and insoluble protein extracts containing His-CycT1;3 recombinant protein. On the left, there is a gel of protein extracts from un-induced control for comparison.

Protein expression was induced by adding different concentrations of IPTG (as shown above each lane) and incubating at 37^{o}C for different durations of time (see Section 2.6.1 for induction conditions). Proteins were extracted under either denaturing conditions (A-C) or under native conditions (D), as described in Section 2.6.2. The positions of Molecular Weight markers are shown to the left of each panel. Arrows in all images indicate the position of the recombinant protein/peptide. Numbers above the IPTG concentrations indicate different colonies expressing the respective GST fusion (Calculated molecular weights: 82kDa for GST-CDKC2 wild type and mutants; 57kDa for His-CycT1;3; 80kDa for GST-CTD).


A.
GST-CDKC2
uninduced
1
2
0.5mM
1mM
0.5mM
1mM
175
80
58
46
B.
GST-CDKC2_D182N
1
2
0.5mM
1mM
0.5mM
1mM
175
80
58
46
C.
GST-CTD
0.5mM
1mM
175
80
58
D.
His-CycT1;3
non-induced
induced
soluble
pellet
175
80
58
46
30
25
17/7

However, the observed size of the expressed CDKC2 protein was smaller than the predicted molecular weight, perhaps due to partial degradation of the recombinant protein. Nevertheless, when affinity purified using the GST tag, there was a good yield of GST-CDKC2 recombinant proteins (wild type and mutant). The reduction in size could be explained by removal of the GST tag during the purification step, probably due to the presence of residual thrombin in one of the buffers, resulting in a proteolytic digestion. In terms of GST-CTD expression, the same yield of recombinant protein was obtained with both 0.5mM and 1mM IPTG (Figure 6.1D). Again the size of the observed protein product was smaller than the calculated peptide size.

Most CDKs require a cognate cyclin for their activation as a kinase; for CDKC, one of the cyclins known to activate is CycT. Induction conditions for the recombinant His-CycT1;3, were as described in (Cui *et al*., 2007). Using native protein extraction conditions, a small amount of protein was soluble whereas the majority stayed in the pellet as an insoluble fraction (Figure 6.1E). The soluble fraction was purified and used to activate the kinase.

Protein extraction, affinity purification and eluate concentration procedures are described in Sections 2.6.2 and 2.6.3. The gel presented in Figure 6.2 shows eluates of affinity-purified GST- and His-tagged proteins. The majority of proteins were of relatively high yield and purity. Only the concentration of His-CycT1;3 fusion protein appeared to be low (Figure 6.2A), because large amounts of the recombinant protein stayed on the beads (Figure 6.2B, see the right hand lane “boiled beads”).

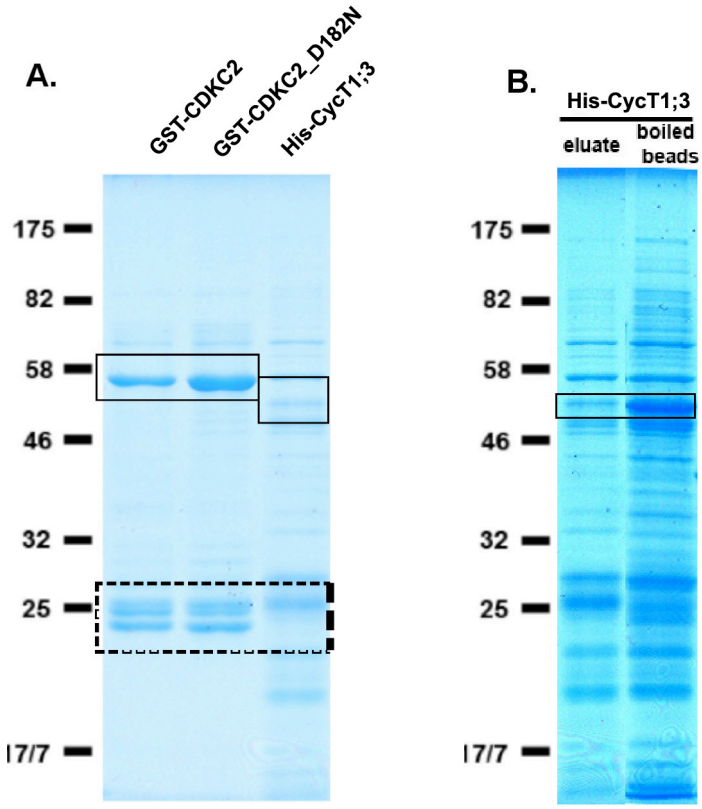


Figure 6.2. Native purification of recombinant proteins using GST-Sepharose or His-Cobalt beads.

Affinity purified recombinant GST- and His-tagged proteins: rectangles surround the recombinant proteins and the dashed rectangle shows probable degradation products. The position of the Molecular Weight markers is shown to the left of Panel A and applies to both gels.

(A) Eluates from beads for each of the recombinant proteins, as indicated above each lane. Approximately equal amounts of protein (4µg) were loaded on a 12% SDS-PAGE gel.

(B) Comparison between the amount of His-CycT1;3 eluted from beads (left lane) and the amount that was retained on the beads (right lane).

6.2 Assessment of the activity of CycT1;3/CDKC2 recombinant complexes

After obtaining soluble recombinant proteins, I examined whether they could form an active complex that could phosphorylate the CTD peptide, a known substrate of CDKC. For doing this I performed an *in vitro* kinase assay. During this assay, the CDKC2 or CDKC2_DN GST fusions were incubated with the CTD peptide or the GST-Magoh fusion in the presence of radioactive ATP (^{32}P-ATP). Incorporatation of ^{32}P-ATP on the substrate is examined by running the reaction on a denaturing SDS gel and exposing the dried gel to an X-ray film. Appearance of a band at the same size as the substrate would indicate that this protein is an *in vitro* substrate of the kinase. Table 6.1 shows the amounts of different reactants used for the kinase assay Since CDKs are activated through binding of the respective cyclin, I included CycT1;3-GST fusion (a known partner of CDKC family members (Cui *et al*., 2007)) in the kinase reaction. As a positive control in my assay I used a sequence containing multiple repeats of the CTD peptide sequence (YSPTSPS; Dietrich *et al*., 1990; see Appendix -IV) fused to GST. In order to show that only the co-existence of CDKC2 and CyclinT1;3 in the kinase reaction is sufficient to form an active heterodimeric complex, I performed one reaction with only CDKC2 and CTD present. CDKC2 alone was not sufficient to phosphorylate the CTD peptide (Figure 6.3; lane 1).

Reaction components	Amounts used			
GST-CDKC2	1µl	1µl	1µl	-
GST-CDKC2_D182N	-	-	-	1µl
HIS-CYCLINT1;3	-	2µl	2µl	2µl
GST-CTD_peptide	1µl	-	1µl	1µl
Reaction cocktail	23µl	22µl	21µl	21µl
End Volume	25µl	25µl	25µl	25µl

Table 6.1 Composition of samples used in kinase assays.

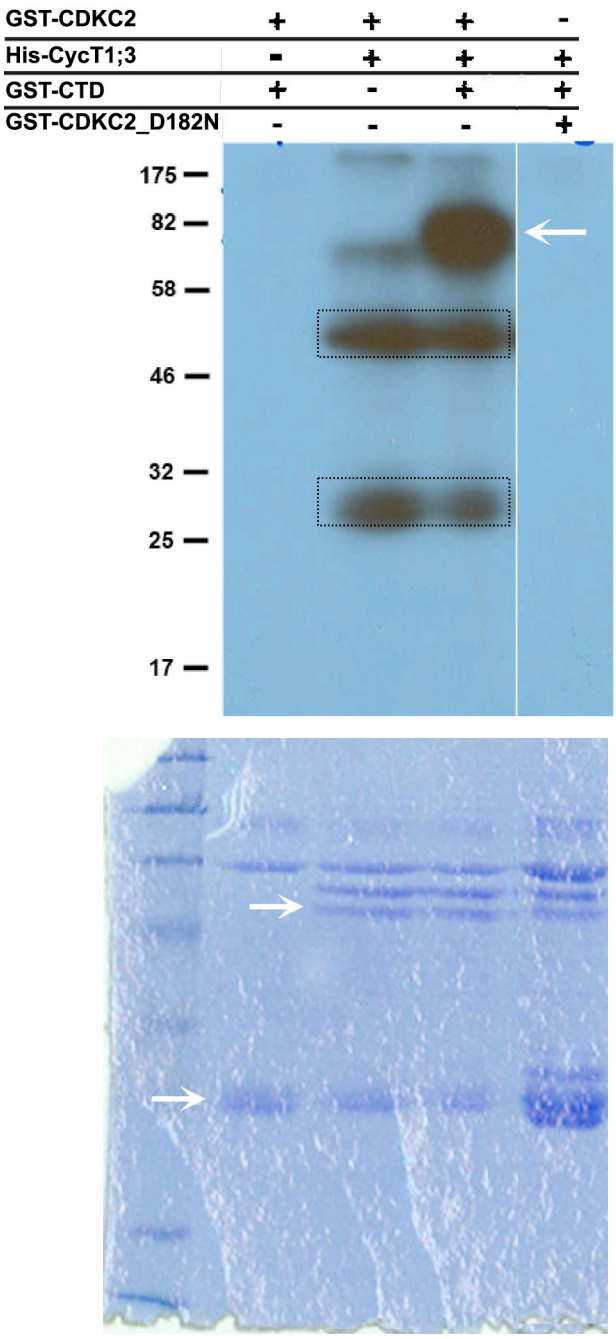
The first column of the table shows recombinant proteins/reagents used for the kinase assay. The other columns indicate which and how much of each recombinant protein or peptide were used in each reaction.

Also the absence of any band when the CDK/Cyclin are incubated alone demonstrates the specificity of the CTD phosphorylation event (Figure 6.3; lane 2). If the CycT1;3/CDKC2 complex is active due to the kinase activity of CDKC2, then the D182N mutation should abolish the CTD phosphorylation. The CDKC2_D182N / CycT1;3 complex was unable to phosphorylate the CTD peptide (Figure 6.3; compare lanes 2 and 3 with 4), confirming that the D182N mutation eliminates activity of the CDKC protein, and so is in agreement with existing reports in the literature where the substitution of conserved Asp by Asn in eukaryotic CDKs resulted in inactivation of related kinases (Hemerly *et al*., 1995; Gould and Nurse, 1989; Garriga *et al*.,1996).

Figure 6.3 Assessment of the kinase activity of CycT1;3/CDKC2 complexes towards the CTD peptide.

Top panel: Autoradiograph showing the kinase activity of hetromeric CDK-cyclin complexes.
Bottom Panel: the dried gel corresponding to the autoradiograph shown in panel above

The identity of recombinant proteins used in each assay is given in the table immediately above the top panel. Rectangles on the gel indicate the major bands phosphorylated in the absence of the CTD substrate. White arrows on the film indicate the band corresponding to phosphorylated CTD peptide, whereas white arrows on the gel indicate bands that are phosphorylated by the CycT1;3/CDKC2 complex, either in the presence or absence of the CTD substrate. The position of molecular weight markers is indicated to the left of the autoradiograph (the presence of the white line between the third and the fourth lane, results from the removal of a kinase CDKC2 mutant that I had tested for altered kinase activity, but which finally was not included in the body of the thesis).


GST-CDKC2 + + + -
His-CycT1;3 - + + +
GST-CTD + - + +
GST-CDKC2_D182N - - - +
175
82
58
46
32
25
17

6.3 Use of active CycT1;3/CDKC2 complexes to assess if Magoh is a substrate

Localisation studies in cell culture indicated that CDKC2 co-localises extensively with Mago-nashi homolog (Magoh), a member of the EJC complex. A physical interaction between Magoh and CDKC, but not other CDKs (A, B and D were tested), was found in a pair-wise yeast 2H interaction test (Pendle and Shaw, unpublished data). Moreover, there is a good consensus (but putative) CDK phosphorylation site in the Magoh protein sequence (Figure 6.4A). Taken together, these data suggested that Magoh maybe a substrate of CDKC2.

My strategy for answering this question was to express recombinant Magoh protein in bacteria, purify and use it as a substrate in an *in vitro* kinase assay. If Magoh is phosphorylated by the CycT1;3/CDKC2, the respective phosphorylated residue/site could be determined.

6.3.1 Production of recombinant Magoh protein and its mutants

To experimentally test whether T57, within the putative CDK phosphorylation site T57PVLK62 is functionally important, and as a control for the kinase assays, I generated two amino acid substitution mutants of Magoh protein by substituting threonine with alanine and glutamine. The first substitution, T57A, produces a protein impossible to phosphorylate at position 57 (called “phosphonull mutant”). The second substitution, T57E, resembles a phosphorylated residue and is called “phosphomimic mutant”. The methodology for T57 substitution is given in Section 2.6. Sequence chromatograms from positive clones confirmed the substitution of T57 to A57 (phosphonull) and E57 (phosphomimic) (Figure 6.4C). Moreover, full sequence profiles for both mutants shows that position 57 carried the only amino acid substitution in both Magoh mutants and this is illustrated by aligning all the three protein sequences: Magoh, Magoh_T57A, Magoh_T57E (Figure 6.4B).

Figure 6.4. Confirmation of in vitro mutagenesis

(A) Sequence of Arabidopsis Magoh protein where the putative CDK phosphorylation site is highlighted.

(B) Multiple sequence alignment of Magoh, Magoh_T57A and Magoh_T57E proteins, generated using the clustalw algorithm (http://align.genome.jp/sit-bin/clustalw). Amino acid substitution at position 57 is framed.

(C) Partial sequencing chromatograms from plasmids containing cDNA sequence encoding for Magoh (1), Magoh_T57A (2) and Magoh_T57E (3). Nucleotide triplet encoding for T57 is highlighted in (1) whereas respective nucleotide substitutions are highlighted in (2) and (3).

(A)

Magoh (At1g02140; NP_171716)

MAAEEATEFYLRYYVGHKGKFGHEFLEFEFREDGKLRYANNSNYKNDTIIRKE
VFLTPAVLKECKRIVSESEILKEDDNNWPEPDRVGKQELEIVLGNEHISFATSKIG
SLVDCQSSNDPEGLRIFYYLVQDLKCLVFSLISLHFKIKPI

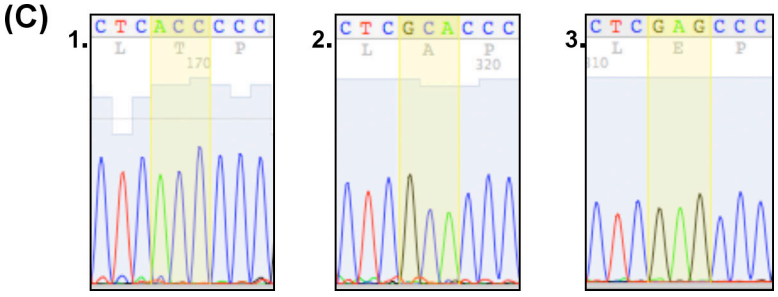
(B)

```
Magoh           MAAEEATEFYLRYYVGHKGKFGHEFLEFEFREDGKLRYANNSNYKNDTIIRKEVFLTPAV
Magoh_T57A      MAAEEATEFYLRYYVGHKGKFGHEFLEFEFREDGKLRYANNSNYKNDTIIRKEVFLAPAV
Magoh_T57E      MAAEEATEFYLRYYVGHKGKFGHEFLEFEFREDGKLRYANNSNYKNDTIIRKEVFLEPAV
                ************************************************************

Magoh           LKECKRIVSESEILKEDDNNWPEPDRVGKQELEIVLGNEHISFATSKIGSLVDCQSSNDP
Magoh_T57A      LKECKRIVSESEILKEDDNNWPEPDRVGKQELEIVLGNEHISFATSKIGSLVDCQSSNDP
Magoh_T57E      LKECKRIVSESEILKEDDNNWPEPDRVGKQELEIVLGNEHISFATSKIGSLVDCQSSNDP
                ************************************************************

Magoh           EGLRIFYYLVQDLKCLVFSLISLHFKIKPI
Magoh_T57A      EGLRIFYYLVQDLKCLVFSLISLHFKIKPI
Magoh_T57E      EGLRIFYYLVQDLKCLVFSLISLHFKIKPI
                ******************************
```

(C)



Recombinant Magoh proteins were produced in BL21 cells. Optimisation of expression utilised 4 different BL21 clones containing the wild type Magoh. Three different IPTG concentrations and three different incubation times were tested and Figure 6.5A shows recombinant protein production in a representative colony. Yield was high, no differences between clones or indication conditions were noted and, in contrast to the other recombinant proteins, GST-Magoh seemed to be expressed at the correct molecular weight. The mutant Magoh GST-tagged proteins were produced and purified under similar conditions (Figure 6.5B).

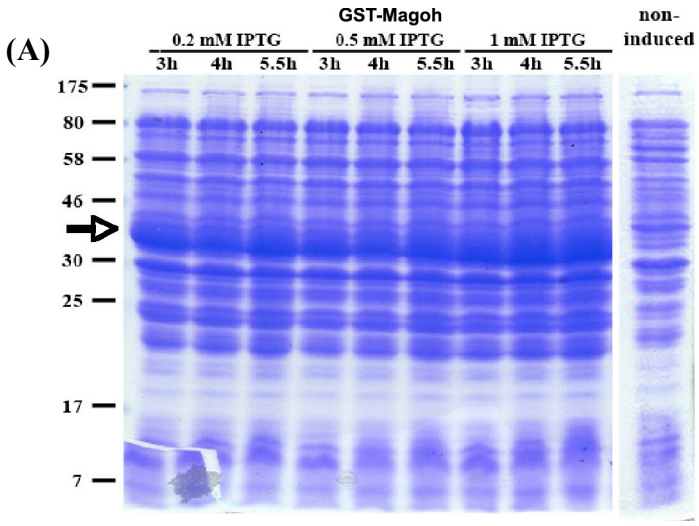
Figure 6.5 Expression and purification of recombinant GST-Magoh fusion proteins.

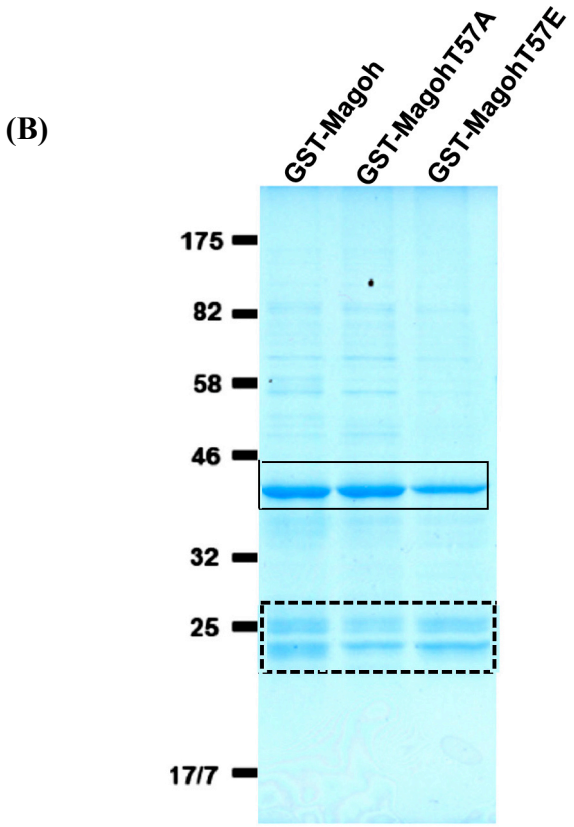
Proteins were extracted under denaturing conditions (A) or native conditions that were used for affinity purification of the tagged recombinant protein (B).

(A) Protein extracts from a representative cell culture expressing wild type GST-Magoh protein, after induction at different concentrations of IPTG and durations (as indicated above), the uninduced control is shown at the far right of the gel The arrow indicates the position of the expressed GST-Magoh protein.

(B) GST-affinity purified proteins from bacterial cell cultures expressing Magoh, Magoh_T57A and Magoh_T57E, separated on a 10% SDS-PAGE. Solid rectangle indicates the recombinant GST-Magoh and dashed rectangle indicates probably cleaved GST.

The position of Molecular Weight markers is shown at the left of each panel.


(A)
GST-Magoh
0.2 mM IPTG
0.5 mM IPTG
1 mM IPTG
non-induced
3h
4h
5.5h
175
80
58
46
30
25
17
7


(B)
GST-Magoh
GST-MagohT57A
GST-MagohT57E
175
82
58
46
32
25
17/7

6.3.2 GST-Magoh is not a substrate of CycT1;3/CDKC2 complex in vitro, but CyclinT1;3 is phosphorylated

An *in vitro* kinase assay to assess whether Magoh is phosphorylated by the CDKC2 is shown in Figure 6.6. The reactants and amounts used in the kinase assay are given at Table 6.2. As a positive control for the kinase assay I used the CTD peptide as a substrate and as a negative control I used the CyclinT1;3/CDKC2 complex alone. Also, in order to show that only the co-existence of CDKC2 and CyclinT1;3 in the kinase reaction is sufficient to form an active heterodimeric complex, I performed one reaction with only CDKC2 and CTD present. CDKC2 alone was not sufficient to phosphorylate the CTD peptide (Figure 6.6A; lane 1). Also from the kinase reaction it could be concluded that Magoh protein, or any of its mutant variants, is not phosphorylated by the active CDKC2/CycT1;3 complex (Figure 6.6A; dashed rectangle), which was able to phosphorylate the CTD peptide (Figure 6.6A; lane 3). This result suggests that Magoh is not a direct substrate of CDKC2, at least *in vitro*.

Reaction components	Amounts used					
GST-CDKC2	1µl	1µl	1µl	1µl	1µl	1µl
HIS-CYCLINT1;3	-	2µl	2µl	2µl	2µl	2µl
GST-Magoh	-	-	-	3µl	-	-
GST-Magoh_T57A	-	-	-	-	3µl	-
GST-Magoh_T57E	-	-	-	-	-	3µl
GST-CTD_peptide	1µl	-	1µl	-	-	-
Reaction cocktail	23µl	22µl	21µl	19µl	19µl	19µl
End Volume	25µl	25µl	25µl	25µl	25µl	25µl

Table 6.2 Reaction components of the *in vitro* kinase assay

However, the increased resolution afforded by the 15% gel revealed the origin of one of the multiple bands appearing on the film, also in the absence of any added substrate (see

page 149). Figure 6.6B shows an overlay of the dried gel and the photographic film from Figure 6.6A. The arrow in this figure shows that the band appearing at ca. 57 kDa matches perfectly with the His-CycT1;3 band on the dried gel. Thus, upon formation of an active CDKC2 / CycT1;3 complex, CycT1;3 is autophosphorylated. Sequence analysis of CycT1;3 protein (NP_174084) showed that there is a single putative CDK phosphorylation site at positions 45-48 with the sequence SPSR. Thus, the CDK/Cyclin complex probably phosphorylates the Ser45 residue, but time limitations did not allow me to verify the site of phosphorylation or its effects on kinase activity or plant growth.

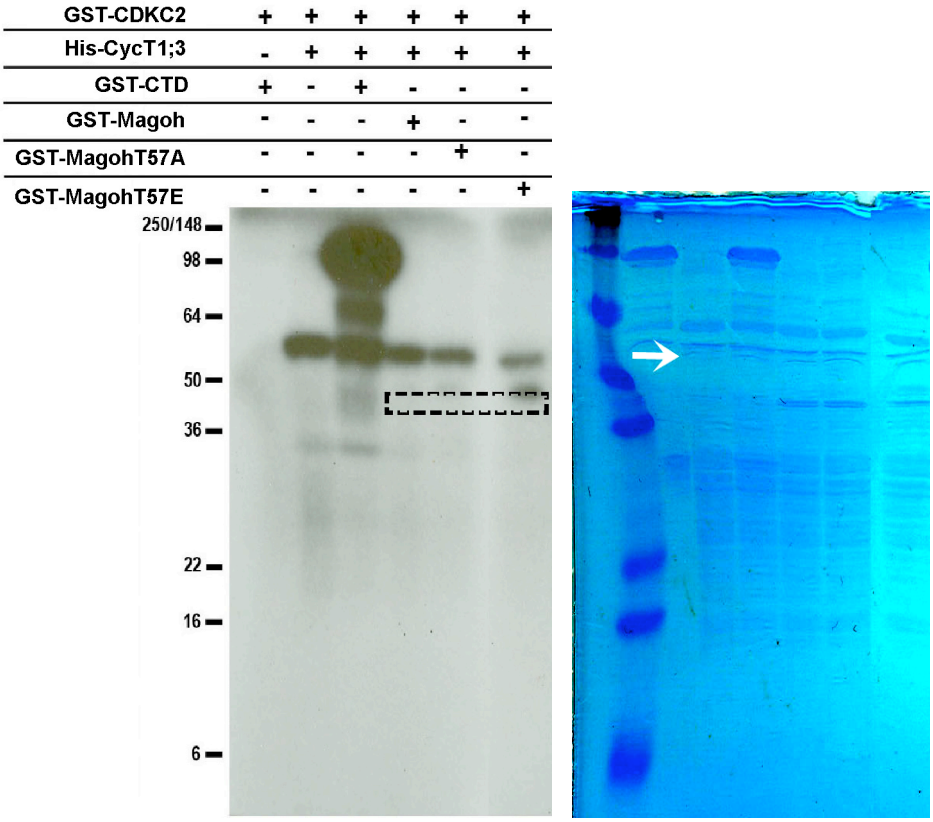
Figure 6.6 Kinase assay for the assessment of Magoh phosphorylation by CDKC2.

(A) Left Panel: Autoradiograph (exposed for 24 hours) of the dried gel shown in the right panel. Dashed rectangle indicates the expected position of the Magoh proteins.

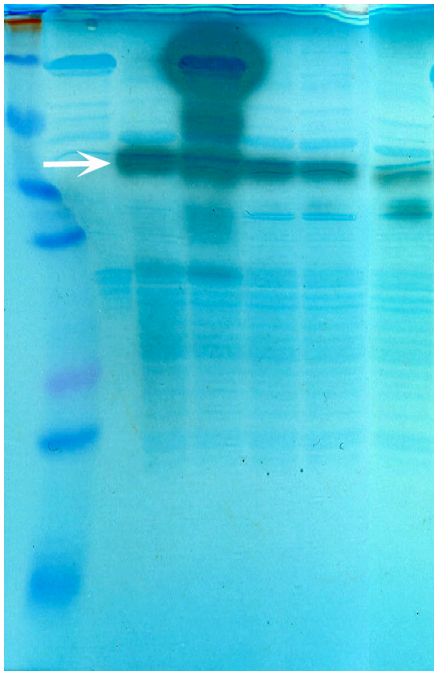
(B) Overlay of autoradiograph and gel to show the complete correspondence between the CycT1;3 band on the gel and band appearing on the film at the same molecular weight.

Above the film the protein components of each sample on each lane are given. Arrows in (A) and (B) indicate the position of CycT1;3 and the overlay between the CycT1;3 and the respective band on the X-ray film, respectively.

(A)


GST-CDKC2 + + + + + +
His-CycT1;3 - + + + + +
GST-CTD + - + - - -
GST-Magoh - - - + - -
GST-MagohT57A - - - - + -
GST-MagohT57E - - - - - +
250/148
98
64
50
36
22
16
6

(B)



6.4 Discussion

6.4.1 CyclinT/CDKC complexes are active in vitro

In the present chapter, I describe the generation of active CycT1;3/CDKC2 complexes *in vitro* that are able to phosphorylate the CTD peptide, a known substrate of CDKC2 (Barocco *et al*., 2003 and Cui *et al*., 2007). Kinase activity is dependent on the presence of both cyclin and CDK subunits without, apparently, the need for extra protein modifications or co-factors. Other recombinant Arabidopsis CDKs, including CDKA1, apart from their cyclin partner require additional co-factors for becoming active kinases (Max Bush, unpublished data). An active CycT1;3/CDKC2 complex is an important tool for screening putative kinase substrates or testing protein-protein interactions.

6.4.2 Putative substrates of the CycT1;3/CDKC2 complexes

The first candidate chosen was Magoh but it is not phosphorylated by the CycT1;3/CDKC2 complex, at least under these conditions. The interaction between CDKC2 and Magoh in a yeast two-hybrid assay (Ali Pendle, unpublished data) has at least two interpretations: first, the interaction in Y2H is a false positive. However, it was strong and reciprocal. Secondly, the two proteins could interact *in vivo* in a common complex but not necessarily as kinase and substrate, like in the case of mammalian CDK11, which forms a complex with transcription elongation factors but the kinase responsible for phosphorylation of these factors was CK2 kinase and not CDK11 (Trembley *et al*., 2003). Experimental verification of the co-existence in a complex of Magoh and CDKC2 could be done by reciprocal immunoprecipitation assays in cell cultures over-expressing different fusions of the two proteins. In addition, the availability of an active CycT1;3/CDKC2 complex opens up the way for identifying novel CDKC2 substrates, allowing as to obtain a more refined view of the functional importance of this kinase in nuclear functions.

In yeast and mammalian cells, cyclin phosphorylation by its associated CDK partner may be a mechanism for controlling the activity of the heterodimeric complex, where cyclin phosphorylation leads to degradation by the ubiquitin-proteasome pathway, leading to cyclin/CDK complex inactivation (Aviram *et al*., 2008; Lanker *et al*., 1996;

Clurman *et al*., 1996; Won and Reed, 1996). The identification of CycT1;3 as an *in vitro* substrate of CDKC2 suggests that a similar mechanism of cyclin/CDK inactivation may be conserved in plants. One common characteristic of all the autophosphorylated cyclins identified so far is that they do not contain a destruction box or PEST sequence at their N-terminus. Such sequences are primary signals for proteasome-mediated degradation and rapid protein turnover, respectively (Renaudin *et al*., 1996; Vandepoele *et al*., 2002; Rechsteiner and Rogers, 1996). CycT1;3 also does not contain a destruction box or a PEST sequence (Wang et al., 2004) at its N-terminus. However, it does contain an N-terminal putative CDK phosphorylation site (S45PSR48) that could be targeted by CDKC2.

In yeast, Pcl5 phosphorylation by Pho85 kinase is related to amino acid starvation (Shemer et al., 2002). Thus, it will be interesting to check the phosphorylation status and stability of CycT1;3 as Arabidopsis cell cultures progress from proliferation to the stationary phase. We also do not know if the formation of the CycT1;3/CDKC2 complex is a pre-requisite for cyclin phosphorylation. Such information will allow us to distinguish whether the phosphorylation occurs by a *trans* or a *cis* mechanism. The first mechanism is suited for cycling proteins, where a certain amount of the cyclin/CDK complex should accumulate for the phosphorylation to take place. In the latter case, the kinetics of phosphorylation is determined by the rate of complex formation, ensuring a limited lifetime of the cyclin/CDK complex. Considering that the CycT1;3 is not regarded as a cycling protein (Menges *et al*., 2005), the *cis* mechanism of phosphorylation is highly probable. Deletion of CycT1;3 cyclin box and determination of protein stability *in vitro* will be required to distinguish experimentally between these two mechanisms.

Chapter 7 – General Discussion

7.1 Transition between proliferation and quiescence in culture *versus* in plants

The transition between proliferation and quiescence lacks a firm conceptual or mechanistic framework, especially in plants. The available evidence suggests that plant cell cultures have very similar cell cycle machinery as in whole plants (De Veylder *et al*., 2007; Hartig and Beck, 2006) and one might expect that the same molecular components regulate the exit from and re-entry into cell cycle. Assuming that this is the case, plant cell cultures could provide an attractive model system in which to study this transition, especially when we consider the ability to synchronize cell cultures and, thus, follow this phase transition in a very controlled and reproducible manner.

Cell cycle activity changes during many normal developmental processes. Differentiation generally involves a reduction in cell cycle activity leading to exit from the cycle and, in some cases, cell death. However, many "quiescent" differentiated tissues can be induced to re-enter the cell cycle either by external stimuli or by developmental cues (e.g. division of the root quiescent centre on exposure to ethylene; Ortega-Martinez *et al*., 2007). Rapid re-entry during wounding of differentiated tissues is associated with strong induction of the *CDKA1* gene, a core regulator of cell cycle (Hemerly *et al*., 1993). This competence for division in differentiated cells mirrors the state of quiescent culture cells, since transferring the latter to fresh media induces cell division. Additional examples where exit from and entry into proliferation is part of the normal developmental process include bud dormancy as well as seed maturation/germination. In *Arabidopsis*, auxiliary buds develop from dedifferentiated cells in the cauline leaf axils of the inflorescence, which proliferate until the bud is formed and then enter a quiescent or dormant state (Shimizu-Sato and Mori, 2001). The dormant state is maintained by signals from the dominant apical bud (apical dominance) or by unfavourable environmental conditions, like low temperature, or both (Cline, 2000; Powell, 1988). Transcript levels of cell cycle-specific genes, like *histone H4*, *CYCB1;2* and *CYCD3;1*, are very low in dormant buds but, as soon as apical dominance is relieved, levels increase dramatically leading to proliferation and growth (Devitt and Stafstrom, 1995; Shimizu and Mori, 1998).

Seed filling represents a process that marks the passage from proliferation to quiescence of a developing seed (Repetto *et al*., 2008). During this process there are dramatic changes in transcript (Agrawal *et al*., 2008) and protein (Le *et al*., 2007) levels before seed maturation, at which stage protein synthesis ceases (Goldberg *et al*, 1994). Proteomic analysis of the nucleus at the seed filling stage of *Medicago truncatula* showed that composition of the nuclear proteome at this stage is considerably different from the nuclear proteome of mature *Medicago* seeds or *Arabidopsis* leaves highlighting the importance of subcellular fractionation in obtaining high-resolution proteomics data as well as the tissue-specific profile of the seed nuclear proteome (Repetto *et al*., 2008). In this study, ribosomal proteins appeared to be more abundant at the onset of seed filling, possibly highlighting the high demands for protein synthesis before seed maturation. Interestingly, transcript levels for these proteins actually declined at this specific stage (Gallardo *et al*., 2007), suggesting that the seed is retaining a pool of ribosomal proteins in the nucleus. The presence of a pool of ribosomal proteins was also observed in my quantitative proteomics analysis at proliferation and stationary cell cultures, again suggesting that cell cultures maintain a constant pool of ribosomal proteins that may allow for prompt cellular responses upon alteration of the physiological conditions of the culture, for example resuspension in fresh liquid medium.

Seed germination involves the re-initiation of cell proliferation (Georgieva *et al*., 1994; De Castro *et al*., 2000) and can be considered as a transition from a quiescent state to a proliferative one. During the first 16 hours of germination, root tip cells remain in a quiescent state, whereas after 24h they enter into the proliferation phase that will sustain the subsequent growth of the root (Sgorbati *et al*., 1988; Chiatante *et al*., 1989). 2D gel-based quantitative analysis of whole seed proteome in *Arabidopsis* showed that 74 out of 1,300 proteins identified by MALDI-ToF MS were differentially regulated prior to radicle emergence and during radicle protrusion (Gallardo *et al*., 2001). Most importantly, ACT7 and WD40-like, with confirmed roles in promoting cell division (Kost *et al*., 1999; McKhann *et al*., 1997), were among the up-regulated proteins, confirming the transition from quiescence to proliferation during seed germination. Despite the proteomics analysis of the seed nuclear proteome (Repetto *et al*., 2008), the dynamics of the seed nuclear proteome during germination has yet to be studied using

MS-based approaches, Gel-based estimations have reported an increase in the nuclear protein content during the proliferative phase of germination in pea root tips (Chiatante *et al*., 1991) but this study has very low resolution.

Seed germination is an attractive process for studying the transition of a plant tissue from proliferation to quiescence, since the process can be synchronized (Heydecker *et al*., 1973) and could be used for a more detailed dynamic and quantitative comparison of global transcriptomes and subcellular proteomes with cell suspensions. Such a comparison would allow one to assess the extent to which cultures can be used as a surrogate experimental system for studying reversible cell cycle exit/entry in plants.

7.2 Ribosome biogenesis and the cost of cellular homeostasis

The majority of the stable proteins between day2 and day6 nuclei were ribosomal and other proteins involved in protein synthesis. However, translation rates are very low during the stationary phase of a cell culture (Dickson *et al*., 1998; Fuge *et al*., 1994) raising the question of why such proteins remain at relatively high abundance in the nuclei of quiescent cells. One answer may lie in the cost of assembling ribosomes. The ribosome contains approximately 80 ribosomal proteins that represent 30-50% of the proteome in a growing yeast cell (Perry, 2007). Considering that the addition of one amino acid to a growing peptide chain consumes around 10 molecules of ATP (Hachiya *et al*., 2007; Noguchi *et al*., 2001), ribosome biogenesis uses energy that could otherwise be used for other cellular processes. The cell achieves energy economy by employing slow translation rates and controlling protein synthesis at the level of initiation (Beilharz and Preiss, 2007; Lackner *et al*., 2007). In actively growing yeast cells, despite the fact that there is an excess number of ribosomes, the ribosome density in polysomes is one fifth of the theoretical value - one ribosome every 156 nucleotides of the mRNA against the expected value of 5 ribosomes for the same length, considering that a ribosome roughly spans 35 nucleotides of the mRNA (Arava *et al*., 2003) - suggesting that translation is mostly regulated at the level of initiation (Arava *et al*., 2003; MacKay *et al*., 2004) and that plants may employ similar control mechanisms (Piques *et al*., 2009). Thus, quiescent cells will be able to promptly recover their

translation rates if conditions become favourable without the cost of completely re-synthesising their ribosomes.

Some quiescent plant tissues are also thought to be in a similar state of readiness; dessicated seeds contain many of the components required for resumption of protein synthesis and these are deployed during seed imbibition (Bewley, 1997). Spliceosomal components are also recruited upon resumption of proliferation activity. Transition from quiescence to proliferation in germinating tomato seeds leads to the redistribution of spliceosomal components from discrete perinucleolar-associated domains into finer nuclear particles (speckles) in nuclei (Echeverría *et al*., 2007). Thus, the relatively constant amount of RNA processing proteins between day2 and day6 nuclei suggests that there is a pool of RNA processing proteins retained in quiescent culture cells, ready to be activated/redistributed upon entry into proliferation. Regulation of splicing factor activity and distribution in mammalian cells is achieved by reversible phosphorylation (Sanford *et al*., 2005; Sanford and Bruzik, 1999). Spliceosomal components in plant cell cultures are sequestered into "megaspeckles" upon inhibition of the activity of CDKC2 kinase or when cellular transcription is inhibited (Chapter 5), suggesting similar post-translational regulatory mechanisms also occur in plants.

7.3 CDKC2 kinase and its role in pre-mRNA transcription and processing

As mentioned above, kinases can regulate protein activity without affecting protein abundance. CTD-kinases phosphorylate serine residues 2, 5 or 7 within the CTD heptad repeat 1-YSPTSPS-7 of the RNA polymerase II. Specific defects of *cdkc2*/CDKC1;RNAi plants in flowering time, viral resistance (Cui *et al*., 2007) and trichome development (this study) suggest that this CTD kinase family may differentially affect the transcription of specific genes and not global mRNA synthesis. Gene-specific transcriptional regulation has been reported for the KIN28 and CTK1 CTD-kinases in yeast (Kanin *et al*., 2007; Ahn *et al*., 2004) and this regulation appeared to be dependent on the structure of gene promoters that dictate the dynamics of serine phosphorylation (Kim *et al*., 2010). So far, no reports are available on which serine residue of the CTD is phosphorylated by CDKC2. Speculation on Ser7 phosphorylation by CDKC2 can only be made based on the facts that this kinase co-localizes with

spliceosomal components in cell culture (Chapter 5) and that modification of this residue is probably associated with a co-transcriptional assembly and disassembly of the splicing machinery in yeast (Kim *et al*., 2010). CDKC2 may phosphorylate the Ser2 of the CTD since transcription of intronless viral genes is taking place in a CDKC2-dependent manner (Cui *et al*., 2007). In yeast, the balance between Ser2 and Ser5 phosphorylation controls transcript elongation and termination, whereas phosphorylation of these residues is mediated by CTK1 and KIN28 kinases, respectively (Kim *et al*., 2010). It is tempting to speculate that the corresponding pair in *Arabidopsis* could be members of the CDKC and CDKD families, since members of the latter are known to phosphorylate the CTD domain and lie in the same phylogenetic clade with KIN28, when protein sequences of CTD-kinases from different eukaryotes were compared (Guo and Stiller, 2004).

7.4 Future directions

Understanding the mechanisms underlying the transition from proliferation to quiescence, in respect to changes in the nuclear proteome profiles, can give us an insight on how plants control the balance between meristematic and differentiated tissues. While suspension cells provide an easily accessible source of nuclear proteins from defined populations of cells, the isolation of nuclei from proliferating and differentiated plant tissues is now required. Enrichment of tissue specific nuclear proteomes can be achieved by using a technique called INTACT, which is based on tissue-specific expression of nuclear envelope protein fusions *in planta*, followed by fluorescence-activated cell sorting of nuclei (Birnbaum *et al*., 2003; Brady *et al*., 2007) and, with the choice of suitable promoters, could be employed to isolate proliferating and quiescent cells directly from plant tissues. Protein yields are likely to be very low but this approach could provide a much-needed comparison with cell suspensions.

Determining the targets of particular kinases such as CDKC during exit/re-entry processes would also be very interesting. Immunoprecipitation of CDKC-containing complexes from nuclear extracts followed by subsequent kinase assay and protein identification by MS-based analysis, would provide a list of potential substrates since many kinases co-precipitate with their substrates. However, kinase assays are prone to

false positive results (i.e. antibody specificity, unspecific binding to beads) and identified proteins will need to be verified independently. Another method that is reported to largely avoid the recovery of false positive substrates would be to generate analog-sensitive mutants of the kinase, (Gegan *et al*., 2007; Ubersax *et al*., 2003). These mutants are unable to bind wild type ATP and can only utilize a chemical variant of ATP, called "ATP-analog" which in turn cannot be used by the wild type kinases. However, it is not always possible to predict the precise mutation that will produce the desired change in substrate utilisation for any particular kinase, and this needs to be determined empirically. Indeed no previous reports have been published that utilize this method in plants.

APPENDICES

Contents of Appendices

Appendix-I

(*Provided in the form of CD*)

Sheet1
Protein report for the K003, K004 and K005 after running them together in Scaffold using a common set of parameters

Sheet2
Peptide report for the dataset in Sheet1

Sheet3
Functional categorization of the proteins from the K005 experiment

Appendix-II

(*Provided in the form of CD*)

Sheet1
Protein report from the K005 experiment that were quantified by Scaffold

Sheet2
Proteins from Sheet1 that were also identified by Scaffold during the combined analysis of the K003, K004 and K005 experiments

Sheet3
Peptide report for proteins of Sheet1

Sheet4
Categorization of proteins from Sheet3 in "up-regulated", "down-regulated" and "constant" and their respective functional categorization. Also an extra column presents the number of unique peptides that were used by Scaffold to quantify each protein

Appendix-III

(*Provided in the form of CD*)

Unknown proteins present in the K005 sample

Appendix-IV

Sequencing of the CTD peptide in the pGEX4T-3 vector using the pGEX-for and pGEX-rev primers

High quality sequencing results were isolated, based on the quality of the chromatograms obtained using both primers, were used to generate a consensus nucleotide sequence, which was translated at +1 frame. Below is the result of translation.

Amino acids in green boxes indicate the presence of a conserved CTD motif (YSPTSPS); amino acids in red boxes indicate the presence of a CTD motif lacking the last conserved amino acid (Serine; S); amino acids in the blue box indicate the end of the GST tag, showing that the translation of the consensus sequence is in frame with the tag.

REFERENCES

Aebersold, R. (2009). A stress test for mass spectrometry-based proteomics. Nat Methods 6, 411-2.

Agrawal, G. K., Hajduch, M., Graham, K. and Thelen, J. J. (2008). In-depth investigation of the soybean seed-filling proteome and comparison with a parallel study of rapeseed. Plant Physiol 148, 504-18.

Ahn, S. H., Kim, M. and Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. Mol Cell 13, 67-76.

Akoulitchev, S., Chuikov, S. and Reinberg, D. (2000). TFIIH is negatively regulated by cdk8-containing mediator complexes. Nature 407, 102-6.

Ali, G. S. and Reddy, A. S. (2006). ATP, phosphorylation and transcription regulate the mobility of plant splicing factors. J Cell Sci 119, 3527-38.

An, G., P.R. Ebert, A. Mitra, and S.B. Ha. (1988). Binary vectors. In Plant Molecular Biology Manual, vol. A3 (ed. G. a. S. Gelvin, R.), pp. 1-19: Kluwer.

Andersen, J. S., Lam, Y. W., Leung, A. K., Ong, S. E., Lyon, C. E., Lamond, A. I. and Mann, M. (2005). Nucleolar proteome dynamics. Nature 433, 77-83.

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

Andrade, L. E., Tan, E. M. and Chan, E. K. (1993). Immunocytochemical analysis of the coiled body in the cell cycle and during cell proliferation. Proc Natl Acad Sci U S A 90, 1947-51.

Antonov, A. V., Dietmann, S., Rodchenkov, I. and Mewes, H. W. (2009). PPI spider: a tool for the interpretation of proteomics data in the context of protein-protein interaction networks. Proteomics 9, 2740-9.

Arava, Y., Wang, Y., Storey, J. D., Liu, C. L., Brown, P. O. and Herschlag, D. (2003). Genome-wide analysis of mRNA translation profiles in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 100, 3889-94.

Arciga-Reyes, L., Wootton, L., Kieffer, M. and Davies, B. (2006). UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in Arabidopsis. Plant J 47, 480-9.

Aviram, S., Simon, E., Gildor, T., Glaser, F. and Kornitzer, D. (2008). Autophosphorylation-induced degradation of the Pho85 cyclin Pcl5 is essential for response to amino acid limitation. Mol Cell Biol 28, 6858-69.

Bae, M. S., Cho, E. J., Choi, E. Y. and Park, O. K. (2003). Analysis of the Arabidopsis nuclear proteome and its response to cold stress. Plant J 36, 652-63.

Baerenfaller, K., Grossmann, J., Grobei, M. A., Hull, R., Hirsch-Hoffmann, M., Yalovsky, S., Zimmermann, P., Grossniklaus, U., Gruissem, W. and Baginsky, S. (2008). Genome-scale proteomics reveals Arabidopsis thaliana gene models and proteome dynamics. Science 320, 938-41.

Bantscheff, M., Boesche, M., Eberhard, D., Matthieson, T., Sweetman, G. and Kuster, B. (2008). Robust and sensitive iTRAQ quantification on an LTQ Orbitrap mass spectrometer. Mol Cell Proteomics 7, 1702-13.

Baricheva, E. A., Berrios, M., Bogachev, S. S., Borisevich, I. V., Lapik, E. R., Sharakhov, I. V., Stuurman, N. and Fisher, P. A. (1996). DNA from Drosophila melanogaster beta-heterochromatin binds specifically to nuclear lamins in vitro and the nuclear envelope in situ. Gene 171, 171-6.

Barneche, F., Steinmetz, F. and Echeverria, M. (2000). Fibrillarin genes encode both a conserved nucleolar protein and a novel small nucleolar RNA involved in ribosomal RNA methylation in Arabidopsis thaliana. J Biol Chem 275, 27212-20.

Barroco, R. M., De Veylder, L., Magyar, Z., Engler, G., Inze, D. and Mironov, V. (2003). Novel complexes of cyclin-dependent kinases and a cyclin-like protein from Arabidopsis thaliana with a function unrelated to cell division. Cell Mol Life Sci 60, 401-12.

Bartl, S., Taplick, J., Lagger, G., Khier, H., Kuchler, K. and Seiser, C. (1997). Identification of mouse histone deacetylase 1 as a growth factor-inducible gene. Mol Cell Biol 17, 5033-43.

Beilharz, T. H. and Preiss, T. (2007). Widespread use of poly(A) tail length control to accentuate expression of the yeast transcriptome. RNA 13, 982-97.

Bell, A. W., Deutsch, E. W., Au, C. E., Kearney, R. E., Beavis, R., Sechi, S., Nilsson, T. and Bergeron, J. J. (2009). A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. Nat Methods 6, 423-30.

Bell, P. (2001). Viral eukaryogenesis: was the ancestor of the nucleus a complex DNA virus? Journal of Molecular Evolution 53, 251-256.

Benjamini, Y. a. H., Y. (1995). Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. Journal of Royal Statistical Society 57, 289-300.

Bentley, D. L. (2005). Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. Curr Opin Cell Biol 17, 251-6.

Berr, A. and Schubert, I. (2007). Interphase chromosome arrangement in Arabidopsis thaliana is similar in differentiated and meristematic tissues and shows a transient mirror symmetry after nuclear division. Genetics 176, 853-63.

Berro, R., Pedati, C., Kehn-Hall, K., Wu, W., Klase, Z., Even, Y., Geneviere, A. M., Ammosova, T., Nekhai, S. and Kashanchi, F. (2008). CDK13, a new potential human immunodeficiency virus type 1 inhibitory factor regulating viral mRNA splicing. J Virol 82, 7155-66.

Bewley, J. D. (1997). Seed Germination and Dormancy. Plant Cell 9, 1055-1066.

Bianchi, M. E. and Agresti, A. (2005). HMG proteins: dynamic players in gene regulation and differentiation. Curr Opin Genet Dev 15, 496-506.

Bindschedler, L. V., Palmblad, M. and Cramer, R. (2008). Hydroponic isotope labelling of entire plants (HILEP) for quantitative plant proteomics; an oxidative stress case study. Phytochemistry 69, 1962-72.
Birnbaum, K., Shasha, D. E., Wang, J. Y., Jung, J. W., Lambert, G. M., Galbraith, D. W. and Benfey, P. N. (2003). A gene expression map of the Arabidopsis root. Science 302, 1956-60.

Bollman, K. M., Aukerman, M. J., Park, M. Y., Hunter, C., Berardini, T. Z. and Poethig, R. S. (2003). HASTY, the Arabidopsis ortholog of exportin 5/MSN5, regulates phase change and morphogenesis. Development 130, 1493-504.

Boube, M., Joulia, L., Cribbs, D.L. and Bourbon, H.M. (2002) Evidence for a mediator of RNA polymerase II transcriptional regulation conserved from yeast to man. Cell, 110, 143-151.

Boudonck, K., Dolan, L. and Shaw, P. J. (1998). Coiled body numbers in the Arabidopsis root epidermis are regulated by cell type, developmental stage and cell cycle parameters. J Cell Sci 111 (Pt 24), 3687-94.

Boudonck, K., Dolan, L. and Shaw, P. J. (1999). The movement of coiled bodies visualized in living plant cells by the green fluorescent protein. Mol Biol Cell 10, 2297-307.

Bradford, M. (1976). A rapid and sensitive method for quantitation of microgram quantities of protein utilizing the principle of protein-dye-binding. Anal Biochem 72, 248-254.

Brady, S. M., Orlando, D. A., Lee, J. Y., Wang, J. Y., Koch, J., Dinneny, J. R., Mace, D., Ohler, U. and Benfey, P. N. (2007). A high-resolution root spatiotemporal map reveals dominant expression patterns. Science 318, 801-6.

Bregman, D. B., Du, L., van der Zee, S. and Warren, S. L. (1995). Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. J Cell Biol 129, 287-98.

Bridger, J. M., Boyle, S., Kill, I. R. and Bickmore, W. A. (2000). Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts. Curr Biol 10, 149-52.

Brotherton, D. H., Dhanaraj, V., Wick, S., Brizuela, L., Domaille, P. J., Volyanik, E., Xu, X., Parisini, E., Smith, B. O., Archer, S. J. et al. (1998). Crystal structure of the complex of the cyclin D-dependent kinase Cdk6 bound to the cell-cycle inhibitor p19INK4d. Nature 395, 244-50.

Brown, E. B., Wu, E. S., Zipfel, W. and Webb, W. W. (1999). Measurement of molecular diffusion in solution by multiphoton fluorescence photobleaching recovery. Biophys J 77, 2837-49.

Brown, J. M., Green, J., das Neves, R. P., Wallace, H. A., Smith, A. J., Hughes, J., Gray, N., Taylor, S., Wood, W. G., Higgs, D. R. et al. (2008). Association between active genes occurs at nuclear speckles and is modulated by chromatin environment. J Cell Biol 182, 1083-97.

Brown, J. W. and Shaw, P. J. (2008). The role of the plant nucleolus in pre-mRNA processing. Curr Top Microbiol Immunol 326, 291-311.

Brown, J. W., Clark, G. P., Leader, D. J., Simpson, C. G. and Lowe, T. (2001). Multiple snoRNA gene clusters from Arabidopsis. RNA 7, 1817-32.

Brown, K. E., Guest, S. S., Smale, S. T., Hahm, K., Merkenschlager, M. and Fisher, A. G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell 91, 845-54.

Browning, K. S. (2004). Plant translation initiation factors: it is not easy to be green. Biochem Soc Trans 32, 589-91.

Brugger, B., Erben, G., Sandhoff, R., Wieland, F. T. and Lehmann, W. D. (1997). Quantitative analysis of biological membrane lipids at the low picomole level by nano-electrospray ionization tandem mass spectrometry. Proc Natl Acad Sci U S A 94, 2339-44.

Buggy, Y., Maguire, T. M., McDermott, E., Hill, A. D., O'Higgins, N. and Duffy, M. J. (2006). Ets2 transcription factor in normal and neoplastic human breast tissue. Eur J Cancer 42, 485-91.

Bunnell, B. A., Heath, L. S., Adams, D. E., Lahti, J. M. and Kidd, V. J. (1990). Increased expression of a 58-kDa protein kinase leads to changes in the CHO cell cycle. Proc Natl Acad Sci U S A 87, 7467-71.
Buratowski, S., Hahn, S., Guarente, L. and Sharp, P. A. (1989). Five intermediate complexes in transcription initiation by RNA polymerase II. Cell 56, 549-61.

Bush, M. S., Hutchins, A. P., Jones, A. M., Naldrett, M. J., Jarmolowski, A., Lloyd, C. W. and Doonan, J. H. (2009). Selective recruitment of proteins to 5' cap complexes during the growth cycle in Arabidopsis. Plant J 59, 400-12.

Calikowski, T. T., Meulia, T. and Meier, I. (2003). A proteomic study of the arabidopsis nuclear matrix. J Cell Biochem 90, 361-78.

Calvert, M. E., Keck, K. M., Ptak, C., Shabanowitz, J., Hunt, D. F. and Pemberton, L. F. (2008). Phosphorylation by casein kinase 2 regulates Nap1 localization and function. Mol Cell Biol 28, 1313-25.

Canetta, E., Kim, S. H., Kalinina, N. O., Shaw, J., Adya, A. K., Gillespie, T., Brown, J. W. and Taliansky, M. (2008). A plant virus movement protein forms ringlike complexes with the major nucleolar protein, fibrillarin, in vitro. J Mol Biol 376, 932-7.

Cao, W., Jamison, S. F. and Garcia-Blanco, M. A. (1997). Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. RNA 3, 1456-67.

Cavalier-Smith, T. (2005). Economy, speed and size matter: evolutionary forces driving nuclear genome miniaturization and expansion. Ann Bot 95, 147-75.

Champion, A., Kreis, M., Mockaitis, K., Picaud, A. and Henry, Y. (2004). Arabidopsis kinome: after the casting. Funct Integr Genomics 4, 163-87.

Charlton, W. L., Johnson, B., Graham, I. A. and Baker, A. (2005). Non-coordinate expression of peroxisome biogenesis, beta-oxidation and glyoxylate cycle genes in mature Arabidopsis plants. Plant Cell Rep 23, 647-53.

Chen, H. H., Wang, Y. C. and Fann, M. J. (2006). Identification and characterization of the CDK12/cyclin L1 complex involved in alternative splicing regulation. Mol Cell Biol 26, 2736-45.

Chen, H. T., Warfield, L. and Hahn, S. (2007). The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. Nat Struct Mol Biol 14, 696-703.

Chen, X. (2005). MicroRNA biogenesis and function in plants. FEBS Lett 579, 5923-31.

Chiatante, D. B., P.; Levi, M and Sparvoli, E. (1991). Nuclear Proteins During the Onset of Cell Proliferation in Pea Root Meristems. J Exp Bot 42, 44-50.

Chiatante, D. L., M.; Sgorbati, S. and Sparvoli, E. (1989). Variations of phosphorylation of protein membrane components in pea root meristems during germination. Journal of Plant Growth Regulation 8, 107-119.

Ching, Y. P., Zhou, H. J., Yuan, J. G., Qiang, B. Q., Kung Hf, H. F. and Jin, D. Y. (2002). Identification and characterization of FTSJ2, a novel human nucleolar protein homologous to bacterial ribosomal RNA methyltransferase. Genomics 79, 2-6.

Choder, M. (1991). A general topoisomerase I-dependent transcriptional repression in the stationary phase in yeast. Genes Dev 5, 2315-26.

Clarke, J. H., Tack, D., Findlay, K., Van Montagu, M. and Van Lijsebettens, M. (1999). The SERRATE locus controls the formation of the early juvenile leaves and phase length in Arabidopsis. Plant J 20, 493-501.

Cline, M. G. (2000). Execution of the auxin replacement apical dominance experiment in temperate woody species. Am J Bot 87, 182-190.

Clurman, B. E., Sheaff, R. J., Thress, K., Groudine, M. and Roberts, J. M. (1996). Turnover of cyclin E by the ubiquitin-proteasome pathway is regulated by cdk2 binding and cyclin phosphorylation. Genes Dev 10, 1979-90.

Cmarko, D., Smigova, J., Minichova, L. and Popov, A. (2008). Nucleolus: the ribosome factory. Histol Histopathol 23, 1291-8.

Cokol, M., Nair, R. and Rost, B. (2000). Finding nuclear localization signals. EMBO Rep 1, 411-5.

Collier, S., Pendle, A., Boudonck, K., van Rij, T., Dolan, L. and Shaw, P. (2006). A distant coilin homologue is required for the formation of cajal bodies in Arabidopsis. Mol Biol Cell 17, 2942-51.

Colwill, K., Pawson, T., Andrews, B., Prasad, J., Manley, J. L., Bell, J. C. and Duncan, P. I. (1996). The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. EMBO J 15, 265-75.

Conrads, T. P., Alving, K., Veenstra, T. D., Belov, M. E., Anderson, G. A., Anderson, D. J., Lipton, M. S., Pasa-Tolic, L., Udseth, H. R., Chrisler, W. B. et al. (2001). Quantitative analysis of bacterial and mammalian proteomes using a combination of cysteine affinity tags and 15N-metabolic labeling. Anal Chem 73, 2132-9.

Conti, E. and Izaurralde, E. (2005). Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. Curr Opin Cell Biol 17, 316-25.

Corden, J. L. (1990). Tails of RNA polymerase II. Trends Biochem Sci 15, 383-7.
Cremer, T., Cremer, M., Dietzel, S., Muller, S., Solovei, I. and Fakan, S. (2006). Chromosome territories--a functional nuclear landscape. Curr Opin Cell Biol 18, 307-16.

Cui, X., Fan, B., Scholz, J. and Chen, Z. (2007). Roles of Arabidopsis cyclin-dependent kinase C complexes in cauliflower mosaic virus infection, plant growth, and development. Plant Cell 19, 1388-402.

Dahl, K. N., Ribeiro, A. J. and Lammerding, J. (2008). Nuclear shape, mechanics, and mechanotransduction. Circ Res 102, 1307-18.

Day, P. J., Cleasby, A., Tickle, I. J., O'Reilly, M., Coyle, J. E., Holding, F. P., McMenamin, R. L., Yon, J., Chopra, R., Lengauer, C. et al. (2009). Crystal structure of human CDK4 in complex with a D-type cyclin. Proc Natl Acad Sci U S A 106, 4166-70.

de Beus, E., Brockenbrough, J. S., Hong, B. and Aris, J. P. (1994). Yeast NOP2 encodes an essential nucleolar protein with homology to a human proliferation marker. J Cell Biol 127, 1799-813.

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

de Carcer, G., Cerdido, A. and Medina, F. J. (1997). NopA64, a novel nucleolar phosphoprotein from proliferating onion cells, sharing immunological determinants with mammalian nucleolin. Planta 201, 487-95.

De Castro RD, van Lammeren AAM, Groot SPC, Bino RJ, Hilhorst HWM (2000) Cell division and subsequent radicle protrusion in tomato seeds are inhibited by osmótica stress but DNA synthesis and formation of microtubular cytoskeleton are not. Plant Physiol 122: 327–335

de Hoffmann, E. S., V. . (2003). Mass Spectrometry: Principles and Applications (Second ed.). John Wiley & Sons, Ltd.. p. 65.

de la Fuente van Bentem, S., Anrather, D., Roitinger, E., Djamei, A., Hufnagl, T., Barta, A., Csaszar, E., Dohnal, I., Lecourieux, D. and Hirt, H. (2006) Phosphoproteomics reveals extensive in vivo phosphorylation of Arabidopsis proteins envolved in RNA metabolismo. Nuc. Acids Res. 34, 3267-3278.

de Nooijer, S., Wellink, J., Mulder, B. and Bisseling, T. (2009). Non-specific interactions are sufficient to explain the position of heterochromatic chromocenters and nucleoli in interphase nuclei. Nucleic Acids Res 37, 3558-68.

De Veylder L, Beeckman T, Inzé D. (2007) The ins and outs of the plant cell cycle. Nat Rev Mol Cell Biol. 8, 655-665.

Deal, R. B. and Henikoff, S. (2010). The INTACT method for cell type-specific gene expression and chromatin profiling in Arabidopsis thaliana. Nat Protoc 6, 56-68.

Deighton, R. F., Kerr, L. E., Short, D. M., Allerhand, M., Whittle, I. R. and McCulloch, J. (2010). Network generation enhances interpretation of proteomic data from induced apoptosis. Proteomics 10, 1307-15.

Dellaire, G., Farrall, R. and Bickmore, W. A. (2003). The Nuclear Protein Database (NPD): sub-nuclear localisation and functional annotation of the nuclear proteome. Nucleic Acids Res 31, 328-30.

Deprost, D., Yao, L., Sormani, R., Moreau, M., Leterreux, G., Nicolai, M., Bedu, M., Robaglia, C. and Meyer, C. (2007). The Arabidopsis TOR kinase links plant growth, yield, stress resistance and mRNA translation. EMBO Rep 8, 864-70.

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

Devitt, M. L. and Stafstrom, J. P. (1995). Cell cycle regulation during growth-dormancy cycles in pea axillary buds. Plant Mol Biol 29, 255-65.

Dickinson, J. a. F., S. (2002). Quantification of proteins on western blots using ECL. In The Protein Protocols, (ed. J. Walker). Totowa, New Jersey: Humana Press.

Dickson, L. M. and Brown, A. J. (1998). mRNA translation in yeast during entry into stationary phase. Mol Gen Genet 259, 282-93.

Dieci, G., Preti, M. and Montanini, B. (2009). Eukaryotic snoRNAs: a paradigm for gene expression flexibility. Genomics 94, 83-8.
DNA and nuclear protein content in plant tissue. Protoplasma 144, 180-184.

Docquier, S., Tillemans, V., Deltour, R. and Motte, P. (2004). Nuclear bodies and compartmentalization of pre-mRNA splicing factors in higher plants. Chromosoma 112, 255-66.

Douglas, D. J., Frank, A. J. and Mao, D. (2005). Linear ion traps in mass spectrometry. Mass Spectrom Rev 24, 1-29.

Dreyfuss, G., Kim, V. N. and Kataoka, N. (2002). Messenger-RNA-binding proteins and the messages they carry. Nat Rev Mol Cell Biol 3, 195-205.

Dunkley, T. P., Hester, S., Shadforth, I. P., Runions, J., Weimar, T., Hanton, S. L., Griffin, J. L., Bessant, C., Brandizzi, F., Hawes, C. et al. (2006). Mapping the Arabidopsis organelle proteome. Proc Natl Acad Sci U S A 103, 6518-23.

Duroux, M., Houben, A., Ruzicka, K., Friml, J. and Grasser, K. D. (2004). The chromatin remodelling complex FACT associates with actively transcribed regions of the Arabidopsis genome. Plant J 40, 660-71.

Echeverria, O., Vazquez-Nin, G., Juarez-Chavero, S. and Moreno Diaz de la Espina, S. (2007). Firing of transcription and compartmentalization of splicing factors in tomato radicle nuclei during germination(1). Biol Cell 99, 519-30.

Elmlund, H., Baraznenok, V., Lindahl, M., Samuelsen, C. O., Koeck, P. J., Holmberg, S., Hebert, H. and Gustafsson, C. M. (2006). The cyclin-dependent kinase 8 module sterically blocks Mediator interactions with RNA polymerase II. Proc Natl Acad Sci U S A 103, 15788-93.

Emmott, E., Dove, B. K., Howell, G., Chappell, L. A., Reed, M. L., Boyne, J. R., You, J. H., Brooks, G., Whitehouse, A. and Hiscox, J. A. (2008). Viral nucleolar localisation signals determine dynamic trafficking within the nucleolus. Virology 380, 191-202.

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, 389-96.

Even, Y., Durieux, S., Escande, M. L., Lozano, J. C., Peaucellier, G., Weil, D. and Geneviere, A. M. (2006). CDC2L5, a Cdk-like kinase with RS domain, interacts with the ASF/SF2-associated protein p32 and affects splicing in vivo. J Cell Biochem 99, 890-904.

Fang, Y. and Spector, D. L. (2007). Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living Arabidopsis plants. Curr Biol 17, 818-23.

Fang, Y., Hearn, S. and Spector, D. L. (2004). Tissue-specific expression and dynamic organization of SR splicing factors in Arabidopsis. Mol Biol Cell 15, 2664-73.

Fatica, A. and Tollervey, D. (2002) Making ribosomes. Curr Opin Cell Biol, 14, 313-318.
Faulk, W.P. and Taylor, G.M. (1971) An immunocolloid method for the electron microscope. Immunochemistry, 8, 1081-1083.

Feaver, W. J., Svejstrup, J. Q., Henry, N. L. and Kornberg, R. D. (1994). Relationship of CDK-activating kinase and RNA polymerase II CTD kinase TFIIH/TFIIK. Cell 79, 1103-9.

Feldherr, C. M. and Akin, D. (1990). The permeability of the nuclear envelope in dividing and nondividing cell cultures. J Cell Biol 111, 1-8.

Feng, B., Li, L., Zhou, X., Stanley, B. and Ma, H. (2009). Analysis of the Arabidopsis floral proteome: detection of over 2 000 proteins and evidence for posttranslational modifications. J Integr Plant Biol 51, 207-23.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F. and Whitehouse, C. M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. Science 246, 64-71.

Fenyo, D. and Beavis, R. C. (2003). A method for assessing the statistical significance of mass spectrometry-based protein identifications using general scoring schemes. Anal Chem 75, 768-74.

Ferreira-Cerca, S., Poll, G., Gleizes, P. E., Tschochner, H. and Milkereit, P. (2005). Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. Mol Cell 20, 263-75.

Fink, J. L., Karunaratne, S., Mittal, A., Gardiner, D. M., Hamilton, N., Mahony, D., Kai, C., Suzuki, H., Hayashizaki, Y. and Teasdale, R. D. (2008). Towards defining the nuclear proteome. Genome Biol 9, R15.

Firestein, R., Bass, A.J., Kim, S.Y., Dunn, I.F., Silver, S.J., Guney, I., Freed, E., Ligon, A.H., Vena, N., Ogino, S., Chheda, M.G., Tamayo, P., Finn, S., Shrestha, Y., Boehm, J.S., Jain, S., Bojarski, E., Mermel, C., Barretina, J., Chan, J.A., Baselga, J., Tabernero, J., Root, D.E., Fuchs, C.S., Loda, M., Shivdasani, R.A., Meyerson, M. and Hahn, W.C. (2008) CDK8 is a colorectal cancer oncogene that regulates beta-catenin activity. Nature, 455, 547-551.

Flanagan, P. M., Kelleher, R. J., 3rd, Sayre, M. H., Tschochner, H. and Kornberg, R. D. (1991). A mediator required for activation of RNA polymerase II transcription in vitro. Nature 350, 436-8.

Flores, C. L., Rodriguez, C., Petit, T. and Gancedo, C. (2000). Carbohydrate and energy-yielding metabolism in non-conventional yeasts. FEMS Microbiol Rev 24, 507-29.

Fraser, P. and Bickmore, W. (2007). Nuclear organization of the genome and the potential for gene regulation. Nature 447, 413-7.

Fricker, M., Hollinshead, M., White, N. and Vaux, D. (1997). Interphase nuclei of many mammalian cell types contain deep, dynamic, tubular membrane-bound invaginations of the nuclear envelope. J Cell Biol 136, 531-44.

Fromont-Racine, M., Senger, B., Saveanu, C. and Fasiolo, F. (2003). Ribosome assembly in eukaryotes. Gene 313, 17-42.

Fuerst, J. A. (2005). Intracellular compartmentation in planctomycetes. Annu Rev Microbiol 59, 299-328.

Fuge, E. K., Braun, E. L. and Werner-Washburne, M. (1994). Protein synthesis in long-term stationary-phase cultures of Saccharomyces cerevisiae. J Bacteriol 176, 5802-13.

Fujinaga, K., Irwin, D., Huang, Y., Taube, R., Kurosu, T. and Peterlin, B.M. (2004) Dynamics of human immunodeficiency virus transcription: P-TEFb phosphorylates RD and dissociates negative effectors from the transactivation response element. Mol Cell Biol, 24, 787-795.

Fulop, K., Pettko-Szandtner, A., Magyar, Z., Miskolczi, P., Kondorosi, E., Dudits, D. and Bako, L. (2005). The Medicago CDKC;1-CYCLINT;1 kinase complex phosphorylates the carboxy-terminal domain of RNA polymerase II and promotes transcription. Plant J 42, 810-20.

Futcher, B. (1996). Cyclins and the wiring of the yeast cell cycle. Yeast 12, 1635-46.

Galiova, G., Bartova, E., Raska, I., Krejci, J. and Kozubek, S. (2008). Chromatin changes induced by lamin A/C deficiency and the histone deacetylase inhibitor trichostatin A. Eur J Cell Biol 87, 291-303.

Gallardo, K., Firnhaber, C., Zuber, H., Hericher, D., Belghazi, M., Henry, C., Kuster, H. and Thompson, R. (2007). A combined proteome and transcriptome analysis of developing Medicago truncatula seeds: evidence for metabolic specialization of maternal and filial tissues. Mol Cell Proteomics 6, 2165-79.

Gan, C. S., Chong, P. K., Pham, T. K. and Wright, P. C. (2007). Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). J Proteome Res 6, 821-7.

Garriga, J., Mayol, X. and Grana, X. (1996) The CDC2-related kinase PITALRE is the catalytic subunit of active multimeric protein complexes. Biochem J, 319 (Pt 1), 293-298.

Garriga, J., Peng, J., Parreno, M., Price, D.H., Henderson, E.E. and Grana, X. (1998) Upregulation of cyclin T1/CDK9 complexes during T cell activation. Oncogene, 17, 3093-3102.

Gasch, A. P., Spellman, P. T., Kao, C. M., Carmel-Harel, O., Eisen, M. B., Storz, G., Botstein, D. and Brown, P. O. (2000). Genomic expression programs in the response of yeast cells to environmental changes. Mol Biol Cell 11, 4241-57.

Gentzel, M., Kocher, T., Ponnusamy, S. and Wilm, M. (2003). Preprocessing of tandem mass spectrometric data to support automatic protein identification. Proteomics 3, 1597-610.

Georgieva EI, Lo`pez-Rodas G, Hittmair A, Feichtinger H, Brosch G, Loidl P (1994) Maize embryo germination. Planta 192: 118–124

Gong, Z., Dong, C. H., Lee, H., Zhu, J., Xiong, L., Gong, D., Stevenson, B. and Zhu, J. K. (2005). A DEAD box RNA helicase is essential for mRNA export and important for development and stress responses in Arabidopsis. Plant Cell 17, 256-67.

Gonzalez-Camacho, F. and Medina, F. J. (2006). The nucleolar structure and the activity of NopA100, a nucleolin-like protein, during the cell cycle in proliferating plant cells. Histochem Cell Biol 125, 139-53.

Gonzalez-Melendi, P., Wells, B., Beven, A. F. and Shaw, P. J. (2001). Single ribosomal transcription units are linear, compacted Christmas trees in plant nucleoli. Plant J 27, 223-33.

Gonzalez, D., Bowen, A. J., Carroll, T. S. and Conlan, R. S. (2007). The transcription corepressor LEUNIG interacts with the histone deacetylase HDA19 and mediator components MED14 (SWP) and CDK8 (HEN3) to repress transcription. Mol Cell Biol 27, 5306-15.

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

Graves, B. J. and Petersen, J. M. (1998). Specificity within the ets family of transcription factors. Adv Cancer Res 75, 1-55.

Graves, B. J., Cowley, D. O., Goetz, T. L., Petersen, J. M., Jonsen, M. D. and Gillespie, M. E. (1998). Autoinhibition as a transcriptional regulatory mechanism. Cold Spring Harb Symp Quant Biol 63, 621-9.

Gray, J. V., Petsko, G. A., Johnston, G. C., Ringe, D., Singer, R. A. and Werner-Washburne, M. (2004). "Sleeping beauty": quiescence in Saccharomyces cerevisiae. Microbiol Mol Biol Rev 68, 187-206.

Gregan, J., Zhang, C., Rumpf, C., Cipak, L., Li, Z., Uluocak, P., Nasmyth, K., and Shokat, K. (2007) Construction of conditional analog-sensitive kinase alleles in the fission yeast Schizosaccharomyces pombe. Nat Protoc. 2, 2996-3000.

Gregory, T. (2005). Genome size evolution in animals. In In The Evolution of the Genome., vol. 1 (ed. T. Gregory), pp. 4-87: Elsevier Academic Press, London.

Griffiths, S. D., Burthem, J., Unwin, R. D., Holyoake, T. L., Melo, J. V., Lucas, G. S. and Whetton, A. D. (2007). The use of isobaric tag peptide labeling (iTRAQ) and mass spectrometry to examine rare, primitive hematopoietic cells from patients with chronic myeloid leukemia. Mol Biotechnol 36, 81-9.

Gruhler, A., Schulze, W. X., Matthiesen, R., Mann, M. and Jensen, O. N. (2005). Stable isotope labeling of Arabidopsis thaliana cells and quantitative proteomics by mass spectrometry. Mol Cell Proteomics 4, 1697-709.

Gui, J. F., Tronchere, H., Chandler, S. D. and Fu, X. D. (1994). Purification and characterization of a kinase specific for the serine- and arginine-rich pre-mRNA splicing factors. Proc Natl Acad Sci U S A 91, 10824-8.

Guiguen, A., Soutourina, J., Dewez, M., Tafforeau, L., Dieu, M., Raes, M., Vandenhaute, J., Werner, M. and Hermand, D. (2007). Recruitment of P-TEFb (Cdk9-Pch1) to chromatin by the cap-methyl transferase Pcm1 in fission yeast. EMBO J 26, 1552-9.

Guo, Z. and Stiller, J. W. (2004). Comparative genomics of cyclin-dependent kinases suggest co-evolution of the RNAP II C-terminal domain and CTD-directed CDKs. BMC Genomics 5, 69.

Gururajan, R., Lahti, J. M., Grenet, J., Easton, J., Gruber, I., Ambros, P. F. and Kidd, V. J. (1998). Duplication of a genomic region containing the Cdc2L1-2 and MMP21-22 genes on human chromosome 1p36.3 and their linkage to D1Z2. Genome Res 8, 929-39.

Guven-Ozkan, T., Nishi, Y., Robertson, S. M. and Lin, R. (2008). Global transcriptional repression in C. elegans germline precursors by regulated sequestration of TAF-4. Cell 135, 149-60.

Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H. and Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nat Biotechnol 17, 994-9.

Han, M. H., Goud, S., Song, L. and Fedoroff, N. (2004). The Arabidopsis double-stranded RNA-binding protein HYL1 plays a role in microRNA-mediated gene regulation. Proc Natl Acad Sci U S A 101, 1093-8.

Hanks, S. K. (2003). Genomic analysis of the eukaryotic protein kinase superfamily: a perspective. Genome Biol 4, 111.

Hardcastle, I.R., Golding, B.T. and Griffin, R.J. (2002) Designing inhibitors of cyclin-dependent kinases. Annu Rev Pharmacol Toxicol, 42, 325-348.

Harris, H. (1967). The reactivation of the red cell nucleus. J Cell Sci 2, 23-32.

Harris, S. D. (1999). Morphogenesis is coordinated with nuclear division in germinating Aspergillus nidulans conidiospores. Microbiology 145 (Pt 10), 2747-56.

Hartig K, Beck E. (2006) Crosstalk between auxin, cytokinins, and sugars in the plant cell cycle. Plant Biol (Stuttg) 8, 389-96.

Hata, S. (1991) cDNA cloning of a novel cdc2+/CDC28-related protein kinase from rice. FEBS Lett, 279, 149-52.

Haustein, E. and Schwille, P. (2007). Fluorescence correlation spectroscopy: novel variations of an established technique. Annu Rev Biophys Biomol Struct 36, 151-69.

Hawkins, J., Davis, L. and Boden, M. (2007). Predicting nuclear localization. J Proteome Res 6, 1402-9.

Heard, E. and Bickmore, W. (2007). The ins and outs of gene regulation and chromosome territory organisation. Curr Opin Cell Biol 19, 311-6.

Hebeler, R., Oeljeklaus, S., Reidegeld, K.A., Eisenacher, M., Stephan, C., Sitek, B., Stuhler, K., Meyer, H.E., Sturre, M.J., Dijkwel, P.P. and Warscheid, B. (2008) Study of early leaf senescence in Arabidopsis thaliana by quantitative proteomics using reciprocal 14N/15N labeling and difference gel electrophoresis. Mol Cell Proteomics, 7, 108-120.

Hedges, J., West, M. and Johnson, A. W. (2005). Release of the export adapter, Nmd3p, from the 60S ribosomal subunit requires Rpl10p and the cytoplasmic GTPase Lsg1p. EMBO J 24, 567-79.

Heidenhain, H. (1894)Arch. Mikroskop. Anat. , 43, 423.

Hemerly, A. S., Ferreira, P., de Almeida Engler, J., Van Montagu, M., Engler, G. and Inze, D. (1993). cdc2a expression in Arabidopsis is linked with competence for cell division. Plant Cell 5, 1711-23.

Hengartner, C. J., Myer, V. E., Liao, S. M., Wilson, C. J., Koh, S. S. and Young, R. A. (1998). Temporal regulation of RNA polymerase II by Srb10 and Kin28 cyclin-dependent kinases. Mol Cell 2, 43-53.

Henras, A. K., Soudet, J., Gerus, M., Lebaron, S., Caizergues-Ferrer, M., Mougin, A. and Henry, Y. (2008). The post-transcriptional steps of eukaryotic ribosome biogenesis. Cell Mol Life Sci 65, 2334-59.

Herman, P. K. (2002). Stationary phase in yeast. Curr Opin Microbiol 5, 602-7.

Heydecker W, Higgins J, Gulliver RL (1973) Accelerated germination by osmotic seed treatment. Nature

246, 42–44

Hiller, M., Chen, X., Pringle, M. J., Suchorolski, M., Sancak, Y., Viswanathan, S., Bolival, B., Lin, T. Y., Marino, S. and Fuller, M. T. (2004). Testis-specific TAF homologs collaborate to control a tissue-specific transcription program. Development 131, 5297-308.

Hirose, Y. and Manley, J. L. (2000). RNA polymerase II and the integration of nuclear events. Genes Dev 14, 1415-29.

Hirose, Y. and Manley, J.L. (2000) RNA polymerase II and the integration of nuclear events. Genes Dev, 14, 1415-1429.

Hochachka, P.W., Buck, L.T., Doll, C.J. and Land, S.C. (1996) Unifying theory of hypoxia tolerance: molecular/metabolic defense and rescue mechanisms for surviving oxygen lack. Proc Natl Acad Sci U S A, 93, 9493-9498.

Hochheimer, A. and Tjian, R. (2003). Diversified transcription initiation complexes expand promoter selectivity and tissue-specific gene expression. Genes Dev 17, 1309-20.

Hoffmann, K., Dreger, C. K., Olins, A. L., Olins, D. E., Shultz, L. D., Lucke, B., Karl, H., Kaps, R., Muller, D., Vaya, A. et al. (2002). Mutations in the gene encoding the lamin B receptor produce an altered nuclear morphology in granulocytes (Pelger-Huet anomaly). Nat Genet 31, 410-4.

Hori, K. and Watanabe, Y. (2005). UPF3 suppresses aberrant spliced mRNA in Arabidopsis. Plant J 43, 530-40.

Hu, D., Mayeda, A., Trembley, J. H., Lahti, J. M. and Kidd, V. J. (2003). CDK11 complexes promote pre-mRNA splicing. J Biol Chem 278, 8623-9.

Hu, Q., Noll, R. J., Li, H., Makarov, A., Hardman, M. and Graham Cooks, R. (2005). The Orbitrap: a new mass spectrometer. J Mass Spectrom 40, 430-43.

Hutchins, A. (2004) Identification of CDKA1-interacting proteins. PhD thesis, University of East Anglia, p. 124.

Iborra, F. J., Jackson, D. A. and Cook, P. R. (2001). Coupled transcription and translation within nuclei of mammalian cells. Science 293, 1139-42.

Ilgenfritz, H., Bouyer, D., Schnittger, A., Mathur, J., Kirik, V., Schwab, B., Chua, N. H., Jurgens, G. and

Hulskamp, M. (2003). The Arabidopsis STICHEL gene is a regulator of trichome branch number and encodes a novel protein. Plant Physiol 131, 643-55.

Ingle, R. A., Schmidt, U. G., Farrant, J. M., Thomson, J. A. and Mundree, S. G. (2007). Proteomic analysis of leaf proteins during dehydration of the resurrection plant Xerophyta viscosa. Plant Cell Environ 30, 435-46.

Inoue, H., Nojima, H. and Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. Gene 96, 23-8.

Iorns, E., Turner, N. C., Elliott, R., Syed, N., Garrone, O., Gasco, M., Tutt, A. N., Crook, T., Lord, C. J. and Ashworth, A. (2008). Identification of CDK10 as an important determinant of resistance to endocrine therapy for breast cancer. Cancer Cell 13, 91-104.

Ivorra, C., Kubicek, M., Gonzalez, J. M., Sanz-Gonzalez, S. M., Alvarez-Barrientos, A., O'Connor, J. E., Burke, B. and Andres, V. (2006). A mechanism of AP-1 suppression through interaction of c-Fos with lamin A/C. Genes Dev 20, 307-20.

Jakel, S. and Gorlich, D. (1998). Importin beta, transportin, RanBP5 and RanBP7 mediate nuclear import of ribosomal proteins in mammalian cells. EMBO J 17, 4491-502.

Jares-Erijman, E. A. and Jovin, T. M. (2003). FRET imaging. Nat Biotechnol 21, 1387-95.

Jeffrey, P. D., Russo, A. A., Polyak, K., Gibbs, E., Hurwitz, J., Massague, J. and Pavletich, N. P. (1995). Mechanism of CDK activation revealed by the structure of a cyclinA-CDK2 complex. Nature 376, 313-20.

Jones, A. M., Bennett, M. H., Mansfield, J. W. and Grant, M. (2006). Analysis of the defence phosphoproteome of Arabidopsis thaliana using differential mass tagging. Proteomics 6, 4155-65.

Jones, A. M., MacLean, D., Studholme, D. J., Serna-Sanz, A., Andreasson, E., Rathjen, J. P. and Peck, S. C. (2009). Phosphoproteomic analysis of nuclei-enriched fractions from Arabidopsis thaliana. J Proteomics 72, 439-51.

Jorgensen, P. and Tyers, M. (2004). How cells coordinate growth and division. Curr Biol 14, R1014-27.
Jorgensen, P., Edgington, N. P., Schneider, B. L., Rupes, I., Tyers, M. and Futcher, B. (2007). The size of the nucleus increases as yeast cells grow. Mol Biol Cell 18, 3523-32.

Joubes, J., Chevalier, C., Dudits, D., Heberle-Bors, E., Inze, D., Umeda, M. and Renaudin, J. P. (2000). CDK-related protein kinases in plants. Plant Mol Biol 43, 607-20.

Junera, H. R., Masson, C., Geraud, G. and Hernandez-Verdun, D. (1995). The three-dimensional organization of ribosomal genes and the architecture of the nucleoli vary with G1, S and G2 phases. J Cell Sci 108 (Pt 11), 3427-41.

Jurica, M. S. and Moore, M. J. (2003). Pre-mRNA splicing: awash in a sea of proteins. Mol Cell 12, 5-14.

Kalyna, M. and Barta, A. (2004). A plethora of plant serine/arginine-rich proteins: redundancy or evolution of novel gene functions? Biochem Soc Trans 32, 561-4.

Kanin, E. I., Kipp, R. T., Kung, C., Slattery, M., Viale, A., Hahn, S., Shokat, K. M. and Ansari, A. Z. (2007). Chemical inhibition of the TFIIH-associated kinase Cdk7/Kin28 does not impair global mRNA synthesis. Proc Natl Acad Sci U S A 104, 5812-7.

Karas, M. and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem 60, 2299-301.

Kasten, M. and Giordano, A. (2001). Cdk10, a Cdc2-related kinase, associates with the Ets2 transcription factor and modulates its transactivation activity. Oncogene 20, 1832-8.

Kataoka, N., Diem, M. D., Kim, V. N., Yong, J. and Dreyfuss, G. (2001). Magoh, a human homolog of Drosophila mago nashi protein, is a component of the splicing-dependent exon-exon junction complex. EMBO J 20, 6424-33.

Kato K, M. T., Koiwai S, Mizusaki S, Nishida K, Nogushi M, Tamaki E. (1972). Liquid suspension culture of tobaco cells. In Ferment Technology Today., (ed. G. Terui), pp. 689-695: Society of Fermentation Technology.

Kelleher, R. J., 3rd, Flanagan, P. M. and Kornberg, R. D. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. Cell 61, 1209-15.

Keller, A., Nesvizhskii, A. I., Kolker, E. and Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem 74, 5383-92.

Khan, M. M. and Komatsu, S. (2004). Rice proteomics: recent developments and analysis of nuclear proteins. Phytochemistry 65, 1671-81.

Kim, H., Erickson, B., Luo, W., Seward, D., Graber, J. H., Pollock, D. D., Megee, P. C. and Bentley, D. L. (2010). Gene-specific RNA polymerase II phosphorylation and the CTD code. Nat Struct Mol Biol 17, 1279-86.

Kim, Y. J., Bjorklund, S., Li, Y., Sayre, M. H. and Kornberg, R. D. (1994). A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. Cell 77, 599-608.

King, M. C., Raposo, G. and Lemmon, M. A. (2004). Inhibition of nuclear import and cell-cycle progression by mutated forms of the dynamin-like GTPase MxB. Proc Natl Acad Sci U S A 101, 8957-62.

Kirkland, D. J. and Kim, N. N. (1995). Special considerations for conducting genotoxicity tests with protein materials. Mutagenesis 10, 393-8.

Kiseleva, E., Allen, T. D., Rutherford, S. A., Murray, S., Morozova, K., Gardiner, F., Goldberg, M. W. and Drummond, S. P. (2007). A protocol for isolation and visualization of yeast nuclei by scanning electron microscopy (SEM). Nat Protoc 2, 1943-53.

Kitsios, G. (2007). Analysis of Cyclin-dendent kinases in Arabidopsis thaliana, vol. PhD (ed., pp. 224. Norwich: University of East Anglia.

Kitsios, G., Alexiou, K. G., Bush, M., Shaw, P. and Doonan, J. H. (2008). A cyclin-dependent protein kinase, CDKC2, colocalizes with and modulates the distribution of spliceosomal components in Arabidopsis. Plant J 54, 220-35.

Kleffmann, T., von Zychlinski, A., Russenberger, D., Hirsch-Hoffmann, M., Gehrig, P., Gruissem, W. and Baginsky, S. (2007). Proteome dynamics during plastid differentiation in rice. Plant Physiol 143, 912-23.

Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Xuong, N. H., Taylor, S. S. and Sowadski, J. M. (1991). Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science 253, 414-20.

Knuesel, M. T., Meyer, K. D., Bernecky, C. and Taatjes, D. J. (2009). The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. Genes Dev 23, 439-51.

Knuesel, M. T., Meyer, K. D., Donner, A. J., Espinosa, J. M. and Taatjes, D. J. (2009). The human CDK8 subcomplex is a histone kinase that requires Med12 for activity and can function independently of mediator. Mol Cell Biol 29, 650-61.

Ko, T. K., Kelly, E. and Pines, J. (2001). CrkRS: a novel conserved Cdc2-related protein kinase that colocalises with SC35 speckles. J Cell Sci 114, 2591-603.

Kobor, M. S. and Greenblatt, J. (2002). Regulation of transcription elongation by phosphorylation. Biochim Biophys Acta 1577, 261-275.

Kolla, V., Jeno, P., Moes, S., Tercanli, S., Lapaire, O., Choolani, M. and Hahn, S. (2010). Quantitative proteomics analysis of maternal plasma in Down syndrome pregnancies using isobaric tagging reagent (iTRAQ). J Biomed Biotechnol 2010, 952047.

Koncz, C., and Schell, J. (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a nove1 type of Agrobacterium binary vector. Mol Gen Genet 204, 383-396.

Kornblihtt, A. R., de la Mata, M., Fededa, J. P., Munoz, M. J. and Nogues, G. (2004). Multiple links between transcription and splicing. RNA 10, 1489-98.

Koroleva, O. A., Calder, G., Pendle, A. F., Kim, S. H., Lewandowska, D., Simpson, C. G., Jones, I. M., Brown, J. W. and Shaw, P. J. (2009). Dynamic behavior of Arabidopsis eIF4A-III, putative core protein of exon junction complex: fast relocation to nucleolus and splicing speckles under hypoxia. Plant Cell 21, 1592-606.

Koroleva, O. A., Tomlinson, M. L., Leader, D., Shaw, P. and Doonan, J. H. (2005). High-throughput protein localization in Arabidopsis using Agrobacterium-mediated transient expression of GFP-ORF fusions. Plant J 41, 162-74.

Kumaran, R. I., Thakar, R. and Spector, D. L. (2008). Chromatin dynamics and gene positioning. Cell 132, 929-34.

Kuroda, M., Tanabe, H., Yoshida, K., Oikawa, K., Saito, A., Kiyuna, T., Mizusawa, H. and Mukai, K. (2004). Alteration of chromosome positioning during adipocyte differentiation. J Cell Sci 117, 5897-903.
Lackner, D. H., Beilharz, T. H., Marguerat, S., Mata, J., Watt, S., Schubert, F., Preiss, T. and Bahler, J. (2007). A network of multiple regulatory layers shapes gene expression in fission yeast. Mol Cell 26, 145-55.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-5.

Lagger, G., O'Carroll, D., Rembold, M., Khier, H., Tischler, J., Weitzer, G., Schuettengruber, B., Hauser, C., Brunmeir, R., Jenuwein, T. et al. (2002). Essential function of histone deacetylase 1 in proliferation control and CDK inhibitor repression. EMBO J 21, 2672-81.

Lahti, J. M., Xiang, J., Heath, L. S., Campana, D. and Kidd, V. J. (1995). PITSLRE protein kinase activity is associated with apoptosis. Mol Cell Biol 15, 1-11.

Lammerding, J., Schulze, P. C., Takahashi, T., Kozlov, S., Sullivan, T., Kamm, R. D., Stewart, C. L. and Lee, R. T. (2004). Lamin A/C deficiency causes defective nuclear mechanics and mechanotransduction. J Clin Invest 113, 370-8.

Lamond, A. I. and Spector, D. L. (2003). Nuclear speckles: a model for nuclear organelles. Nat Rev Mol Cell Biol 4, 605-12.

Lanker, S., Valdivieso, M. H. and Wittenberg, C. (1996). Rapid degradation of the G1 cyclin Cln2 induced by CDK-dependent phosphorylation. Science 271, 1597-601.

Lanquar, V., Kuhn, L., Lelievre, F., Khafif, M., Espagne, C., Bruley, C., Barbier-Brygoo, H., Garin, J. and Thomine, S. (2007). 15N-metabolic labeling for comparative plasma membrane proteomics in Arabidopsis cells. Proteomics 7, 750-4.

Larochelle, S., Pandur, J., Fisher, R.P., Salz, H.K. and Suter, B. (1998) Cdk7 is essential for mitosis and for in vivo Cdk-activating kinase activity. Genes Dev, 12, 370-381.

Lau, S., Ehrismann, J. S., Schlereth, A., Takada, S., Mayer, U. and Jurgens, G. (2010). Cell-cell communication in Arabidopsis early embryogenesis. Eur J Cell Biol 89, 225-30.

Launholt, D., Merkle, T., Houben, A., Schulz, A. and Grasser, K. D. (2006). Arabidopsis chromatin-associated HMGA and HMGB use different nuclear targeting signals and display highly dynamic localization within the nucleus. Plant Cell 18, 2904-18.

Le Hir, H., Izaurralde, E., Maquat, L. E. and Moore, M. J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J 19, 6860-9.

Le Hir, H., Nott, A. and Moore, M. J. (2003). How introns influence and enhance eukaryotic gene expression. Trends Biochem Sci 28, 215-20.

Lee, H., Guo, Y., Ohta, M., Xiong, L., Stevenson, B. and Zhu, J. K. (2002). LOS2, a genetic locus required for cold-responsive gene transcription encodes a bi-functional enolase. EMBO J 21, 2692-702.

Lelievre, S. A. and Bissell, M. J. (1998). Communication between the cell membrane and the nucleus: role of protein compartmentalization. J Cell Biochem Suppl 30-31, 250-63.

Lerch-Gaggl, A., Haque, J., Li, J., Ning, G., Traktman, P. and Duncan, S. A. (2002). Pescadillo is essential for nucleolar assembly, ribosome biogenesis, and mammalian cell proliferation. J Biol Chem 277, 45347-55.

Leung, A. K. and Lamond, A. I. (2003). The dynamics of the nucleolus. Crit Rev Eukaryot Gene Expr 13, 39-54.

Li, C. F., Henderson, I. R., Song, L., Fedoroff, N., Lagrange, T. and Jacobsen, S. E. (2008). Dynamic regulation of ARGONAUTE4 within multiple nuclear bodies in Arabidopsis thaliana. PLoS Genet 4, e27.

Li, C. F., Pontes, O., El-Shami, M., Henderson, I. R., Bernatavichute, Y. V., Chan, S. W., Lagrange, T., Pikaard, C. S. and Jacobsen, S. E. (2006). An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in Arabidopsis thaliana. Cell 126, 93-106.

Li, L., Li, Q., Rohlin, L., Kim, U., Salmon, K., Rejtar, T., Gunsalus, R. P., Karger, B. L. and Ferry, J. G. (2007). Quantitative proteomic and microarray analysis of the archaeon Methanosarcina acetivorans grown with acetate versus methanol. J Proteome Res 6, 759-71.

Lidke, D. S. and Wilson, B. S. (2009). Caught in the act: quantifying protein behaviour in living cells. Trends Cell Biol 19, 566-74.

Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M. and Yates, J. R., 3rd. (1999). Direct analysis of protein complexes using mass spectrometry. Nat Biotechnol 17, 676-82.

Lippert, D. N., Ralph, S. G., Phillips, M., White, R., Smith, D., Hardie, D., Gershenzon, J., Ritland, K., Borchers, C. H. and Bohlmann, J. (2009). Quantitative iTRAQ proteome and comparative transcriptome analysis of elicitor-induced Norway spruce (Picea abies) cells reveals elements of calcium signaling in the early conifer defense response. Proteomics 9, 350-67.

Lodha, M., Marco, C. F. and Timmermans, M. C. (2008). Genetic and epigenetic regulation of stem cell homeostasis in plants. Cold Spring Harb Symp Quant Biol 73, 243-51.

Loewith, R., Jacinto, E., Wullschleger, S., Lorberg, A., Crespo, J. L., Bonenfant, D., Oppliger, W., Jenoe, P. and Hall, M. N. (2002). Two TOR complexes, only one of which is rapamycin sensitive, have distinct roles in cell growth control. Mol Cell 10, 457-68.

Lohrig, K., Muller, B., Davydova, J., Leister, D. and Wolters, D. A. (2009). Phosphorylation site mapping of soluble proteins: bioinformatical filtering reveals potential plastidic phosphoproteins in Arabidopsis thaliana. Planta 229, 1123-34.

Loidl, P. (2004). A plant dialect of the histone language. Trends Plant Sci 9, 84-90.

Lolas, I. B., Himanen, K., Gronlund, J. T., Lynggaard, C., Houben, A., Melzer, M., Van Lijsebettens, M. and Grasser, K. D. (2010). The transcript elongation factor FACT affects Arabidopsis vegetative and reproductive development and genetically interacts with HUB1/2. Plant J 61, 686-97.

Lolli, G. (2009). Binding to DNA of the RNA-polymerase II C-terminal domain allows discrimination between Cdk7 and Cdk9 phosphorylation. Nucleic Acids Res 37, 1260-8.

Lolli, G., Lowe, E. D., Brown, N. R. and Johnson, L. N. (2004). The crystal structure of human CDK7 and its protein recognition properties. Structure 12, 2067-79.

Lopez-Garcia, P. and Moreira, D. (2006). Selective forces for the origin of the eukaryotic nucleus. Bioessays 28, 525-33.

Lorkovic, Z. J. and Barta, A. (2004). Compartmentalization of the splicing machinery in plant cell nuclei. Trends Plant Sci 9, 565-8.

Lorkovic, Z. J., Hilscher, J. and Barta, A. (2008). Co-localisation studies of Arabidopsis SR splicing factors reveal different types of speckles in plant cell nuclei. Exp Cell Res 314, 3175-86.

Loyer, P., Trembley, J. H., Katona, R., Kidd, V. J. and Lahti, J. M. (2005). Role of CDK/cyclin complexes in transcription and RNA splicing. Cell Signal 17, 1033-51.

Loyer, P., Trembley, J. H., Lahti, J. M. and Kidd, V. J. (1998). The RNP protein, RNPS1, associates with specific isoforms of the p34cdc2-related PITSLRE protein kinase in vivo. J Cell Sci 111 (Pt 11), 1495-506.

Lu, H. and Schulze-Gahmen, U. (2006). Toward understanding the structural basis of cyclin-dependent kinase 6 specific inhibition. J Med Chem 49, 3826-31.

Lucker, J., Laszczak, M., Smith, D. and Lund, S. T. (2009). Generation of a predicted protein database from EST data and application to iTRAQ analyses in grape (Vitis vinifera cv. Cabernet Sauvignon) berries at ripening initiation. BMC Genomics 10, 50.

Lundergardh, H. (1913) Archiv. Zellforsch. , 9, 203.

Luo, D. and Oppenheimer, D. G. (1999). Genetic control of trichome branch number in Arabidopsis: the roles of the FURCA loci. Development 126, 5547-57.

Ma, X. M. and Blenis, J. (2009). Molecular mechanisms of mTOR-mediated translational control. Nat Rev Mol Cell Biol 10, 307-18.

Ma, X. M., Yoon, S. O., Richardson, C. J., Julich, K. and Blenis, J. (2008). SKAR links pre-mRNA splicing to mTOR/S6K1-mediated enhanced translation efficiency of spliced mRNAs. Cell 133, 303-13.

McKay, V. L., Li, X., Flory, M. R., Turcott, E., Law, G. L., Serikawa, K. A., Xu, X. L., Lee, H., Goodlett, D. R., Aebersold, R. et al. (2004). Gene expression analyzed by high-resolution state array analysis and quantitative proteomics: response of yeast to mating pheromone. Mol Cell Proteomics 3, 478-89.

Maere, S., Heymans, K. and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics 21, 3448-9.

Majeran, W., Zybailov, B., Ytterberg, A. J., Dunsmore, J., Sun, Q. and van Wijk, K. J. (2008). Consequences of C4 differentiation for chloroplast membrane proteomes in maize mesophyll and bundle sheath cells. Mol Cell Proteomics 7, 1609-38.

Makarov, A. (2000). Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. Anal Chem 72, 1156-62.

Makarov, A., Denisov, E., Lange, O. and Horning, S. (2006). Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer. J Am Soc Mass Spectrom 17, 977-82.

Maldonado, E., Shiekhattar, R., Sheldon, M., Cho, H., Drapkin, R., Rickert, P., Lees, E., Anderson, C. W., Linn, S. and Reinberg, D. (1996). A human RNA polymerase II complex associated with SRB and DNA-repair proteins. Nature 381, 86-9.

Manning, G., Whyte, D. B., Martinez, R., Hunter, T. and Sudarsanam, S. (2002). The protein kinase complement of the human genome. Science 298, 1912-34.

Maor, R., Jones, A., Nuhse, T. S., Studholme, D. J., Peck, S. C. and Shirasu, K. (2007). Multidimensional protein identification technology (MudPIT) analysis of ubiquitinated proteins in plants. Mol Cell Proteomics 6, 601-10.

Maquat, L. E. (2004). Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nat Rev Mol Cell Biol 5, 89-99.

Marion, J., Bach, L., Bellec, Y., Meyer, C., Gissot, L. and Faure, J. D. (2008). Systematic analysis of protein subcellular localization and interaction using high-throughput transient transformation of Arabidopsis seedlings. Plant J 56, 169-79.

Maris, C., Dominguez, C. and Allain, F. H. (2005). The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. FEBS J 272, 2118-31.

Marques, F., Moreau, J. L., Peaucellier, G., Lozano, J. C., Schatt, P., Picard, A., Callebaut, I., Perret, E. and Geneviere, A. M. (2000). A new subfamily of high molecular mass CDC2-related kinases with PITAI/VRE motifs. Biochem Biophys Res Commun 279, 832-7.

Marsh, E., Alvarez, S., Hicks, L. M., Barbazuk, W. B., Qiu, W., Kovacs, L. and Schachtman, D. (2010). Changes in protein abundance during powdery mildew infection of leaf tissues of Cabernet Sauvignon grapevine (Vitis vinifera L.). Proteomics 10, 2057-64.

Marshall, N.F., Peng, J., Xie, Z. and Price, D.H. (1996) Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. J Biol Chem, 271, 27176-27183.

Martin, B., Brenneman, R., Becker, K. G., Gucek, M., Cole, R. N. and Maudsley, S. (2008). iTRAQ analysis of complex proteome alterations in 3xTgAD Alzheimer's mice: understanding the interface between physiology and disease. PLoS One 3, e2750.

Mathur, J. a. K., C. (1998). Callus culture and regeneration in Arabidopsis thaliana. In Arabidopsis Protocols, (ed. J. a. S. Martinez-Zapater, J.), pp. 31-34: Humana Press.

Mathur, J., Szabados, L., Schaefer, S., Grunenberg, B., Lossow, A., Jonas-Straube, E., Schell, J., Koncz, C. and Koncz-Kalman, Z. (1998). Gene identification with sequenced T-DNA tags generated by transformation of Arabidopsis cell suspension. Plant J 13, 707-16.

Max, T., Sogaard, M. and Svejstrup, J.Q. (2007) Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. J Biol Chem, 282, 14113-14120.

May, M. a. L., C. (1993). Oxidative stimulation of glutathione synthesis in Arabidopsis thaliana suspension cultures. Plant Physiol 103, 621-627.

Mayeda, A., Badolato, J., Kobayashi, R., Zhang, M.Q., Gardiner, E.M. and Krainer, A.R. (1999) Purification and characterization of human RNPS1: a general activator of pre-mRNA splicing. Embo J, 18, 4560-4570.

McKendrick, L., Morley, S. J., Pain, V. M., Jagus, R. and Joshi, B. (2001). Phosphorylation of eukaryotic initiation factor 4E (eIF4E) at Ser209 is not required for protein synthesis in vitro and in vivo. Eur J Biochem 268, 5375-85.

Meaburn, K. J., Misteli, T. and Soutoglou, E. (2007). Spatial genome organization in the formation of chromosomal translocations. Semin Cancer Biol 17, 80-90.

Medina, F. J., Cerdido, A. and de Carcer, G. (2000). The functional organization of the nucleolus in proliferating plant cells. Eur J Histochem 44, 117-31.

Mehta, I. S., Amira, M., Harvey, A. J. and Bridger, J. M. (2010). Rapid chromosome territory relocation by nuclear motor activity in response to serum removal in primary human fibroblasts. Genome Biol 11, R5.

Meier, I. (2007). Composition of the plant nuclear envelope: theme and variations. J Exp Bot 58, 27-34.

Melaragno, J. E., Mehrotra, B. and Coleman, A. W. (1993). Relationship between Endopolyploidy and Cell Size in Epidermal Tissue of Arabidopsis. Plant Cell 5, 1661-1668.

Menand, B., Desnos, T., Nussaume, L., Berger, F., Bouchez, D., Meyer, C. and Robaglia, C. (2002). Expression and disruption of the Arabidopsis TOR (target of rapamycin) gene. Proc Natl Acad Sci U S A 99, 6422-7.

Mendenhall, M.D., Richardson, H.E. and Reed, S.I. (1988) Dominant negative protein kinase mutations that confer a G1 arrest phenotype. Proc Natl Acad Sci U S A, 85, 4426-4430.

Menges, M., de Jager, S. M., Gruissem, W. and Murray, J. A. (2005). Global analysis of the core cell cycle regulators of Arabidopsis identifies novel genes, reveals multiple and highly specific profiles of expression and provides a coherent model for plant cell cycle control. Plant J 41, 546-66.

Menges, M., Hennig, L., Gruissem, W. and Murray, J. A. (2003). Genome-wide gene expression in an Arabidopsis cell suspension. Plant Mol Biol 53, 423-42.

Merrick, M. J., Gibbins, J. R. and Postgate, J. R. (1987). A rapid and efficient method for plasmid transformation of Klebsiella pneumoniae and Escherichia coli. J Gen Microbiol 133, 2053-7.

Michalet, X., Pinaud, F. F., Bentolila, L. A., Tsay, J. M., Doose, S., Li, J. J., Sundaresan, G., Wu, A. M., Gambhir, S. S. and Weiss, S. (2005). Quantum dots for live cells, in vivo imaging, and diagnostics. Science 307, 538-44.

Minguez, P., Gotz, S., Montaner, D., Al-Shahrour, F. and Dopazo, J. (2009). SNOW, a web-based tool for the statistical analysis of protein-protein interaction networks. Nucleic Acids Res 37, W109-14.

Misteli, T. (2000). Cell biology of transcription and pre-mRNA splicing: nuclear architecture meets nuclear function. J Cell Sci 113 (Pt 11), 1841-9.

Misteli, T., Caceres, J. F. and Spector, D. L. (1997). The dynamics of a pre-mRNA splicing factor in living cells. Nature 387, 523-7.
mitotic cycle of proliferative and stationary phase exised root meristems. Exp H Cell Res 51, 167-176.

Mitulovic, G., Stingl, C., Steinmacher, I., Hudecz, O., Hutchins, J. R., Peters, J. M. and Mechtler, K. (2009). Preventing carryover of peptides and proteins in nano LC-MS separations. Anal Chem 81, 5955-60.

Moen, P. T., Jr., Johnson, C. V., Byron, M., Shopland, L. S., de la Serna, I. L., Imbalzano, A. N. and Lawrence, J. B. (2004). Repositioning of muscle-specific genes relative to the periphery of SC-35 domains during skeletal myogenesis. Mol Biol Cell 15, 197-206.

Moore, M. J. and Proudfoot, N. J. (2009). Pre-mRNA processing reaches back to transcription and ahead to translation. Cell 136, 688-700.

Moreno Díaz de la Espina, S., Barthellemy, I. and Cerezuela, M. A. (1991). Isolation and ultraestructural characterization of the residual nuclear matrix in a plant cell system. Chromosoma 100, 110-117.

Morgan, D. O. (1995). Principles of CDK regulation. Nature 374, 131-4.

Morgan, D. O. (1997). Cyclin-dependent kinases: engines, clocks, and microprocessors. Annu Rev Cell Dev Biol 13, 261-91.

Morris, E.J., Ji, J.Y., Yang, F., Di Stefano, L., Herr, A., Moon, N.S., Kwon, E.J., Haigis, K.M., Naar, A.M. and Dyson, N.J. (2008) E2F1 represses beta-catenin transcription and is antagonized by both pRB and CDK8. Nature, 455, 552-556.

Myers, L. C. and Kornberg, R. D. (2000). Mediator of transcriptional regulation. Annu Rev Biochem 69, 729-49.

Myers, L.C. and Kornberg, R.D. (2000) Mediator of transcriptional regulation. Annu Rev Biochem, 69, 729-749.

Navarro, L., Dunoyer, P., Jay, F., Arnold, B., Dharmasiri, N., Estelle, M., Voinnet, O. and Jones, J. D. (2006). A plant miRNA contributes to antibacterial resistance by repressing auxin signaling. Science 312, 436-9.

Nebreda, A. R. (2006). CDK activation by non-cyclin proteins. Curr Opin Cell Biol 18, 192-8.

Nesic, D., Tanackovic, G. and Kramer, A. (2004). A role for Cajal bodies in the final steps of U2 snRNP biogenesis. J Cell Sci 117, 4423-33.

Nesvizhskii, A.I., Keller, A., Kolker, E. and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem, 75, 4646-4658.

Neumann, F. R. and Nurse, P. (2007). Nuclear size control in fission yeast. J Cell Biol 179, 593-600.

Nicolai, M., Roncato, M. A., Canoy, A. S., Rouquie, D., Sarda, X., Freyssinet, G. and Robaglia, C. (2006). Large-scale analysis of mRNA translation states during sucrose starvation in arabidopsis cells identifies cell proliferation and chromatin structure as targets of translational control. Plant Physiol 141, 663-73.

Nie, L., Wu, G., Culley, D. E., Scholten, J. C. and Zhang, W. (2007). Integrative analysis of transcriptomic and proteomic data: challenges, solutions and applications. Crit Rev Biotechnol 27, 63-75.

Nigg, E. A. (1995). Cyclin-dependent protein kinases: key regulators of the eukaryotic cell cycle. Bioessays 17, 471-80.

Nilo, R., Saffie, C., Lilley, K., Baeza-Yates, R., Cambiazo, V., Campos-Vargas, R., Gonzalez, M., Meisel, L. A., Retamales, J., Silva, H. et al. (2010). Proteomic analysis of peach fruit mesocarp softening and chilling injury using difference gel electrophoresis (DIGE). BMC Genomics 11, 43.

Nowak, S.J. and Corces, V.G. (2000) Phosphorylation of histone H3 correlates with transcriptionally active loci. Genes Dev, 14, 3003-3013.

Nuhse, T. S., Bottrill, A. R., Jones, A. M. and Peck, S. C. (2007). Quantitative phosphoproteomic analysis of plasma membrane proteins reveals regulatory mechanisms of plant innate immune responses. Plant J 51, 931-40.

Nurse, P. (2002). Cyclin dependent kinases and cell cycle control (nobel lecture). Chembiochem 3, 596-603.

Nurse, P. and Bissett, Y. (1981). Gene required in G1 for commitment to cell cycle and in G2 for control of mitosis in fission yeast. Nature 292, 558-60.

Oelgeschlager, T. (2002) Regulation of RNA polymerase II activity by CTD phosphorylation and cell cycle control. J Cell Physiol, 190, 160-169.

Old, W. M., Meyer-Arendt, K., Aveline-Wolf, L., Pierce, K. G., Mendoza, A., Sevinsky, J. R., Resing, K. A. and Ahn, N. G. (2005). Comparison of label-free methods for quantifying human proteins by shotgun proteomics. Mol Cell Proteomics 4, 1487-502.

Olins, A. L., Zwerger, M., Herrmann, H., Zentgraf, H., Simon, A. J., Monestier, M. and Olins, D. E. (2008). The human granulocyte nucleus: Unusual nuclear envelope and heterochromatin composition. Eur J Cell Biol 87, 279-90.

Olson, M.O., Dundr, M. and Szebeni, A. (2000) The nucleolus: an old factory with unexpected capabilities. Trends Cell Biol, 10, 189-196.

Ong, S. E. and Mann, M. (2005). Mass spectrometry-based proteomics turns quantitative. Nat Chem Biol 1, 252-62.

Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A. and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1, 376-86.

Ortega-Martinez, O., Pernas, M., Carol, R. J. and Dolan, L. (2007). Ethylene modulates stem cell division in the Arabidopsis thaliana root. Science 317, 507-10.

Osborne, C. S., Chakalova, L., Brown, K. E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J. A., Lopes, S., Reik, W. et al. (2004). Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet 36, 1065-71.

Pandey, A., Choudhary, M. K., Bhushan, D., Chattopadhyay, A., Chakraborty, S., Datta, A. and Chakraborty, N. (2006). The nuclear proteome of chickpea (Cicer arietinum L.) reveals predicted and unexpected proteins. J Proteome Res 5, 3301-11.

Patterson, J., Ford, K., Cassin, A., Natera, S. and Bacic, A. (2007). Increased abundance of proteins involved in phytosiderophore production in boron-tolerant barley. Plant Physiol 144, 1612-31.

Pei, Y. and Shuman, S. (2002). Interactions between fission yeast mRNA capping enzymes and elongation factor Spt5. J Biol Chem 277, 19639-48.

Pei, Y., Du, H., Singer, J., Stamour, C., Granitto, S., Shuman, S. and Fisher, R. P. (2006). Cyclin-dependent kinase 9 (Cdk9) of fission yeast is activated by the CDK-activating kinase Csk1, overlaps functionally with the TFIIH-associated kinase Mcs6, and associates with the mRNA cap methyltransferase Pcm1 in vivo. Mol Cell Biol 26, 777-88.

Pei, Y., Schwer, B. and Shuman, S. (2003). Interactions between fission yeast Cdk9, its cyclin partner Pch1, and mRNA capping enzyme Pct1 suggest an elongation checkpoint for mRNA quality control. J Biol Chem 278, 7180-8.

Pendle, A. F., Clark, G. P., Boon, R., Lewandowska, D., Lam, Y. W., Andersen, J., Mann, M., Lamond, A. I., Brown, J. W. and Shaw, P. J. (2005). Proteomic analysis of the Arabidopsis nucleolus suggests novel nucleolar functions. Mol Biol Cell 16, 260-9.

Peng, J., Zhu, Y., Milton, J.T. and Price, D.H. (1998) Identification of multiple cyclin subunits of human P-TEFb. Genes Dev, 12, 755-762.

Perkins, D. N., Pappin, D. J., Creasy, D. M. and Cottrell, J. S. (1999). Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20, 3551-67.

Perry, R. P. (2007). Balanced production of ribosomal proteins. Gene 401, 1-3.

Pestova, T. V., Kolupaeva, V. G., Lomakin, I. B., Pilipenko, E. V., Shatsky, I. N., Agol, V. I. and Hellen, C. U. (2001). Molecular mechanisms of translation initiation in eukaryotes. Proc Natl Acad Sci U S A 98, 7029-36.

Petersen-Mahrt, S. K., Estmer, C., Ohrmalm, C., Matthews, D. A., Russell, W. C. and Akusjarvi, G. (1999). The splicing factor-associated protein, p32, regulates RNA splicing by inhibiting ASF/SF2 RNA binding and phosphorylation. EMBO J 18, 1014-24.

Petricka, J. J. and Nelson, T. M. (2007). Arabidopsis nucleolin affects plant development and patterning. Plant Physiol 144, 173-86.

Phair, R. D. and Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. Nature 404, 604-9.

Pierrat, O. A., Mikitova, V., Bush, M. S., Browning, K. S. and Doonan, J. H. (2007). Control of protein translation by phosphorylation of the mRNA 5'-cap-binding complex. Biochem Soc Trans 35, 1634-7.
Pinhero, R., Liaw, P., Bertens, K. and Yankulov, K. (2004) Three cyclin-dependent kinases preferentially phosphorylate different parts of the C-terminal domain of the large subunit of RNA polymerase II. Eur J Biochem, 271, 1004-1014.

Piques, M., Schulze, W. X., Hohne, M., Usadel, B., Gibon, Y., Rohwer, J. and Stitt, M. (2009). Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in Arabidopsis. Mol Syst Biol 5, 314.

Platani, M., Goldberg, I., Swedlow, J. R. and Lamond, A. I. (2000). In vivo analysis of Cajal body movement, separation, and joining in live human cells. J Cell Biol 151, 1561-74.

Powell, L. (1988). In Plant Hormones and their role in plant growth and development. 539-552.
Powers, T. and Walter, P. (1999). Regulation of ribosome biogenesis by the rapamycin-sensitive TOR-signaling pathway in Saccharomyces cerevisiae. Mol Biol Cell 10, 987-1000.

Prabhakar, V., Lottgert, T., Gigolashvili, T., Bell, K., Flugge, U. I. and Hausler, R. E. (2009). Molecular and functional characterization of the plastid-localized Phosphoenolpyruvate enolase (ENO1) from Arabidopsis thaliana. FEBS Lett 583, 983-91.

Prakash, A., Piening, B., Whiteaker, J., Zhang, H., Shaffer, S. A., Martin, D., Hohmann, L., Cooke, K., Olson, J. M., Hansen, S. et al. (2007). Assessing bias in experiment design for large scale mass spectrometry-based quantitative proteomics. Mol Cell Proteomics 6, 1741-8.

Price, D.H. (2000) P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. Mol Cell Biol, 20, 2629-2634.

Proud, C. G. (2002). Regulation of mammalian translation factors by nutrients. Eur J Biochem 269, 5338-49.

Proudfoot, N. J., Furger, A. and Dye, M. J. (2002). Integrating mRNA processing with transcription. Cell 108, 501-12.

Puvion-Dutilleul, F., Besse, S., Diaz, J. J., Kindbeiter, K., Vigneron, M., Warren, S. L., Kedinger, C., Madjar, J. J. and Puvion, E. (1997). Identification of transcription factories in nuclei of HeLa cells transiently expressing the Us11 gene of herpes simplex virus type 1. Gene Expr 6, 315-32.

Ragoczy, T., Bender, M. A., Telling, A., Byron, R. and Groudine, M. (2006). The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. Genes Dev 20, 1447-57.

Ranish, J.A. and Hahn, S. (1996) Transcription: basal factors and activation. Curr Opin Genet Dev, 6, 151-158.

Rappsilber, J., Ryder, U., Lamond, A. I. and Mann, M. (2002). Large-scale proteomic analysis of the human spliceosome. Genome Res 12, 1231-45.

Raska, I., Shaw, P. J. and Cmarko, D. (2006). New insights into nucleolar architecture and activity. Int Rev Cytol 255, 177-235.

Raught, B., Gingras, A. C., Gygi, S. P., Imataka, H., Morino, S., Gradi, A., Aebersold, R. and Sonenberg, N. (2000). Serum-stimulated, rapamycin-sensitive phosphorylation sites in the eukaryotic translation initiation factor 4GI. EMBO J 19, 434-44.

Rechsteiner, M. and Rogers, S. W. (1996). PEST sequences and regulation by proteolysis. Trends Biochem Sci 21, 267-71.

Reeves, R. (2001). Molecular biology of HMGA proteins: hubs of nuclear function. Gene 277, 63-81.

Renaudin, J. P., Doonan, J. H., Freeman, D., Hashimoto, J., Hirt, H., Inze, D., Jacobs, T., Kouchi, H., Rouze, P., Sauter, M. et al. (1996). Plant cyclins: a unified nomenclature for plant A-, B- and D-type cyclins based on sequence organization. Plant Mol Biol 32, 1003-18.

Repetto, O., Rogniaux, H., Firnhaber, C., Zuber, H., Kuster, H., Larre, C., Thompson, R. and Gallardo, K. (2008). Exploring the nuclear proteome of Medicago truncatula at the switch towards seed filling. Plant J 56, 398-410.

Richard, C., Granier, C., Inze, D. and De Veylder, L. (2001). Analysis of cell division parameters and cell cycle gene expression during the cultivation of Arabidopsis thaliana cell suspensions. J Exp Bot 52, 1625-33.

Rino, J., Carvalho, T., Braga, J., Desterro, J. M., Luhrmann, R. and Carmo-Fonseca, M. (2007). A stochastic view of spliceosome assembly and recycling in the nucleus. PLoS Comput Biol 3, 2019-31.

Riou-Khamlichi, C., Menges, M., Healy, J. M. and Murray, J. A. (2000). Sugar control of the plant cell cycle: differential regulation of Arabidopsis D-type cyclin gene expression. Mol Cell Biol 20, 4513-21.

Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., Khainovski, N., Pillai, S., Dey, S., Daniels, S. et al. (2004). Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Mol Cell Proteomics 3, 1154-69.

Rossignol, M., Kolb-Cheynel, I. and Egly, J.M. (1997) Substrate specificity of the cdk-activating kinase (CAK) is altered upon association with TFIIH. EMBO J, 16, 1628-1637.

Rusinol, A. E. and Sinensky, M. S. (2006). Farnesylated lamins, progeroid syndromes and farnesyl transferase inhibitors. J Cell Sci 119, 3265-72.

Sacco-Bubulya, P. and Spector, D. L. (2002). Disassembly of interchromatin granule clusters alters the coordination of transcription and pre-mRNA splicing. J Cell Biol 156, 425-36.

Saitoh, N., Spahr, C. S., Patterson, S. D., Bubulya, P., Neuwald, A. F. and Spector, D. L. (2004). Proteomic analysis of interchromatin granule clusters. Mol Biol Cell 15, 3876-90.

Sanford, J. R. and Bruzik, J. P. (1999). Developmental regulation of SR protein phosphorylation and activity. Genes Dev 13, 1513-8.

Sanford, J. R., Ellis, J. D., Cazalla, D. and Caceres, J. F. (2005). Reversible phosphorylation differentially affects nuclear and cytoplasmic functions of splicing factor 2/alternative splicing factor. Proc Natl Acad Sci U S A 102, 15042-7.

Sansam, C. L., Wells, K. S. and Emeson, R. B. (2003). Modulation of RNA editing by functional nucleolar sequestration of ADAR2. Proc Natl Acad Sci U S A 100, 14018-23.

Schafer, T., Strauss, D., Petfalski, E., Tollervey, D. and Hurt, E. (2003). The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes. EMBO J 22, 1370-80.

Schaff, J. E., Mbeunkui, F., Blackburn, K., Bird, D. M. and Goshe, M. B. (2008). SILIP: a novel stable isotope labeling method for in planta quantitative proteomic analysis. Plant J 56, 840-54.

Schaffert, N., Hossbach, M., Heintzmann, R., Achsel, T. and Luhrmann, R. (2004). RNAi knockdown of hPrp31 leads to an accumulation of U4/U6 di-snRNPs in Cajal bodies. EMBO J 23, 3000-9.

Scheer, U. and Rose, K. M. (1984). Localization of RNA polymerase I in interphase cells and mitotic chromosomes by light and electron microscopic immunocytochemistry. Proc Natl Acad Sci U S A 81, 1431-5.

Schirmer, E. C., Florens, L., Guan, T., Yates, J. R., 3rd and Gerace, L. (2005). Identification of novel integral membrane proteins of the nuclear envelope with potential disease links using subtractive proteomics. Novartis Found Symp 264, 63-76; discussion 76-80, 227-30.

Schmidt, E. E. and Schibler, U. (1995). Cell size regulation, a mechanism that controls cellular RNA accumulation: consequences on regulation of the ubiquitous transcription factors Oct1 and NF-Y and the liver-enriched transcription factor DBP. J Cell Biol 128, 467-83.

Schneider, T., Schellenberg, M., Meyer, S., Keller, F., Gehrig, P., Riedel, K., Lee, Y., Eberl, L. and Martinoia, E. (2009). Quantitative detection of changes in the leaf-mesophyll tonoplast proteome in dependency of a cadmium exposure of barley (Hordeum vulgare L.) plants. Proteomics 9, 2668-77.

Schulze, W. X. and Usadel, B. (2010). Quantitation in mass-spectrometry-based proteomics. Annu Rev Plant Biol 61, 491-516.

Schwartz, R., Ting, C. S. and King, J. (2001). Whole proteome pI values correlate with subcellular localizations of proteins for organisms within the three domains of life. Genome Res 11, 703-9.

Serizawa, H., Makela, T.P., Conaway, J.W., Conaway, R.C., Weinberg, R.A. and Young, R.A. (1995) Association of Cdk-activating kinase subunits with transcription factor TFIIH. Nature, 374, 280-282.

Sgorbati, S. S., E.; Levi, M.; Chiatante, D. and Giordano, P. . (1988). Bivariate cytofluorimetric analysis of DNA and nuclear protein content in plant tissue. Protoplasma 144, 180-184.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498-504.

Shav-Tal, Y., Blechman, J., Darzacq, X., Montagna, C., Dye, B.T., Patton, J.G., Singer, R.H. and Zipori, D. (2005) Dynamic sorting of nuclear components into distinct nucleolar caps during transcriptional inhibition. Mol Biol Cell, 16, 2395-2413.

Shaw, P. and Doonan, J. (2005). The nucleolus. Playing by different rules? Cell Cycle 4, 102-5.

Shaw, P. J. and Brown, J. W. (2004). Plant nuclear bodies. Curr Opin Plant Biol 7, 614-20.

Shaw, P. J. and Jordan, E. G. (1995). The nucleolus. Annu Rev Cell Dev Biol 11, 93-121.

Shemer, R., Meimoun, A., Holtzman, T. and Kornitzer, D. (2002). Regulation of the transcription factor Gcn4 by Pho85 cyclin PCL5. Mol Cell Biol 22, 5395-404.

Shiekhattar, R., Mermelstein, F., Fisher, R.P., Drapkin, R., Dynlacht, B., Wessling, H.C., Morgan, D.O. and Reinberg, D. (1995) Cdk-activating kinase complex is a component of human transcription factor TFIIH. Nature, 374, 283-287.

Shimanuki, M., Chung, S. Y., Chikashige, Y., Kawasaki, Y., Uehara, L., Tsutsumi, C., Hatanaka, M., Hiraoka, Y., Nagao, K. and Yanagida, M. (2007). Two-step, extensive alterations in the transcriptome from G0 arrest to cell division in Schizosaccharomyces pombe. Genes Cells 12, 677-92.

Shimizu-Sato, S. and Mori, H. (2001). Control of outgrowth and dormancy in axillary buds. Plant Physiol 127, 1405-13.

Shimizu, S. and Mori, H. (1998). Analysis of cycles of dormancy and growth in pea axillary buds based on mRNA accumulation patterns of cell cycle-related genes. Plant Cell Physiol 39, 255-62.

Shimotohno, A., Matsubayashi, S., Yamaguchi, M., Uchimiya, H. and Umeda, M. (2003) Differential phosphorylation activities of CDK-activating kinases in Arabidopsis thaliana. FEBS Lett, 534, 69-74.

Shimotohno, A., Ohno, R., Bisova, K., Sakaguchi, N., Huang, J., Koncz, C., Uchimiya, H. and Umeda, M. (2006) Diverse phosphoregulatory mechanisms controlling cyclin-dependent kinase-activating kinases in Arabidopsis. Plant J, 47, 701-710.

Shimotohno, A., Umeda-Hara, C., Bisova, K., Uchimiya, H. and Umeda, M. (2004) The plant-specific kinase CDKF;1 is involved in activating phosphorylation of cyclin-dependent kinase-activating kinases in Arabidopsis. Plant Cell, 16, 2954-2966.

Shopland, L. S., Johnson, C. V., Byron, M., McNeil, J. and Lawrence, J. B. (2003). Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. J Cell Biol 162, 981-90.

Sigmon, I., Lee, L. W., Chang, D. K., Krusberski, N., Cohen, D., Eng, J. K. and Martin, D. B. (2010) ChromEval: a software application for the rapid evaluation of HPLC system performance in proteomic applications. Anal Chem 82, 5060-8.

Skop, A. R., Liu, H., Yates, J., 3rd, Meyer, B. J. and Heald, R. (2004). Dissection of the mammalian midbody proteome reveals conserved cytokinesis mechanisms. Science 305, 61-6.

Sormani, R., Yao, L., Menand, B., Ennar, N., Lecampion, C., Meyer, C. and Robaglia, C. (2007). Saccharomyces cerevisiae FKBP12 binds Arabidopsis thaliana TOR and its expression in plants leads to rapamycin susceptibility. BMC Plant Biol 7, 26.

Sridha, S. and Wu, K. (2006). Identification of AtHD2C as a novel regulator of abscisic acid responses in Arabidopsis. Plant J 46, 124-33.

Sridhar, V. V., Surendrarao, A. and Liu, Z. (2006). APETALA1 and SEPALLATA3 interact with SEUSS to mediate transcription repression during flower development. Development 133, 3159-66.

Sridhar, V. V., Surendrarao, A., Gonzalez, D., Conlan, R. S. and Liu, Z. (2004). Transcriptional repression of target genes by LEUNIG and SEUSS, two interacting regulatory proteins for Arabidopsis flower development. Proc Natl Acad Sci U S A 101, 11494-9.

Stamm, S. (2008). Regulation of alternative splicing by reversible protein phosphorylation. J Biol Chem 283, 1223-7.

Strasburger, E. (1880) Zellbindung und Zelltheilung: Gustav Fisher, Jene.

Strasburger, E. (1884)Arch. Mikroskop. Anat, 23, 246

Strezoska, Z., Pestov, D. G. and Lau, L. F. (2002). Functional inactivation of the mouse nucleolar protein Bop1 inhibits multiple steps in pre-rRNA processing and blocks cell cycle progression. J Biol Chem 277, 29617-25.

Strudwick, S. and Borden, K. L. (2002). The emerging roles of translation factor eIF4E in the nucleus. Differentiation 70, 10-22.

Su, S. S., Tanaka, Y., Samejima, I., Tanaka, K. and Yanagida, M. (1996). A nitrogen starvation-induced dormant G0 state in fission yeast: the establishment from uncommitted G1 state and its delay for return to proliferation. J Cell Sci 109 (Pt 6), 1347-57.

Suelter, C. H. and DeLuca, M. (1983). How to prevent losses of protein by adsorption to glass and plastic. Anal Biochem 135, 112-9.

Sugimoto-Shirasu, K. and Roberts, K. (2003). "Big it up": endoreduplication and cell-size control in plants. Curr Opin Plant Biol 6, 544-53.

Sugimoto-Shirasu, K., Roberts, G. R., Stacey, N. J., McCann, M. C., Maxwell, A. and Roberts, K. (2005). RHL1 is an essential component of the plant DNA topoisomerase VI complex and is required for ploidy-dependent cell growth. Proc Natl Acad Sci U S A 102, 18736-41.

Sutherland HGE, M. G., Newton K, Ford LV, Farral R, Dellaire G, Cáceres JF and Bickmore WA. (2001). Large-scale identification of mammalian proteins localized to nuclear sub-compartments. Hum Mol Genet 10, 1995-2011.

Swift, H. (1950). The constancy of desoxyribose nucleic acid in plant nuclei. Proc Natl Acad Sci U S A 36, 643-54.

Taddei, A., Hediger, F., Neumann, F. R. and Gasser, S. M. (2004). The function of nuclear architecture: a genetic approach. Annu Rev Genet 38, 305-45.

Takahashi, Y., Rayman, J. B. and Dynlacht, B. D. (2000). Analysis of promoter binding by the E2F and pRB families in vivo: distinct E2F proteins mediate activation and repression. Genes Dev 14, 804-16.

Tan, F. L., G.; Chitteti, B. and Peng, Z. . (2007). Proteome and phosphoproteome analysis of chromatin associated proteins in rice (Oryza sativa). Proteomics 7, 4511-4527.

Thomas, C. H., Collier, J. H., Sfeir, C. S. and Healy, K. E. (2002). Engineering gene expression and protein synthesis by modulation of nuclear shape. Proc Natl Acad Sci U S A 99, 1972-7.

Thomas, G. (2000). An encore for ribosome biogenesis in the control of cell proliferation. Nat Cell Biol 2, E71-2.

Tillemans, V., Dispa, L., Remacle, C., Collinge, M. and Motte, P. (2005). Functional distribution and dynamics of Arabidopsis SR splicing factors in living plant cells. Plant J 41, 567-82.

Tillemans, V., Leponce, I., Rausin, G., Dispa, L. and Motte, P. (2006). Insights into nuclear organization in plants as revealed by the dynamic distribution of Arabidopsis SR splicing factors. Plant Cell 18, 3218-34.

Tobin H, S. T., Gordon J,. (1979). Electrophoretic transfer of proteins from polyacrylamide to nitrocellulose sheet: procedure and some applications. Proc Natl Acad Sci U S A 76, 4350-4.

Trembley, J.H., Hu, D., Slaughter, C.A., Lahti, J.M. and Kidd, V.J. (2003) Casein kinase 2 interacts with cyclin-dependent kinase 11 (CDK11) in vivo and phosphorylates both the RNA polymerase II carboxyl-terminal domain and CDK11 in vitro. J Biol Chem, 278, 2265-2270.

Trimarchi, J. M. and Lees, J. A. (2002). Sibling rivalry in the E2F family. Nat Rev Mol Cell Biol 3, 11-20.

Trotta, C. R., Lund, E., Kahan, L., Johnson, A. W. and Dahlberg, J. E. (2003). Coordinated nuclear export of 60S ribosomal subunits and NMD3 in vertebrates. EMBO J 22, 2841-51.

Tschochner, H. and Hurt, E. (2003). Pre-ribosomes on the road from the nucleolus to the cytoplasm. Trends Cell Biol 13, 255-63.

Tsutsui, T., Umemura, H., Tanaka, A., Mizuki, F., Hirose, Y. and Ohkuma, Y. (2008). Human mediator kinase subunit CDK11 plays a negative role in viral activator VP16-dependent transcriptional regulation. Genes Cells 13, 817-26.

Tzafrir, I., McElver, J. A., Liu Cm, C. M., Yang, L. J., Wu, J. Q., Martinez, A., Patton, D. A. and Meinke, D. W. (2002). Diversity of TITAN functions in Arabidopsis seed development. Plant Physiol 128, 38-51.

Ubersax, J., Woodbury, E., Quang, P., Paraz, M., Blethrow, J., Shah, K., Shokat, K., and Morgan, D. (2003) Targets of the cyclin-dependent kinase Cdk1. Nature 425, 859-64

Ueno, Y., Ishikawa, T., Watanabe, K., Terakura, S., Iwakawa, H., Okada, K., Machida, C. and Machida, Y. (2007). Histone deacetylases and ASYMMETRIC LEAVES2 are involved in the establishment of polarity in leaves of Arabidopsis. Plant Cell 19, 445-57.

Umen, J. G. (2005). The elusive sizer. Curr Opin Cell Biol 17, 435-41.

Unlu, M., Morgan, M. E. and Minden, J. S. (1997). Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. Electrophoresis 18, 2071-7.

Usheva, A., Maldonado, E., Goldring, A., Lu, H., Houbavi, C., Reinberg, D. and Aloni, Y. (1992). Specific interaction between the nonphosphorylated form of RNA polymerase II and the TATA-binding protein. Cell 69, 871-81.

van den Heuvel, S. and Harlow, E. (1993) Distinct roles for cyclin-dependent kinases in cell cycle control. Science, 262, 2050-2054.

Van't Hoff, J. (1968). The action of IAA and kinetin on the

Vandepoele, K., Raes, J., De Veylder, L., Rouze, P., Rombauts, S. and Inze, D. (2002). Genome-wide analysis of core cell cycle genes in Arabidopsis. Plant Cell 14, 903-16.

Vermeulen, L., Vanden Berghe, W., Beck, I. M., De Bosscher, K. and Haegeman, G. (2009). The versatile role of MSKs in transcriptional regulation. Trends Biochem Sci 34, 311-8.

Wada, T., Takagi, T., Yamaguchi, Y., Ferdous, A., Imai, T., Hirose, S., Sugimoto, S., Yano, K., Hartzog, G. A., Winston, F. et al. (1998). DSIF, a novel transcription elongation factor that regulates RNA polymerase II processivity, is composed of human Spt4 and Spt5 homologs. Genes Dev 12, 343-56.

Wada, T., Takagi, T., Yamaguchi, Y., Watanabe, D. and Handa, H. (1998) Evidence that P-TEFb alleviates the negative effect of DSIF on RNA polymerase II-dependent transcription in vitro. Embo J, 17, 7395-7403.

Wang, A., Kurdistani, S. K. and Grunstein, M. (2002). Requirement of Hos2 histone deacetylase for gene activity in yeast. Science 298, 1412-4.

Wang, B. a. (2006). Additives for reversed-phase HPLC mobile phases, vol. US Patent 7,125,492 B2.

Wang, G., Kong, H., Sun, Y., Zhang, X., Zhang, W., Altman, N., DePamphilis, C. W. and Ma, H. (2004). Genome-wide analysis of the cyclin family in Arabidopsis and comparative phylogenetic analysis of plant cyclin-like proteins. Plant Physiol 135, 1084-99.

Wang, H. and Hanash, S. M. (2008). Increased throughput and reduced carryover of mass spectrometry-based proteomics using a high-efficiency nonsplit nanoflow parallel dual-column capillary HPLC system. J Proteome Res 7, 2743-55.

Wang, W. and Chen, X. (2004). HUA ENHANCER3 reveals a role for a cyclin-dependent protein kinase in the specification of floral organ identity in Arabidopsis. Development 131, 3147-56.

Warner, A. K. and Sloboda, R. D. (1999). C-terminal domain of the mitotic apparatus protein p62 targets the protein to the nucleolus during interphase. Cell Motil Cytoskeleton 44, 68-80.

Washburn, M. P., Koller, A., Oshiro, G., Ulaszek, R. R., Plouffe, D., Deciu, C., Winzeler, E. and Yates, J. R., 3rd. (2003). Protein pathway and complex clustering of correlated mRNA and protein expression analyses in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 100, 3107-12.

Weart, R. B., Lee, A. H., Chien, A. C., Haeusser, D. P., Hill, N. S. and Levin, P. A. (2007). A metabolic sensor governing cell size in bacteria. Cell 130, 335-47.

Wei, P., Garber, M.E., Fang, S.M., Fischer, W.H. and Jones, K.A. (1998) A novel CDK9-associated C-type cyclin interacts directly with HIV-1 Tat and mediates its high-affinity, loop-specific binding to TAR RNA. Cell, 92, 451-462.

Weingartner, M., Binarova, P., Drykova, D., Schweighofer, A., David, J. P., Heberle-Bors, E., Doonan, J. and Bogre, L. (2001). Dynamic recruitment of Cdc2 to specific microtubule structures during mitosis. Plant Cell 13, 1929-43.

White, S. L., Gharbi, S., Bertani, M. F., Chan, H. L., Waterfield, M. D. and Timms, J. F. (2004). Cellular responses to ErbB-2 overexpression in human mammary luminal epithelial cells: comparison of mRNA and protein expression. Br J Cancer 90, 173-81.

Wilkins, M. R., Gasteiger, E., Gooley, A. A., Herbert, B. R., Molloy, M. P., Binz, P. A., Ou, K., Sanchez, J. C., Bairoch, A., Williams, K. L. et al. (1999). High-throughput mass spectrometric discovery of protein post-translational modifications. J Mol Biol 289, 645-57.

Wilson, E. (1925). The Cell in Development and Heredity. New York: Macmillan.

Won, K. A. and Reed, S. I. (1996). Activation of cyclin E/CDK2 is coupled to site-specific autophosphorylation and ubiquitin-dependent degradation of cyclin E. EMBO J 15, 4182-93.

Wood, A., Schneider, J., Dover, J., Johnston, M. and Shilatifard, A. (2005) The Bur1/Bur2 complex is required for histone H2B monoubiquitination by Rad6/Bre1 and histone methylation by COMPASS. Mol Cell, 20, 589-599.

Wyce, A., Xiao, T., Whelan, K.A., Kosman, C., Walter, W., Eick, D., Hughes, T.R., Krogan, N.J., Strahl, B.D. and Berger, S.L. (2007) H2B ubiquitylation acts as a barrier to Ctk1 nucleosomal recruitment prior to removal by Ubp8 within a SAGA-related complex. Mol Cell, 27, 275-288.

Xia, Q., Hendrickson, E. L., Zhang, Y., Wang, T., Taub, F., Moore, B. C., Porat, I., Whitman, W. B., Hackett, M. and Leigh, J. A. (2006). Quantitative proteomics of the archaeon Methanococcus maripaludis validated by microarray analysis and real time PCR. Mol Cell Proteomics 5, 868-81.

Xiang, J., Lahti, J. M., Grenet, J., Easton, J. and Kidd, V. J. (1994). Molecular cloning and expression of alternatively spliced PITSLRE protein kinase isoforms. J Biol Chem 269, 15786-94.

Xu, L. and Massague, J. (2004). Nucleocytoplasmic shuttling of signal transducers. Nat Rev Mol Cell Biol 5, 209-19.

Yamaguchi, M., Fabian, T., Sauter, M., Bhalerao, R.P., Schrader, J., Sandberg, G., Umeda, M. and Uchimiya, H. (2000) Activation of CDK-activating kinase is dependent on interaction with H-type cyclins in plants. Plant J, 24, 11-20.

Yamaguchi, M., Umeda, M. and Uchimiya, H. (1998). A rice homolog of Cdk7/MO15 phosphorylates both cyclin-dependent protein kinases and the carboxy-terminal domain of RNA polymerase II. Plant J 16, 613-9.

Yanagida, M. (2009). Cellular quiescence: are controlling genes conserved? Trends Cell Biol 19, 705-15.

Yankulov, K. Y. and Bentley, D. L. (1997). Regulation of CDK7 substrate specificity by MAT1 and TFIIH. EMBO J 16, 1638-46.

Yankulov, K.Y. and Bentley, D.L. (1997) Regulation of CDK7 substrate specificity by MAT1 and TFIIH. EMBO J, 16, 1638-1646.

Yates, J. R., Ruse, C. I. and Nakorchevsky, A. (2009). Proteomics by mass spectrometry: approaches, advances, and applications. Annu Rev Biomed Eng 11, 49-79.

Zandomeni, R., Zandomeni, M.C., Shugar, D. and Weinmann, R. (1986) Casein kinase type II is involved in the inhibition by 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole of specific RNA polymerase II transcription. J Biol Chem, 261, 3414-3419.

Zhang, X., Grey, P. H., Krishnakumar, S. and Oppenheimer, D. G. (2005). The IRREGULAR TRICHOME BRANCH loci regulate trichome elongation in Arabidopsis. Plant Cell Physiol 46, 1549-60.

Zhao, R., Bodnar, M. S. and Spector, D. L. (2009). Nuclear neighborhoods and gene expression. Curr Opin Genet Dev 19, 172-9.

Zhou, Z., Licklider, L. J., Gygi, S. P. and Reed, R. (2002). Comprehensive proteomic analysis of the human spliceosome. Nature 419, 182-5.

Zhu, M., Dai, S., McClung, S., Yan, X. and Chen, S. (2009). Functional differentiation of Brassica napus guard cells and mesophyll cells revealed by comparative proteomics. Mol Cell Proteomics 8, 752-66.

Zhu, Y., Dong, A., Meyer, D., Pichon, O., Renou, J. P., Cao, K. and Shen, W. H. (2006). Arabidopsis NRP1 and NRP2 encode histone chaperones and are required for maintaining postembryonic root growth. Plant Cell 18, 2879-92.

Zilberman, D., Cao, X. and Jacobsen, S. E. (2003). ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science 299, 716-9.

Zimmermann, B. and Adhami, H. (1976). Influence of cell density of nuclear size in monolayer cells from embryonic mouse brains. Acta Anat (Basel) 95, 399-407.

Zink, D., Fischer, A. H. and Nickerson, J. A. (2004). Nuclear structure in cancer cells. Nat Rev Cancer 4, 677-87.