UNIVERSITY
of
GLASGOW

Structural and functional studies of the DSC1 cell cycle transcription factor complex in fission yeast

A thesis submitted for the degree of
Doctor of Philosophy

Allan John Dunlop B.Sc. (Hons.)

Division of Biochemistry and Molecular Biology
Institute of Biomedical and Life Sciences
University of Glasgow
October 2004

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

Declaration

The research reported in this work is my own work except where otherwise stated, and has not been submitted for any other degree.

Allan John Dunlop

Acknowledgements

Firstly I would like to thank my supervisors, Dr. Chris McInerny and Professor Gordon Lindsay for giving me the opportunity to work in their laboratories and for their continued help, support and advice throughout the last three years. I am extremely grateful.

Thanks also extend to Professor Richard Cogdell for advice and encouragement, Dr. Sharon Kelly and Mr. Tom Jess for their expertise in performing circular dichroism experiments and Dr. Andy Pitt for mass spectrometric analyses. I am also indebted to the BBSRC for funding. I would also like to thank all members of the McInerny and Lindsay laboratories (past and present): Heather Lindsay, Dr. Mark Anderson, Dr. Stewart White, Dr. Lesley Cunliffe, Dr. Tracy Riddell, Dr. Zakia Maqbool, Dr. Audrey Brown, Dr. Donna McGow, Dr. Louise Gourlay, Dr. Alison Prior, Farzana Khaliq, Szu Shien Ng, Kyriaki Papadopoulou and Saeeda Bhatti.

Thanks also to friends and family, especially my grandparents, Karen, and Paul and Stuart, for their patience and encouragement and my parents for their continued support.

Abstract

The fission yeast DSC1 (DNA synthesis control) transcription factor complex regulates cell cycle-specific periodic transcription of a group of genes at the G1-S phase transition during the mitotic cell cycle, by binding to MCB (*Mlu*I cell cycle box) sequence elements common to their promoters. Included in this group are several genes whose functions are required for the onset and progression of S phase, such as *cdc22*$^{+}$ (the large subunit of ribonucleotide reductase), *cig2*$^{+}$ (the major S phase cyclin) and the DNA replication licensing factors *cdc18*$^{+}$ and *cdt1*$^{+}$. In concert with cyclin-dependent kinase activity, DSC1 function is required for passage of START and entry into the mitotic cell division cycle. Similar gene expression programmes exist in both budding yeast and humans controlled by the SBF/MBF and E2F transcription factors, respectively.

Fission yeast DSC1 comprises two related DNA-binding subunits, Res1p and Res2p, each bound to a single molecule of the regulatory Cdc10p protein. In addition, the Rep2p protein has a transcriptional activator function (replaced by Rep1p in the meiotic cycle). Knowledge about functional aspects of each of the DSC1 components has been greatly enhanced by genetic and biochemical studies. However, to date, these proteins remain poorly characterised at the atomic level, with little known about structure beyond their amino acid sequence. The aim of this study was to clone and bacterially express the individual DSC1 genes, to provide sufficient protein to carry out more detailed biophysical and functional studies.

The *cdc10*$^{+}$, *res1*$^{+}$, *res2*$^{+}$ and *rep2*$^{+}$ genes, together with the putative meiotic subunit *rep1*$^{+}$, were cloned and overexpressed in *E. coli* as N-terminal histidine-tagged fusion proteins. Inclusion of the His-tag facilitated purification of the proteins by affinity chromatography. Each recombinant protein (with the exception of His-Cdc10p) was shown to function *in vivo*; ectopic expression of *His-res1*$^{+}$, *His-res2*$^{+}$, *His-rep1*$^{+}$ or *His-rep2*$^{+}$ rescued the cold-sensitive lethality of the fission yeast Δ*res1* mutant strain. Bacterially expressed His-Res1p, His-Res2p and His-Rep2p were recovered in soluble form, whilst His-Rep1p and His-Cdc10p were detergent-solubilised from inclusion bodies. His-Res2p was expressed and purified in yields sufficient to undertake biophysical analyses. Both His-Rep1p and His-Cdc10p were solubilised and purified from inclusion bodies in yields sufficient for structural studies, although initial biophysical data suggests that re-folding strategies will be required to obtain active preparations of these proteins.

In electrophoretic mobility shift assay experiments, neither His-Res1p nor His-Res2p displayed detectable MCB-specific DNA-binding *in vitro*. Intriguingly, replacement of the His-tag with an N-terminal GST-tag conferred detectable MCB-specific DNA-binding upon both proteins. These results suggest that efficient DNA-binding requires dimerisation, a property

that, at least *in vitro*, is apparently not naturally intrinsic to either Resp protein. Taken together, the results presented in this study provide a significant basis with which to undertake future structural analyses of these proteins. The implications of these results for further *in vitro* studies are discussed.

Contents

Chapter 1: Introduction

Chapter 2: Materials and Methods

Index of Figures

Chapter 1

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Index of Tables

Abbreviations

A	adenine
Å	angstrom
APC/C	anaphase promoting complex/cyclosome
ARS	autonomously replicating sequence
ATP	adenosine triphosphate
bp	base pair
C	cytosine
CAP	catabolite activator protein
CD	circular dichroism
Cdc, CDC	cell division cycle
CDK	cyclin-dependent kinase
cm	centimetre
dATP	deoxyadenosine triphosphate
dCTP	deoxycytosine triphosphate
Da	dalton
DDK	Dfp1-dependent kinase
dH_2O	distilled water
dIdC	deoxyinosinic-deoxycytidylic acid
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleosidetriphosphate
DP	DRTF1 protein
DRTF1	differentiation-regulated transcription factor 1
DSC1	DNA synthesis control 1
DTT	dithiothreitol
E2F	E2A binding factor
ECL	enhanced chemiluminescence
EDTA	ethylene diamine tetraacetic acid
EMM	Edinburgh minimal media
g	gram
G	guanine

GB	glasgow collection number for bacteria
GG	glasgow collection number for fission yeast
GO	glasgow collection number for oligonucleotides
GST	glutathione *S*-transferase
G0	quiescence
G1	gap 1
G2	gap 2
h	hour
His	histidine
HRP	horseradish peroxidase
IPTG	isopropyl β-thiogalactopyranoside
ITC	isothermal titration calorimetry
kb	kilobasepairs
kDa	kilodalton
kV	kilovolts
l	litre
LB	luria bertani
M	molar
M (phase)	mitosis
mA	milliamp
mM	millimolar
Mb	megabasepairs
MBF	MCB-binding factor
MCB	*Mlu*I cell cycle box
MCM	mini-chromosome maintenance
mg	milligram
min	minute
ml	millilitre
mRNA	messenger RNA
M_r	relative molecular mass
nm	nanometre
NMR	nuclear magnetic resonance
nmt	no message in thiamine
ORC	origin recognition complex
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction

PEG	polyethyleneglycol
PMSF	phenylmethanesulfonylfluoride
pRB	retinoblastoma gene product
pre-RC	pre-replication complex
R	restriction point
RNA	ribonucleic acid
RNase	ribonuclease
rpm	revolutions per minute
S (phase)	synthesis (phase)
SBF	SCB-binding factor
SCB	Swi4-Swi6 cell cycle box
SDS	sodium dodecyl sulphate
sec	second
SPR	surface plasmon resonance
T	thymine
TBE	Tris/Borate/EDTA buffer
TE	Tris/EDTA buffer
TEMED	N, N, N', N'-tetramethylethylenediamine
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
ts	temperature sensitive
U	units
UV	ultraviolet
µCi	microcuries
µg	microgram
µl	microlitre
µM	micromolar
V	volts
v/v	volume to volume
wHTH	winged helix-turn-helix
w/v	weight to volume
YE	yeast extract
°C	degrees Celsius
Δ	deletion

Chapter 1

Introduction

1.1 Introduction

The aim of cell division is to distribute a single, complete and accurate copy of the genome to each of two daughter cells. The most prominent events in this process are chromosomal replication and segregation. In eukaryotes, the mitotic cell division cycle comprises a highly complex series of events, which are precisely regulated and co-ordinated. The entire process is typically divided into four phases of the order: G1, S and G2 (collectively known as interphase), followed by M phase and cellular division. The chromosomes are replicated during S (synthesis) phase and a single copy of each is then segregated into two identical daughter nuclei during M (mitotic) phase, followed by cytokinesis. Two gap phases, G1 and G2, intercede between these two main processes, allowing time for growth and repair (Alberts et al., 1994). During G1, the cell monitors its environment before making the decision to enter into S phase. Entry into S phase marks an irreversible decision point, committing the cell to a new round of mitotic cell division. Prior to entry into S phase, the cell can undergo alternative developmental fates. The G2 phase provides a safety gap, ensuring that DNA replication is complete (and any DNA damage has been repaired) and the cell has grown to a sufficient size prior to division. Mitosis can be further sub-divided into prophase, metaphase, anaphase and telophase. The chromosomes condense during prophase, align in metaphase, separate during anaphase, decondense in telophase and are segregated into separate cells following cytokinesis. This is the basic pattern of cell division followed by all eukaryotic cells (Alberts et al., 1994).

The molecular mechanisms that are responsible for controlling the major events of the mitotic cell cycle are widely conserved in evolution, from yeast to man. This has facilitated the use of several diverse organisms, each with their own particular experimental advantages, as model systems in cell cycle research. The conservation of many of the salient features has led to a unified theory of eukaryotic cell cycle control (Nurse et al., 1998; Nurse, 2000). This has allowed the application of knowledge gained from more basic systems to higher eukaryotes and humans in particular. This is especially relevant with respect to humans, where aberrant control of the cell cycle can manifest in potentially lethal conditions such as cancer (Ford and Pardee, 1999).

Prominent in these investigations have been the budding yeast *Saccharomyces cerevisiae* and the fission yeast *Schizosaccharomyces pombe*, which have pioneered the elucidation of many of the major cell cycle controls (Hartwell, 1991; Nurse et al., 1998).

1.2 Life cycles and genetics of the budding and fission yeasts

Studies in both budding yeast and the distantly related fission yeast have been particularly valuable in cell cycle research. Despite being unicellular, both yeasts have most of the basic features that are typical of more complex eukaryotes, making them excellent model systems for studying the eukaryotic cell cycle (Lew et al., 1997; Forsburg and Nurse, 1991). Significantly, budding and fission yeasts are widely divergent in evolutionary terms and so cell cycle controls conserved between these two organisms most likely reflect mechanisms conserved throughout all eukaryotes (Sipiczki, 2000).

The haploid genomes of both budding and fission yeasts have been sequenced and are similar in size; 12 Mb and 14 Mb, respectively (Goffeau et al., 1996; Wood et al., 2002). The budding yeast genome is distributed amongst 16 chromosomes, whereas fission yeast contains only 3 chromosomes that are 5.7, 4.7 and 3.5 Mb in size. In budding yeast, cells can grow stably in either the haploid or diploid state. The mitotic cell division cycle begins with formation of a growing bud on the parent cell, increasing in size as the cycle progresses. Eventually, the bud pinches off producing an initially smaller, though genetically identical, daughter cell (Forsburg and Nurse, 1991; Lew et al., 1997). In contrast, fission yeast cells are stable only as haploids, although the transient diploid state can be maintained under appropriate conditions. Fission yeast cells grow by increasing length, dividing by septation and medial fission. The life cycles of these two yeast species are summarised in Figures 1.1 and 1.2.

Important advances in cell cycle research have been primarily due to the isolation of mutants in both yeasts, which block at specific points or exhibit altered regulation of the cell cycle. Cells that have arrested in cell cycle progress, due to mutation of a gene whose product is required for cell cycle progress, are called cell division cycle mutants (Nurse et al., 1976). Mutants with cell cycle defects can be easily distinguished microscopically in both yeasts. In fission yeast, *cdc* mutants continue to grow without dividing, forming elongated cells at the non-permissive temperature. Similarly, *CDC* mutants of budding yeast are recognised by their bud morphology, which is broadly indicative of the point of arrest (Hartwell et al., 1974).

The fission yeast cell cycle has distinct G1, S, G2 and M phases similar to that of higher eukaryotes. In budding yeast, however, many of the common cytological markers of the G2 and M phases (i.e. spindle pole body duplication, mitotic spindle formation) appear early in the cycle, to allow formation of the bud and migration of the nucleus to the bud neck. Budding yeast therefore lacks clear definition between the S, G2 and M phases of its mitotic cycle (Forsburg and Nurse, 1991).

There are two major control points in both yeasts mitotic cell cycle, one in late G1 called 'START' and the other in late G2 regulating mitotic entry. Similar to higher eukaryotes, budding yeast chiefly regulates its cell cycle at the G1-S transition (Hartwell, 1974). In rapidly

growing fission yeast S phase is normally initiated before the completion of cytokinesis and thus the G1 phase is very short. Under these conditions, fission yeast chiefly regulates its cell cycle at the G2-M transition where information on cell size and nutrition is monitored (MacNeill and Nurse, 1997). This situation is advantageous to a haploid organism as the majority of the cycle is spent in possession of two full copies of the genome, which may facilitate protection against DNA damage (Forsburg & Nurse, 1991; Humphrey, 2000). Indeed, the size control operating at G1 is normally cryptic, only becoming apparent under nutritional limitation or in certain cell cycle mutants. In the so-called *wee* mutants, cells divide before the parent has grown to its normal size, thus forming daughter cells that are smaller than normal. Consequently, the G1 phase is lengthened to allow cells to reach a critical size before entry into S phase (Nurse, 1975; Nurse, 1990).

In both organisms, passage of START and entry into S phase requires prior completion of mitosis in the previous cycle and a minimal cell size, in addition to propitious environmental conditions. When nutrients are in plentiful supply, cells grow and enter the mitotic cell cycle leading to DNA replication, chromosome segregation and the production of two genetically identical daughter cells. In the absence of nutrients the cell can undergo either of two alternative fates. Cells can exit from the mitotic cycle and accumulate in stationary phase (a metabolically dormant state where cells await the return of propitious conditions before returning to the mitotic cycle). Alternatively, in the presence of cells of the opposite mating type, conjugation can occur, producing a diploid zygote. In budding yeast, diploid cells are stable and can undergo mitotic as well as meiotic division. In contrast, the diploid state in fission yeast is short-lived, almost immediately entering meiosis and sporulation.

A control point similar to START exists in mammalian cells, known as the 'Restriction point' (R), which also corresponds to an irreversible commitment to mitotic division (Pardee, 1974; Blagosklonny and Pardee, 2002). Similarly, mammalian cells can also exit the cell cycle prior to the Restriction point, entering into a quiescent state known as G0. This state is often confused with the stationary phase of the two yeasts, which results from nutrient limitation that causes cells to accumulate before the size control that operates at START (Bartlett and Nurse, 1990). Withdrawal from the cell cycle in higher eukaryotes results from the absence of mitogenic growth factors (or indeed the presence of negative regulatory growth factors). Consequently, as the cell is not starved of nutrients, it remains metabolically active, albeit at a reduced level (Lodish et al., 1995).

In all cells, once START/R is passed the cell is committed to the mitotic cell cycle, until it returns to G1 (Forsburg and Nurse, 1991). Passage of the G2 control point and entry into mitosis is also dependent on the cell reaching a certain minimum size, as well as prior completion of S phase and repair of any DNA damage.

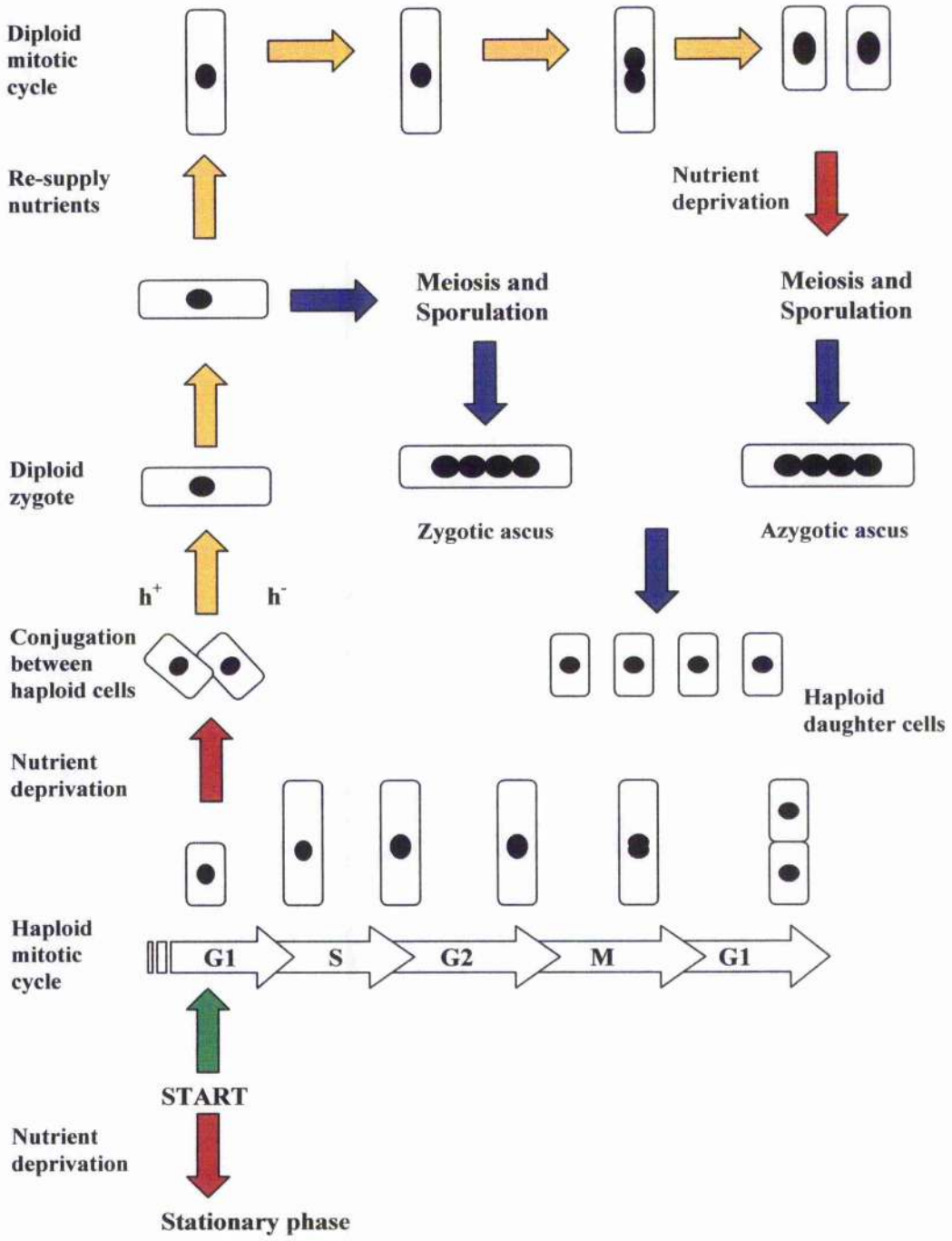


Figure 1.1 The life cycle of the fission yeast *Schizosaccharomyces pombe*
Fission yeast cells grow stably as haploids and propagate by mitotic division. In the absence of sufficient nutrients cells can exit from mitotic growth and enter into stationary phase. Alternatively, in the presence of cells of opposite mating type (h^+ and h^-) conjugation occurs forming a diploid zygote. The diploid state is unstable in fission yeast and proceeds directly into the meiotic cycle, producing a zygotic ascus containing four haploid ascospores. In favourable conditions the four haploid spores can again re-enter the mitotic cycle. The transient diploid state can be maintained if nutrients are re-supplied immediately following conjugation. In this case, the diploid cell can undergo diploid mitotic growth until nutrients become limiting. Starved diploid cells proceed into the meiotic cycle producing an azygotic ascus.

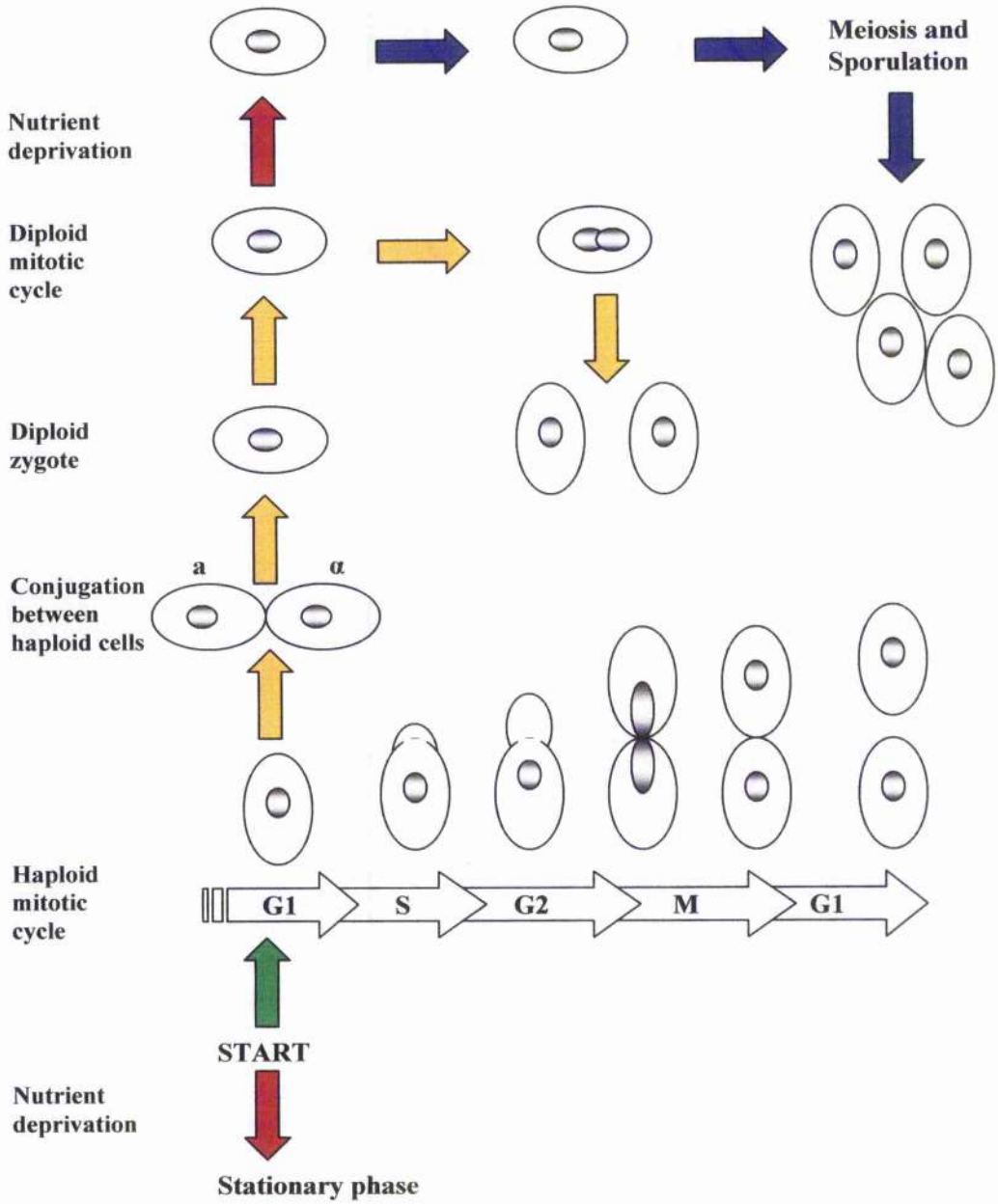


Figure 1.2 The life cycle of the budding yeast *Saccharomyces cerevisiae*
Budding yeast cells are capable of growth both as stable haploids and diploids. In the absence of sufficient nutrients cells can exit from mitotic growth and enter into stationary phase. Alternatively, in the presence of cells of opposite mating type (a and α) conjugation occurs forming a diploid zygote. If nutrients are re-supplied, the diploid cell enters into the diploid mitotic cycle. However, if starved, diploids cells proceed into meiosis and sporulation forming four haploid cells, which can re-enter the mitotic cell cycle under favourable conditions.

1.2.1 A note on nomenclature

Throughout this thesis, reference will be made primarily to proteins involved in fission yeast cell cycle control. However, in the course of the discussion, proteins with analogous functions in both budding yeast and humans will be referred to. Therefore, in the interest of clarity, proteins from different organisms will be distinguished by the following prefixes: Sp (*Schizosaccharomyces pombe*), Sc (*Saccharomyces cerevisiae*) and Hs (*Homo sapiens*).

1.3 The molecular mechanisms of the mitotic cell cycle

As described in Section 1.2, in both budding and fission yeasts and mammalian cells, the decision to enter a new round of chromosome duplication and cellular division is made at a control point in late G1 phase known as START/R. There are two main molecular mechanisms responsible for controlling passage through START/R and entry into S phase, cyclin-dependent kinase (CDK) activity and the cell cycle-regulated periodic transcription of genes that are essential for S phase onset and progression.

The latter of these two mechanisms, which is central to this thesis, is discussed in more detail later (Section 1.6). The temporal association of distinct cyclin-CDK complexes, whose activities regulate the phase-specific events, primarily control the orderly progression of the cell cycle. In budding yeast a single CDK, ScCdc28p, is the major regulator of the cell cycle. In G1 phase ScCdc28p associates with the G1 cyclins ScCln1p, ScCln2p and ScCln3p driving passage of START and entry into S phase, in tandem with the MBF and SBF factors responsible for G1-S phase-specific transcription (Section 1.5.1). ScCdc28p then controls the onset of S phase and DNA replication in association with the ScClb5p and ScClb6p cyclins, whilst the ScClb1-4p cyclins govern the onset and progression of mitosis (Nasmyth, 1996).

In mammalian cells there is more than one major CDK controlling the cell cycle, with HsCdk1p, HsCdk2p, HsCdk4p and HsCdk6p the most prominent. Passage of the Restriction point and entry into S phase requires HsCdk4p/HsCdk6p-cyclin D and HsCdk2p-cyclin E kinase activities and the expression of several genes under control of the E2F family of transcription factors (Section 1.7). HsCdk2p then associates with cyclin A during S phase and G2 whilst HsCdk1p-cyclin B drives mitosis (Sherr, 1996).

Similarly to budding yeast, a single CDK, SpCdc2p, drives the cell cycle in fission yeast and requires the expression of several genes under control of the DSC1 complex, the focus of this thesis (Section 1.6). The molecular control mechanisms of the fission yeast mitotic cell cycle are described in more detail below. Many of the main regulatory processes are conserved throughout evolution and consequently, several of the molecules and mechanisms are similar in both budding yeast and higher eukaryotes, and are alluded to where appropriate.

1.3.1 The molecular mechanisms of the mitotic cell cycle in fission yeast

1.3.2 The G1-S transition

In fission yeast both the SpCdc2p kinase and DSC1-dependent transcription activities are required for passage of START and entry into S phase, (Nurse and Bissett, 1981; MacNeill and Nurse, 1997). In early G1 phase the mitotic form of SpCdc2p (associated with the SpCdc13p cyclin) predominates, enduring from the previous M-phase (Booher et al., 1989; Moreno et al., 1989). However, SpCdc2p-SpCdc13p kinase activity is prohibited at this stage by CDK inhibition and cyclin proteolysis (Moser and Russell, 2000). The CDK inhibitor SpRum1p plays a major role in cell cycle regulation during G1 phase by binding to and inhibiting SpCdc13p. In addition, SpRum1p targets SpCdc13p for ubiquitin-mediated proteolysis by the 26S proteasome (Correa-Bordes et al., 1997; Benito et al., 1998).

Ubiquitin-mediated proteolysis plays an important role in the cell cycle, targeting proteins for degradation by the 26S proteasome in G1-S phases and also during mitosis (Stone and Gordon, 2003). In mitosis, proteolysis is required for sister chromatid separation (at the metaphase to anaphase transition) and the degradation of mitotic cyclins, essential for the exit from mitosis (Section 1.3.6). The anaphase-promoting complex/cyclosome (APC/C) is a cell cycle-regulated multimeric complex that catalyses the transfer of ubiquitin to target proteins (Morgan, 1999; Harper et al., 2002). The APC/C is regulated by the reversible binding of activator proteins, which provide substrate specificity. In fission yeast, the SpSlp1p and SpSte9p proteins are important for APC/C function. SpSte9p only interacts with APC in G1, where it targets the mitotic cyclins, SpCdc13p and SpCig1p, leftover from the previous M phase. SpSte9p itself is negatively regulated by SpCdc2p-dependent phosphorylation, causing dissociation from the APC/C and subsequent degradation (Blanco et al., 2000).

The combined effects of SpRum1p and APC/C-SpSte9p ensure that any SpCdc2p-SpCdc13p kinase activity remaining from the previous M phase cannot act in G1. This is crucial, since inappropriate triggering of mitotic events at this point could induce cells to divide with unreplicated chromosomes, leading to aneuploidy.

Meanwhile, during G1, the DSC1 complex activates transcription of the major S-phase cyclin, SpCig2p (Connolly and Beach, 1994; Obara-Ishihara and Okayama, 1994). Initially, SpRum1p inhibits SpCdc2p-SpCig2p kinase activity (Benito et al., 1998). However, the SpPuc1p and SpCig1p cyclins (both of which are insensitive to SpRum1p) associate with SpCdc2p, leading to the phosphorylation and inactivation of SpRum1p (Martin-Castellanos et al., 2000; Moser and Russell, 2000). Consequently, SpRum1p is degraded allowing SpCdc2p-SpCig2p activity to increase, thereby promoting entry into S-phase (Martin-Castellanos et al., 1996; Mondesert et al., 1996; Benito et al., 1998).

1.3.3 Onset of S phase

The molecular mechanisms that regulate DNA replication in eukaryotes are only partly understood. However, many of the factors involved are largely conserved, with the result that a general picture of the regulation of DNA replication initiation has emerged from genetic and biochemical studies in both yeasts and higher eukaryotes (Kelly and Brown, 2000; Lei and Tye, 2001).

1.3.3.1 Origins of replication

At the onset of S phase, DNA replication is initiated at multiple specific sites throughout the genome known as origins of replication (Kelly and Brown, 2000). These *cis*-acting DNA elements have been particularly well characterised in budding yeast, and are known as autonomous replicating sequences (ARS) based on their ability to promote the autonomous replication of plasmids (Marahens and Stillman, 1992). Budding yeast ARSs are defined by an 11 bp AT-rich consensus sequence that is distributed over sequence blocks of 100-200 bp (Marahens and Stillman, 1992). The sequences defining replication origins in fission yeast and higher eukaryotes are less well defined, although AT-rich regions also appear to be important (Okuno et al., 1999; Masukata et al., 2003).

1.3.3.2 The origin recognition complex

In all eukaryotes examined, replication origins are bound by a hetero-hexameric origin recognition complex (ORC), which in fission yeast comprises the SpOrp1p-SpOrp6p proteins (Moon et al., 1999; Chaung et al., 2002). The ORC was originally identified in budding yeast based on its ability to bind the well-defined ARS and is widely conserved; analogous complexes are found in higher eukaryotes (Bell and Stillman, 1992; Kelly and Brown, 2000). Studies in fission yeast have shown that ORC remains bound to chromatin, via SpOrp4p, throughout the cell cycle (Lygerou and Nurse, 1999; Chuang and Kelly, 1999). ORC appears to function as a scaffold, upon which occurs the ordered assembly of several proteins, culminating in the recruitment and activation of the DNA replication machinery (Brown and Kelly, 2000; Lei and Tye, 2001).

1.3.3.3 The mini-chromosome maintenance complex

Tight control over the initiation of DNA replication is crucial to ensure that it occurs only once, and at the appropriate time, in each cell cycle. Of particular importance, in this context, is the recruitment of the mini-chromosome maintenance complex (MCM) to origins, in a process known as DNA licensing. Similarly to ORC, the *MCM* genes were originally identified in budding yeast, as mutants unable to support the maintenance of mini-chromosomes (i.e. replication of plasmids; Tye, 1999). The MCM complex is a hetero-hexamer composed of the

SpMcm2p-SpMcm7p proteins, all of which share conserved nucleotide binding domains, and is widely conserved throughout eukaryotes (Forsburg, 2004). Accessory proteins, which themselves are precisely regulated, tightly control the loading of the MCM complex onto origins.

1.3.3.4 SpCdc18p, SpCdt1p and the control of DNA licensing

The critical factors for DNA licensing in fission yeast are the SpCdc18p and SpCdt1p proteins, homologues of which exist in budding yeast and higher eukaryotes, ScCdc6p/HsCdc6p and ScTah11p/HsCdt1p, respectively (Kelly et al., 1993; Hofmann and Beach, 1994; Lygerou and Nurse, 2000; Tanaka and Diffley, 2002). Their presence is essential for loading the SpMcm2p-SpMcm7p complex onto origins associated with ORC, thereby licensing DNA for replication (Nishitani et al., 2000; Lygerou and Nurse, 2000). SpCdc18p and SpCdt1p associate with chromatin independently of each other, yet act co-operatively to bring about the construction of the pre-replicative complex (pre-RC) on origins (Figure 1.3; Lygerou and Nurse, 2000). Both proteins are transcribed in late M phase, under the control of the DSC1 complex (Baum et al, 1998; Nishitani et al., 2000). At this point, SpCdc18p levels are kept low by SpCdc2p-SpCdc13p-dependent phosphorylation and subsequent degradation (Jallepalli et al, 1997; Lopez-Girona et al, 1998). However, as cells exit mitosis and enter G1, the mitotic SpCdc2p-SpCdc13p kinase activity decreases (Section 1.3.2) and accumulation of SpCdc18p ensues (Baum et al, 1998; Jallepalli et al, 1997; Lopez-Girona et al, 1998). The combination of transcriptional and post-translational controls ensures a sharp increase in SpCdc18p as cells exit mitosis and enter G1, with the presence of both SpCdc18p and SpCdt1p proteins bound to ORC facilitating recruitment of the SpMcm2p-SpMcm7p complex (Nishitani et al, 2000).

1.3.3.5 Activation of DNA replication: CDK and DDK

Following assembly of the pre-RC, DNA replication is activated by the actions of two sets of protein kinases, the CDK SpCdc2p-SpCig2p and the DDK (Dfp1-dependent kinase) SpHsk1p-SpDfp1p. The SpCdc2p-SpCig2p kinase is the major S phase promoting CDK in fission yeast (Section 1.3.2). The SpHsk1p protein is a member of the Cdc7p family of protein kinases that are essential in all eukaryotes for the initiation of DNA replication (Brown and Kelly, 1998). Although SpHsk1p alone has kinase activity, association with SpDfp1p provides substrate specificity (Brown and Kelly, 1998; Brown and Kelly, 1999). The abundance of SpDfp1p is regulated throughout the cell cycle both transcriptionally and post-transcriptionally (Brown and Kelly, 1999).

Although the specific targets of SpCdc2p-SpCig2p are unclear, with respect to replication initiation, SpHsk1p-SpDfp1p can phosphorylate the SpMcm2p-SpMcm7p complex *in vitro*, on SpMcm2p and SpMcm4p (Brown and Kelly, 1998; Lee et al., 2003). In fission

yeast, a sub-complex of the SpMcm-4p -6p and -7p proteins has been shown to possess helicase activity *in vitro*, and it has been proposed that phosphorylation of the SpMcm2p-SpMcm7p complex could activate intrinsic DNA helicase activity, which may be important for DNA unwinding at replication forks (Lee and Hurwitz, 2001). SpHsk1p-SpDfp1p-mediated phosphorylation of SpMcm2p is dependent on the association of the DDK with the SpCdc23p protein, suggesting an accessory role for this protein in recruiting SpHsk1p-SpDfp1p to origins (Lee et al, 2003). Recent experiments have revealed a further role for SpCdc23p in recruiting the ScCdc45p homologue SpSna41p to origins (Gregan et al., 2003).

The recruitment of SpSna41p is likely to be important for loading of DNA polymerase α onto the SpMcm2p-7p complex (Uchiyama et al., 2001). Activation of DNA replication then begins following recruitment of other components of the replication machinery (Diffley and Labib, 2002). Despite the lack of clarity regarding specific phosphorylation targets, the CDK and DDK activities initiate a chain of events that ultimately leads to establishment of functional replication forks (Diffley and Labib, 2002). Following replication, the two daughter DNA molecules remain tightly associated in a process called sister chromatid cohesion. This requires a number of factors and can only occur during DNA replication (Section 1.3.6). Cohesion occurs at specific sites, including centromeres, and is crucial for mitosis and also important for DNA repair. Termination occurs when two replication forks meet and the nascent DNA from the two forks is ligated together (Diffley and Labib, 2002).

1.3.3.6 Prevention of re-replication

In addition to functioning as an activator of DNA replication, SpCdc2p-SpCig2p also has a negative role in preventing re-initiation from origins following activation (Jallepalli et al., 1997; Lopez-Girona et al., 1998). This apparently contradictory behaviour is crucial, both to ensure proper initiation of DNA replication and the prevention of re-replication. Whilst SpCdc18p accumulates (in late mitosis) as a result of declining SpCdc2p-SpCdc13p activity, it is phosphorylated and targeted for degradation by the rising SpCdc2p-SpCig2p kinase activity, in late G1 (Jallepalli et al., 1997; Lopez-Girona et al., 1998). Furthermore, *cdc18*$^{+}$ expression declines during S phase (Kelly et al., 1993). As a consequence of this dual control, SpCdc18p dissociates from origins, thereby preventing re-initiation at origins that have already fired. Consistent with *cdt1*$^{+}$ transcription being under DSC1 control, the expression pattern is similar to *cdc18*$^{+}$ (Hofmann and Beach, 1994; Nishitani et al., 2000). The protein levels also fluctuate in parallel with SpCdc18p, suggesting that SpCdt1p may also be degraded in a CDK-dependent process (Nishitani et al., 2000). Thus, both CDK-dependent proteolysis and cell cycle-regulated periodic transcription conspire to ensure that re-initiation is prevented until the following M phase has passed. It has recently been shown that the SpCdc2p-SpCdc13p kinase also plays a

crucial role in preventing re-initiation by binding to replication origins (via ORC), presumably by phosphorylating and inactivating components of the pre-RC (Wuarin et al., 2002).

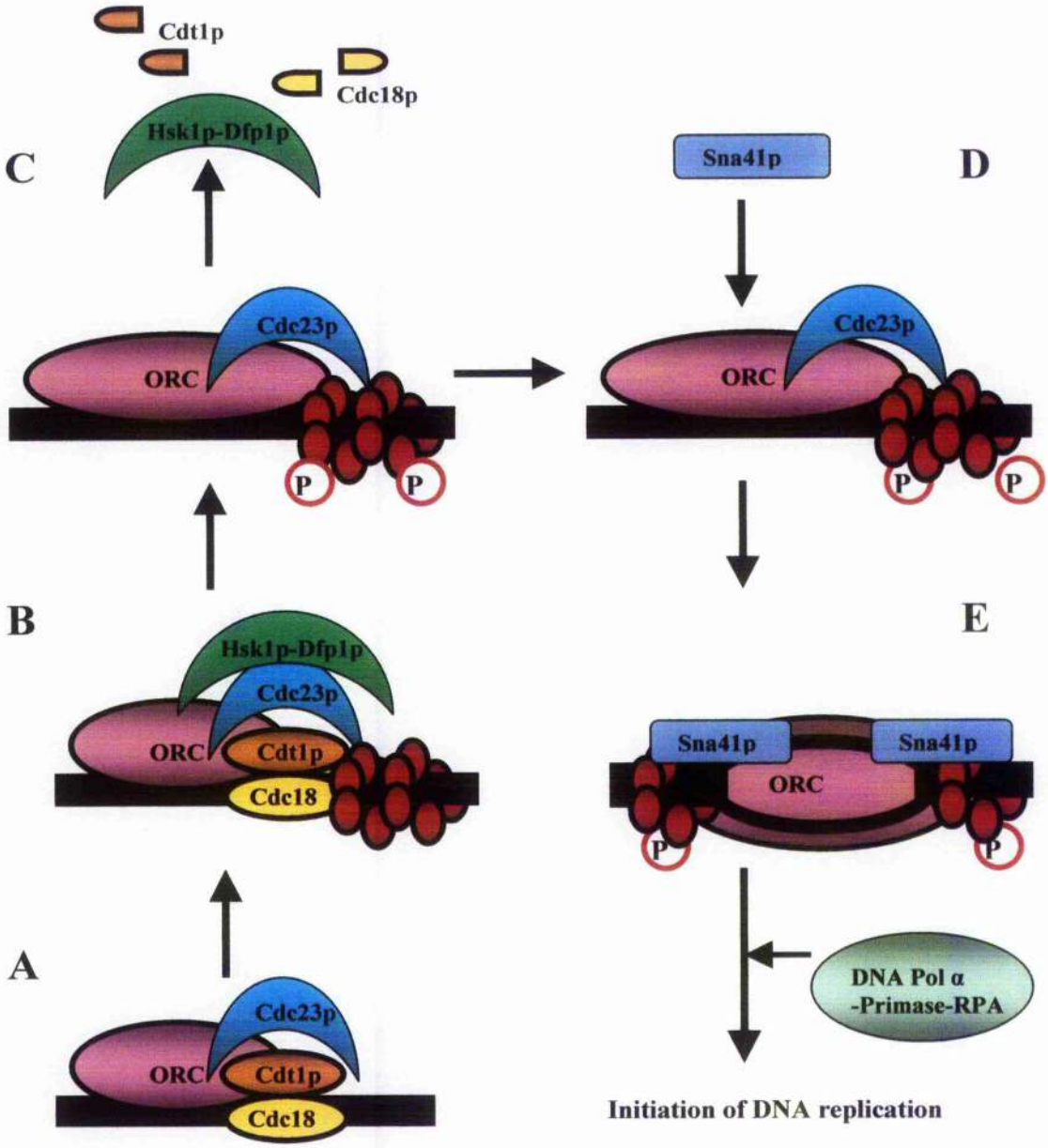


Figure 1.3 Model for fission yeast DNA licensing and replication
(**A**) ORC is bound to origins throughout the cell cycle. At the end of M phase, Cdc18p and Cdt1p are recruited independently to ORC at origins to which Cdc23p binds. (**B**) Cdc18p and Cdt1p then recruit the MCM complex, thereby forming the pre-RC. This is known as DNA licensing. (**C**) The SpCig2p-SpCdc2p and SpHsk1p-SpDfp1p protein kinases then play essential roles in triggering the initiation of DNA replication. SpCdc18p and SpCdt1p are removed for degradation and SpHsk1p-SpDfp1p phosphorylates the SpMcm2p-SpMcm7p complex, thereby promoting its helicase activity. (**D**) After activation of these protein kinases, additional factors such as SpSna41p are recruited, and origins are unwound. (**E**) Finally, DNA polymerase α, single strand DNA-binding protein (RPA) and primase are recruited and semi-conservative replication ensues (Adapted from Lei and Tye, 2001).

1.3.4 The G2-M transition

At the end of G2 phase, when cells reach a critical size, M phase is induced due to increasing SpCdc2p-SpCdc13p kinase activity (Nurse, 1990; Moser and Russell, 2000). Activation of SpCdc2p-SpCdc13p is essential for entry into mitosis and represents the rate-limiting step (Nurse, 1990). After passing START, SpCdc2p is associated with SpCdc13p in S phase, staying associated throughout the entire length of G2, but activity is kept low by inhibitory phosphorylation on Tyrosine-15 (Y-15), carried out by the SpWee1p and SpMik1p kinases (MacNeill and Nurse, 1997). To induce mitosis, SpCdc25p de-phosphorylates SpCdc2p at Y-15. Activation of SpCdc2p-SpCdc13p is promoted by the activation of SpCdc25p and the SpWee1p inhibitory kinases, SpCdr1p and SpCdr2p (MacNeill and Nurse, 1997).

The *mik1*$^{+}$ gene is under transcriptional control of the DSC1 complex (Ng et al., 2001). Consequently, *mik1*$^{+}$ mRNA and protein levels increase in S-phase (Baber-Furnari et al., 2000; Christensen et al., 2000). This suggests that the presence of SpMik1p in S phase is important for maintaining the low SpCdc2p-SpCdc13p activity during G2 phase (Baber-Furnari et al., 2000; Christensen et al., 2000). In contrast, SpWee1p levels are constant throughout the cell cycle (Moser and Russell, 2000). SpCdc25p appears to be regulated at a translational level, since SpCdc25p synthesis is particularly sensitive to reduced translation initiation activity (Daga and Jimenez, 1999). This suggests that control of SpCdc25p translation initiation may form part of the mechanism coupling cell growth with cell division (Daga and Jimenez, 1999; Kellogg, 2003). Co-ordination of these mitotic inhibitors and activators ultimately leads to an increase in SpCdc2p-SpCdc13p kinase activity, which facilitates the subsequent mitotic events (Nurse, 1990; Su and Yanagida, 1997).

1.3.5 Checkpoint controls: DNA integrity

Before a cell can exit G2 phase and enter into mitosis, it must ensure that the genome has been fully and faithfully replicated. Two signal transduction cascades (known as checkpoints) operate in eukaryotes to ensure that when DNA is damaged or replication perturbed, normal cell cycle progression is delayed (Hartwell and Weinert, 1989). Delayed mitotic entry permits completion of either the replication or repair processes. These control mechanisms are crucial to prevent the propagation of potentially genotoxic mutations and cell division with an incomplete genome.

The main effectors of the checkpoint pathways are the SpChk1p and SpCds1p protein kinases, which function in response to DNA-damage and stalled DNA replication signals, respectively (Rhind and Russell, 2000; Boddy and Russell, 2001). These signals are sensed and relayed by the checkpoint SpRadp proteins, which in turn control the kinase activity of SpChk1p and SpCds1p. Several of the gene products involved in checkpoint control in fission yeast are widely conserved in evolution (Carr and Caspari, 2003).

In fission yeast, the checkpoint kinases indirectly maintain the inhibitory Y-15 phosphorylation of SpCdc2p, thereby preventing mitotic entry (Section 1.3.4; Rhind et al., 2000). Accordingly, the targets of SpChk1p and SpCds1p are the SpCdc25p and SpMik1p proteins, although the exact mechanisms are unclear (Rhind et al., 2000). Regulation of SpWee1p does not seem to be important for checkpoint function (Rhind and Russell, 2001). It has been proposed that, in response to DNA damage, SpChk1p negatively regulates SpCdc25p phosphatase activity (Rhind et al., 2000). Consistent with its role during S phase, SpMik1p levels are increased following DNA damage, and therefore may also contribute to the checkpoint response by increasing its SpCdc2p Y-15 directed kinase activity (Baber-Furnari, et al., 2000). Checkpoint-mediated inhibition of SpCdc25p and activation of SpMik1p ensures that SpCdc2p Y-15 phosphorylation is maintained, thereby preventing entry into mitosis.

1.3.6 Mitosis and cytokinesis

The major processes that occur during M phase, required to divide the replicated genome equally into two cells, are essentially similar in all eukaryotes examined (Alberts et al, 1994).

Mitosis is characterised by a highly complex series of cytoskeletal and nuclear re-arrangements, which together achieve chromosome condensation, segregation and cytokinesis (Alberts et al, 1994). These events are triggered by the M phase CDK-dependent phosphorylation of numerous substrates. When M phase CDK activity reaches a critical level, cells proceed from G2 into mitosis. Conversely, mitotic exit relies on inactivation of this CDK activity via ubiquitin-mediated proteolysis of the mitotic cyclins (Morgan, 1999). Loss of mitotic CDK activity is controlled by the actions of the highly regulated APC/C (Section 1.3.2; Morgan, 1999; Harper et al., 2002).

In fission yeast, the numerous events of mitosis are brought about by the rise in SpCdc2p-SpCdc13p kinase activity, controlled at the G2-M transition by the action of SpCdc25p (Section 1.3.4; MacNeill and Nurse, 1997). SpCdc2p-SpCdc13p is the major mitotic kinase in fission yeast, although more recently additional roles for the Polo, Aurora and NIMA kinases have been described (Nigg, 2001).

Preparation for mitosis begins during S phase, when newly replicated sister chromatids are joined together by a complex of proteins collectively known as cohesins, and the spindle pole body is duplicated (Nasmyth et al., 2000). In fission yeast, the cohesin complex is composed of four proteins, SpPsm1p, SpPsm3p, SpRad21p and SpPsc3p (Hagstrom and Meyer, 2003). Cohesion between sister chromatids is required to prevent premature separation prior to the onset of anaphase, important in preventing aneuploidy.

Prophase: The onset of M phase is then initiated by compaction of chromosomes, in a process known as condensation. Chromosome condensation is required to re-organize the loosely packed interphase assortment into highly compact structures. This is necessary to

permit attachment of the mitotic spindle to centromeres and allow separation without entanglement. A multi-protein complex, known as condensin, effects the necessary changes in chromosome structure. In fission yeast, condensin is composed of an SpCut14p-SpCut3p hetero-dimer associated with SpCnd1p, SpCnd2p and SpCnd3p (Hagstrom and Meyer, 2003).

Metaphase: The opposing forces of the mitotic spindles, pulling sister chromatids toward opposite poles of the cell, are counteracted by the maintenance of chromosome cohesion. As a result, the sister chromatids align along the metaphase plate. At this point, a spindle checkpoint mechanism monitors the attachment of the mitotic spindle to chromosomes, delaying anaphase until all chromosomes are properly attached and under tension.

Anaphase: Following bipolar attachment of the chromosomes to the spindle, anaphase begins. During anaphase, cohesion between sister chromatids is dissolved and chromosomes are separated. Anaphase onset is induced by the activating phosphorylation of the APC/C, at this point associated with SpSlp1p. APC/C-SpSlp1p facilitates separation of sister chromatids by ubiquitylating a protein commonly known as securin. The securin protein is normally complexed with a protease separin, which is held inactive in this heterodimeric state (Yanagida, 2000). It is thought that ubiquitin-mediated proteolysis of securin liberates the separin subunit, thereby releasing its proteolytic potential. In fission yeast, SpCut2p and SpCut1p are the putative securin and separin proteins, respectively (Yanagida, 2000). The liberated separin protease then degrades the cohesins responsible for sister chromatid attachment. Loss of cohesion releases the physical tension at the metaphase plate, thereby allowing mitotic spindles to elongate, segregating the sister chromatids to opposite poles of the cell.

Mitotic exit (telophase and cytokinesis): Exit from mitosis requires the inactivation of mitotic CDK activity. Loss of CDK activity is mediated by the destruction of the mitotic cyclins, by ubiquitin-mediated proteolysis (Morgan, 1999). A conserved signalling cascade governs the co-ordination of late mitotic events in both budding and fission yeasts, known as the mitotic exit network (MEN) and septation initiation network (SIN), respectively (McCollum and Gould, 2001; Bardin and Amon, 2001). The SIN is a GTPase-regulated protein kinase cascade, which functions to regulate the initiation of cytokinesis at the end of anaphase.

1.4 Cell cycle-regulated transcription

One of the key mechanisms deployed in controlling the cell cycle is the phase-specific regulation of transcription. Typically, several groups of genes are co-ordinately expressed at a particular period in the cell cycle, commonly corresponding to the phase during which their function is required (Breeden, 2003; McInerny, 2004). Co-ordinated transcription programmes are dependent on the recognition of specific *cis*-acting DNA sequences in gene promoters, by trans-acting proteins. These sequence-specific transcription factors in turn recruit chromatin re-modelling machines and components of the general transcriptional apparatus, thereby controlling transcriptional initiation (Levine and Tjian, 2003).

It has been proposed that phase-specific transcription programmes may serve to limit the functions of particular gene products to a particular point in the cycle at which they are required (Breeden, 2003; McInerny, 2004). In some cases this may be important for cell cycle progression *per se*. For example, inappropriate expression of a cyclin (or another key cell cycle regulator) outwith its normal period of activity may be deleterious to the cell. However, in the majority of cases, a gene may be periodically expressed simply as a means of conserving resources. Nevertheless, the fact that phase-specific transcription programmes are a universal feature of the eukaryotic cell cycle (being widely conserved from yeast to humans) is indicative of its importance. Recent microarray studies have revealed the extent of this regulatory mechanism during the cell cycle. In budding yeast, ~800 genes have been identified whose transcripts are cell cycle-regulated (Spellman et al, 1998). Similarly, in fission yeast ~400 genes display a cell cycle-periodic increase in their transcription (Rustici et al., 2004) and in human cells approximately 700 genes display a cell cycle-periodic increase in their transcription (Cho et al., 2001).

1.5 Cell-cycle regulated transcription at the G1-S transition

As described previously, START/R marks the point of irreversible commitment to the mitotic cell division cycle, requiring propitious conditions and CDK activity (Section 1.2). A key function of the CDK at this stage is to activate transcription of several genes, whose functions are required for progression into S phase. Of these genes, perhaps the most salient are the G1 and S phase cyclins, which associate with the CDK to allow passage of START/R and drive downstream events. Correct temporal expression of these genes is, therefore, crucial for correct cell cycle progression.

Cell cycle-regulated transcription at the G1-S phase boundary has been particularly well studied in the budding and fission yeast systems. In budding yeast, two DNA-binding factors, known as SBF and MBF, are responsible for initiating this co-ordinated programme of gene expression (Merrill et al., 1992). In fission yeast, a single MBF-like complex, referred to

here as DSC1*, performs an analogous function (Whitehall et al., 1999). The remarkable structural and functional homology shared by the MBF/SBF and DSC1 systems, suggests that a co-ordinated programme of G1-S phase-specific transcription is widely conserved in eukaryotes. Regulation of transcription at the G1-S phase transition has been well documented in mammalian cells, under the control of the E2F complex (Dyson, 1998). Although not related to either of the yeast DNA-binding factors at the primary structure level, E2F shares some striking similarities and therefore appears to be responsible for regulating the equivalent process in humans (Section 1.7).

* DSC1 is the traditional name for the G1-S-regulatory transcription factor complex in fission yeast; however, it is now commonly referred to as MBF. Therefore, in this thesis, to avoid confusion during the discussion of both MBF complexes in the fission and budding yeasts, the former is referred to as DSC1 whilst the latter MBF.

1.5.1 G1-S phase-specific transcription in budding yeast

In budding yeast passage of START and entry into S phase requires the activity of the ScCdc28p kinase in association with the G1 cyclins, ScCln1p, ScCln2p and ScCln3p (Nasmyth, 1996; Lew et al., 1997). The ScClnp-ScCdc28p kinase activities are required for passage of START and the early events of the cell cycle: initiation of DNA replication, spindle pole body duplication and bud formation (Nasmyth, 1996; Lew et al., 1997). Transcription of *CLN1* and *CLN2* is phase-specific, peaking in late G1, whereas *CLN3* is transcribed earlier, peaking at the M-G1 boundary (Ogas et al., 1991; McInerny et al., 1997; MacKay et al., 2001). In addition, transcription of the *PCL1* and *PCL2* cyclins, which may also have a role at START, is regulated similarly to *CLN1* and *CLN2* (Espinoza et al., 1994; Measday et al., 1994). A large number of genes (>200), in addition to the G1 cyclins (excluding *CLN3*), have been identified, including the *CLB5* and *CLB6* cyclins required for entry into S phase, that are periodically transcribed in late G1 phase (Epstein and Cross, 1992; Schwob and Nasmyth, 1993; Spellman et al., 1998). These genes can be sub-divided into two groups based on the *cis*-acting regulatory sequence element present in their promoter and their cognate DNA-binding factors. Activation of both of these transcription factors requires the ScCln3p-ScCdc28p kinase activity (Wijnen et al., 2002).

1.5.2 The SBF complex

The first group of genes, which include the *CLN1*, *CLN2*, *PCL1* and *PCL2* cyclins, contain a common 5' CACGAAA 3' sequence element in their promoter, known as the ScSwi4p-ScSwi6p cell cycle-box (SCB). Transcription of these genes is activated in late G1, dependent on the heterodimeric SCB-binding factor (SBF), composed of the ScSwi6p and ScSwi4p proteins (Koch and Nasmyth, 1994; Iyer et al., 2001).

The first evidence of a role for the *SWI4* and *SWI6* gene products in transcriptional regulation was the discovery that both were required, together with a repeated promoter element (now known as the SCB), for the cell cycle-regulated and START-dependent transcription of the *HO* endonuclease gene, involved in the mating type switch (Breeden and Nasmyth, 1987a). Furthermore, the pleiotropic effects displayed in either *swi4* or *swi6* mutants (which are not seen in *ho* mutants) and the lethality of the *swi4 swi6* double mutant indicated that these gene products must function in some other essential cellular process (Breeden and Nasmyth, 1987a). The role of these proteins as transcription factors was confirmed subsequently, when they were biochemically identified (in band-shift assays), as components of an SCB-specific DNA-binding factor (Andrews and Herskowitz, 1989a; Andrews and Herskowitz, 1989b; Taba et al., 1991).

Both proteins display a modular architecture typical of transcription factors (Figure 1.4). In the SBF complex, sequence-specific DNA-binding is mediated by an N-terminal domain in

the 123 kDa ScSwi4p subunit. ScSwi4p also contains centrally located ankyrin-repeat motifs, thought to be involved in protein-protein interactions, and a C-terminal heterodimerization domain (Andrews and Herskowitz 1989b; Primig et al., 1992; Bork et al., 1993; Sedgwick and Smerdon, 1999).

The 92 kDa ScSwi6p protein has no intrinsic DNA-binding capability, instead playing a regulatory role, containing distinct N and C-terminal transcriptional activation domains (Sedgwick et al., 1998). ScSwi6p also contains the centrally located ankyrin-repeat motifs and heterodimerises with ScSwi4p via their C-termini (Breeden and Nasmyth 1987b; Primig et al, 1992; Andrews and Moore, 1992; Foord et al., 1999). Mutational analyses of ScSwi6p indicate that the ankyrin-repeat motifs and a leucine zipper region (located in the C-terminal third of the protein) are required for DNA-binding by the ScSwi4p/ScSwi6p complex (Sidorova and Breeden, 1993). The ankyrin-repeat motifs are dispensable for the association of ScSwi4p and ScSwi6p (Andrews and Moore, 1992).

1.5.3 The MBF complex

A second group of genes are also periodically expressed in late G1, and encode a wide range of factors required for S phase, including several proteins necessary for DNA synthesis (Johnston and Lowndes, 1992; Koch and Nasmyth, 1994). Although many of these gene products are stable throughout the cell cycle, and therefore not rate-limiting, others such as the *CLB5* and *CLB6* cyclins are important for cell cycle progression *per se* (Epstein and Cross, 1992; Schwob and Nasmyth, 1993). Similarly to SBF-regulated genes, this second group are also defined by a *cis*-acting regulatory element common to their promoters. This element has the consensus sequence 5' ACGCGT 3', which coincidentally corresponds to the *MluI* restriction enzyme recognition site. Thus it is known as the *MluI* cell cycle box (MCB), and is both necessary and sufficient for the cell cycle-regulated late G1 expression of these genes (McIntosh et al, 1991; Lowndes et al, 1991). The MCB motif is bound by the heterodimeric MCB-binding factor (MBF), which is composed of ScSwi6p and ScMbp1p (Lowndes et al., 1991; Lowndes et al., 1992b; Dirick et al., 1992; Koch et al., 1993; Iyer et al., 2001). Analogous to SBF, the regulatory function is provided by ScSwi6p, whilst the 120 kDa ScMbp1p protein provides the sequence-specific DNA-binding capability (Dirick et al., 1992; Koch et al., 1993). The ScMbp1p protein resembles ScSwi4p in structure as well as function, consisting of N-terminal DNA-binding and C-terminal heterodimerization domains, separated by the centrally located ankyrin-repeat motifs (Figure 1.4; Koch et al., 1993).

1.5.4 Functional redundancy in SBF and MBF-dependent gene expression

The molecular mechanisms by which SBF and MBF-dependent transcription are regulated are complex and not fully understood (Section 1.5.8). The situation is further complicated by the apparent functional redundancy between the two systems, as suggested by phenotypic analyses, most notably the lethality of *mbp1 swi4* double mutants, due to inadequate expression of the *CLN1* and *CLN2* cyclins (Koch et al., 1993). In contrast, single mutants of either gene are viable, indicating a requirement for the function of at least one complex for cell cycle progression (Koch et al., 1993). That some functional overlap should exist is perhaps not surprising, given that both systems share several common features. The SCB and MCB elements are related in sequence and when present at high concentration *in vitro*, each element can cross-compete with the other for complex binding, although SBF and MBF bind preferentially to the SCB and MCB elements, respectively (Taylor et al., 2000). In at least one case *in vivo*, G1-S-regulated transcription is apparently controlled by SBF binding to MCB-like motifs, thus indicating an overlap in binding specificity (Partridge et al., 1997).

1.5.5 A family of transcription factors

The ScSwi4p/ScMbp1p/ScSwi6p proteins of SBF/MBF and the SpRes1p/SpRes2p/SpCdc10p proteins of the fission yeast MBF-homologue DSC1 (Section 1.6) constitute a family of transcription factors that share a common function and architectural design. The sequence homology shared amongst these six proteins locates to three main regions. The N-termini show significant similarity at the primary structure level, particularly between the ScSwi4p/ScMbp1p and SpRes1p/SpRes2p proteins, which each harbour DNA-binding domains within this region (Figure 1.5). The centrally located ankyrin-repeat motifs are common to all six members, and the C-terminal domains required for heterodimerization also show similarity (Figure 1.7). X-ray crystallographic analysis of a 36 kDa domain from ScSwi6p, containing the ankyrin domain, suggests that intramolecular interactions within this region may allow the N and C termini to come into contact (Foord et al., 1999). This may be important for ScSwi4p since it has been proposed that C-terminally mediated auto-inhibition of the N-terminal DNA-binding domain is important for its regulation (Baetz and Andrews, 1999).

The 3-dimensional (3-D) structure of the N-terminal DNA-binding domain of ScMbp1p has also been solved by X-ray crystallographic analysis (Xu et al., 1997; Taylor et al., 1997). In addition, characterisation of this same N-terminal fragment of ScMbp1p at the atomic level has been enhanced by NMR studies (McIntosh et al., 2000; Nair et al., 2003).

1.5.6 The N-terminal DNA-binding domain

The crystal structure of the DNA-binding domain of ScMbp1p has been solved at 2.1 Å resolution, and is shown in Figure 1.5B (Xu et al., 1997; Taylor et al., 1997). This fragment corresponds to amino acids 1-124 and reveals a globular molecule consisting of a twisted six-stranded anti-parallel β-sheet (β1-β6) with two pairs of α-helices (αA- αB and αC- αD), which fold into a motif similar to the winged-helix-turn-helix (wHTH) family of proteins, including HNF3γ and the bacterial catabolite activator protein, CAP (Schultz et al., 1991; Clarke et al., 1993; Xu et al., 1997; Taylor et al., 1997). Helices αA and αB form the helix-turn-helix motif with the hairpin between the β5-β6 strands forming the putative 'wing'. The β1-β4 and β5-β6 β-sheets are at right angles to one another, forming a barrel with a hydrophobic core. Several of the most highly conserved non-polar residues within the ScMbp1p/ScSwi4p/SpResp proteins, particularly the aromatic amino acids, are situated within this hydrophobic core (Figure 1.5A). The greatest amino acid sequence conservation amongst these proteins is situated within the wHTH region (from β3 to β6) whilst C-terminal to this the sequences diverge (Xu et al., 1997; Taylor et al., 1997).

1.5.7 Model for DNA-binding

The conserved charged and polar residues of the wHTH all lie on one face of the molecule, most notably within the αB helix, thereby forming a positively charged surface and suggesting an involvement in DNA-binding. The αB helix has been proposed to be responsible for the major sequence interactions, within the major groove of the DNA double helix (Xu et al., 1997; Taylor et al., 1997;Taylor et al., 2000; Nair et al., 2003). The β-hairpin ('wing') may also make contact with the DNA (Figure 1.6). Surprisingly, the non-conserved residues C-terminal to the core are required for efficient DNA binding, suggesting that two distinct domains are involved in DNA-binding (Taylor et al., 2000; Nair et al., 2003). The sequence homology between the ScMbp1p/ScSwi4p/SpResp proteins within the DNA-binding domains suggests they all adopt a similar fold and thus share a common mode of DNA-binding. However, the SpResp proteins are believed to bind to DNA as dimers, whereas both ScMbp1p and ScSwi4p bind to their respective recognition sequences as monomers (Section 1.6.6; Taylor et al., 2000). In this respect the SpRes1p/SpRes2p proteins might resemble the CAP and E2F/DP proteins, which also utilise a similar wHTH fold for DNA-binding (Schultz et al., 1991; Zheng et al., 1999).

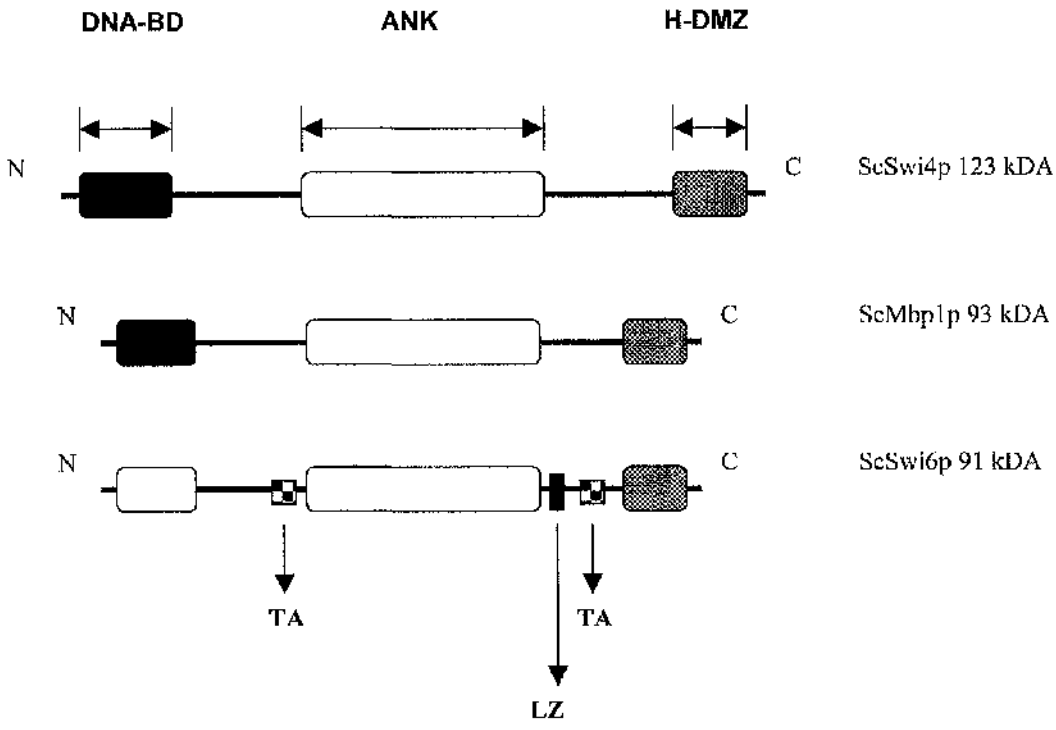
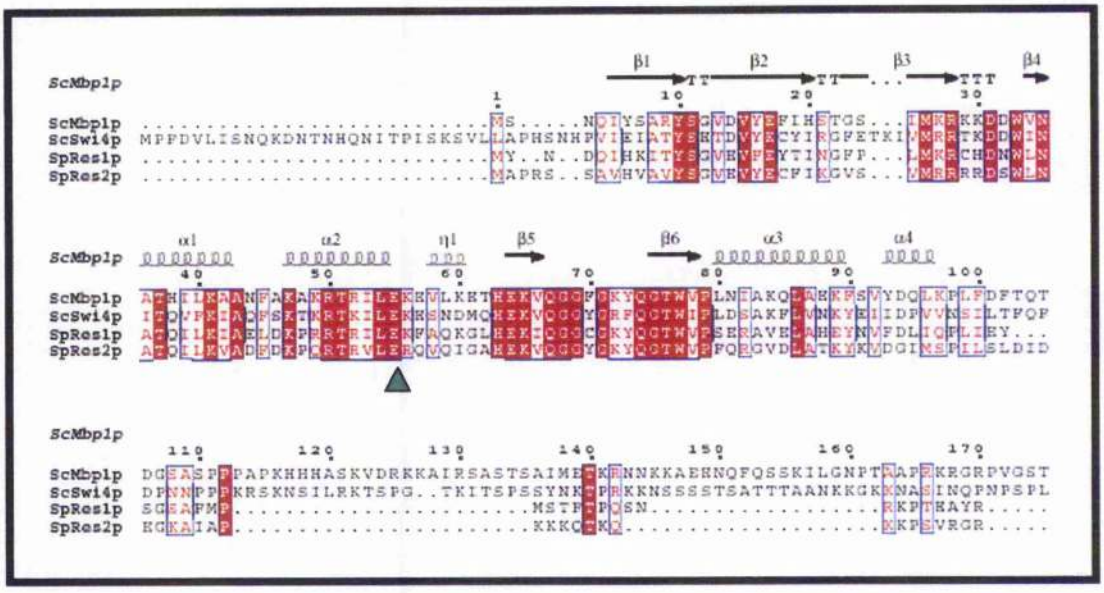


Figure 1.4 Domain architecture of ScSwi4p, ScMbp1p and ScSwi6p
A schematic representation of the major domains in ScSwi4p, ScMbp1p and ScSwi6p: DNA-BD = DNA-binding domain, ANK = Ankyrin repeat domain, H-DMZ = Heterodimerization domain, TA = Transcriptional activation domain, LZ = Leucine zipper (note that the N-terminal domain of ScSwi6p does not bind to DNA).

A



B

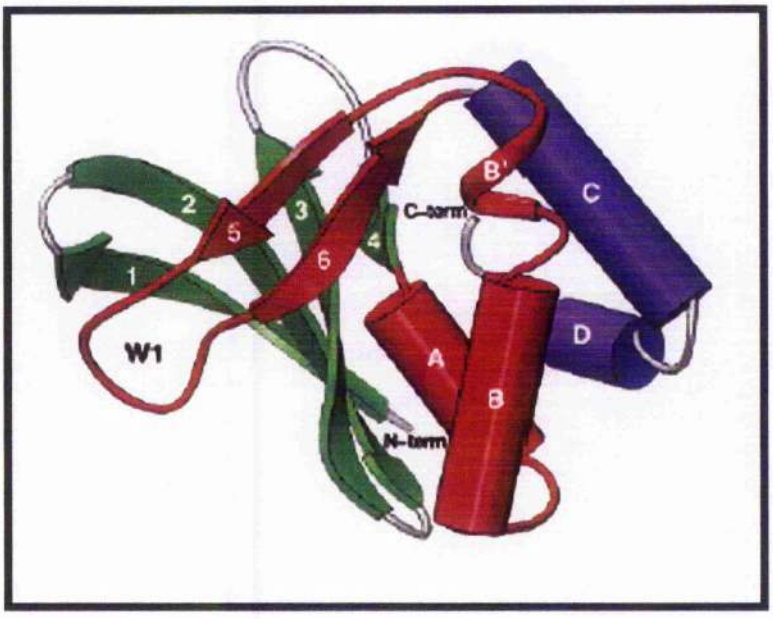


Figure 1.5

A. Alignment of ScMbp1p, ScSwi4p, SpRes1p and SpRes2p DNA-binding domains

Sequence alignment of the N-terminal DNA-binding domains of ScMbp1p, ScSwi4p, SpRes1p and SpRes2p. The residue numbers and secondary structure elements are annotated for ScMbp1p. Identical residues are shown as white characters boxed in red with similarity denoted by red characters boxed in white. β strands (arrows), α helices (squiggle) and β turns (TT) are also shown. The position of the Glu 56 → Lys mutation of SpRes1p is indicated (▲ - Section 1.6.6). This figure was generated using the ESPript program (Gouet et al., 1999).

B. ScMbp1p DNA-binding domain

Ribbon diagram of the X-ray crystal structure of the ScMbp1p DNA-binding domain (Taylor et al., 2000).

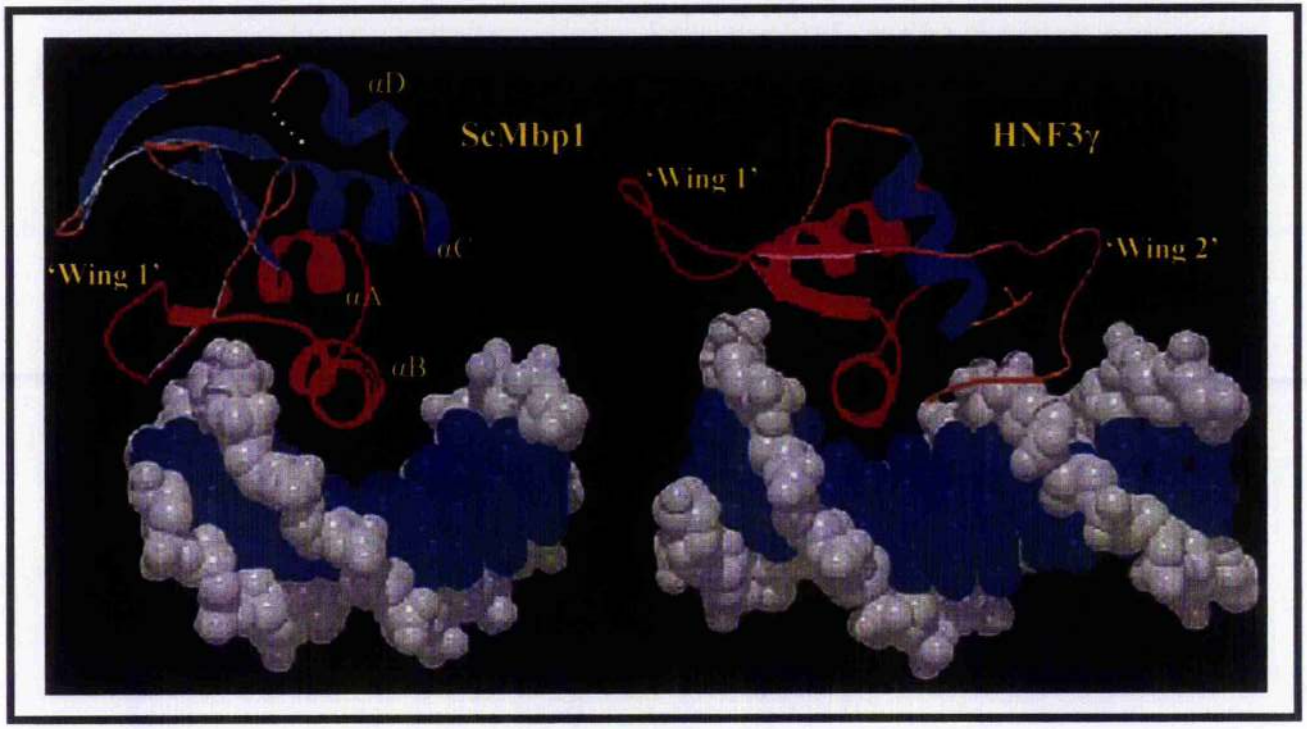


Figure 1.6 Model for ScMbp1p DNA-binding
A putative model for ScMbp1p DNA-binding compared to the co-crystal structure of the HNF3γ-DNA complex (Clarke et al., 1993; adapted from Taylor et al., 2000).

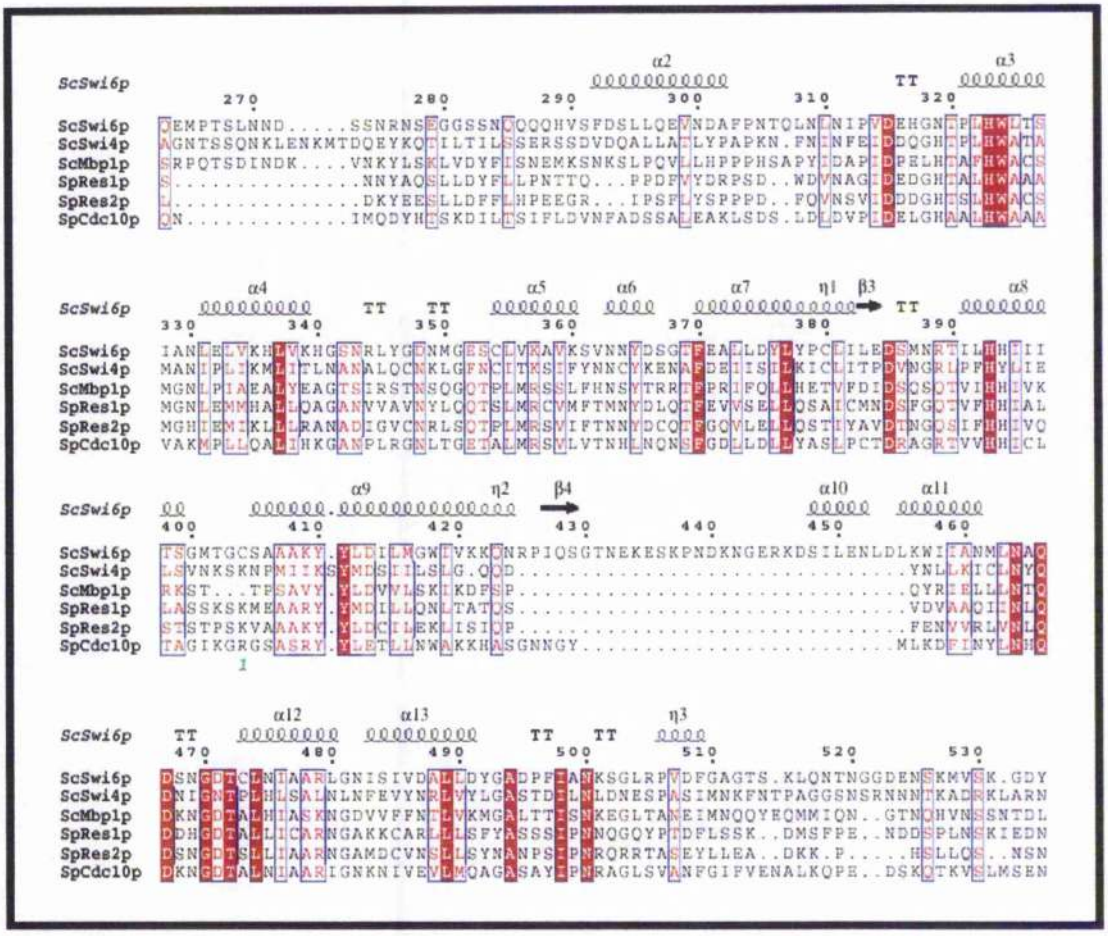


Figure 1.7 Alignment of the ScSwi6p/SpCdc10p ankyrin domains
Sequence alignment between the members of the ScSwi6p/SpCdc10p family within the ankyrin-repeat region. The residue numbers and secondary structure elements are annotated for ScSwi6p. Identical residues are shown as white characters boxed in red with similarity denoted by red characters boxed in white. β strands (arrows), α helices (squiggle) and β turns (TT) are also shown. This figure was generated using the ESPript program (Gouet et al., 1999).

1.5.8 Regulation of SBF and MBF-dependent gene expression

Despite the structural and functional similarities displayed by the SBF and MBF systems, it appears that distinct mechanisms exist to regulate their shared aim of G1-S phase-specific transcription.

Activation of both SBF and MBF-dependent transcription at START is dependent on the ScCln3p-ScCdc28p kinase activity. Although the precise mechanism by which this occurs is unknown, it appears that neither complex is a direct target of the CDK (Wijnen et al., 2002). A recent study suggests that the ScCln3p-ScCdc28p kinase may activate SBF-dependent transcription indirectly by antagonising an SBF-associated repressor (De Bruin et al., 2004). ScCln3p-ScCdc28p-dependent phosphorylation of the ScWhi5p protein promotes its dissociation from SBF, thereby allowing recruitment of the RNA pol II holoenzyme, leading to transcriptional activation (De Bruin et al., 2004; Cosma et al., 1999).

Paradoxically, in addition to a positive role in MBF/SBF-dependent transcription, the ScCln3p-ScCdc28p kinase seemingly also negatively regulates MBF activity. The ScStb1p protein has recently been identified as an MBF-associated activator protein (via its interaction with ScSwi6p). However, ScCln3p-ScCdc28p-dependent phosphorylation leads to dissociation of ScStb1p and the consequent down-regulation of MBF-dependent transcription (Costanzo et al., 2003). In contrast, repression of SBF-regulated genes during G2 and M is dependent on the ScClbp-ScCdc28p kinases, which have no apparent effect on MBF-regulation genes (Amon et al, 1993). Consistent with this SBF-specific function, it has been proposed that down-regulation of SBF activity may occur through the direct interaction of ScSwi4p with the ScClb2p-ScCdc28p kinase during G2 and M (Siegmund and Nasmyth, 1996).

The subcellular localisation of ScSwi6p is cell cycle-regulated by phosphorylation. Predominantly nuclear during late M and early G1 phases, ScSwi6p becomes largely cytoplasmic in S, G2 and early M phases, only re-appearing in the nucleus following completion of mitosis (Taba et al., 1991; Sidorova et al., 1995). It has been proposed that phosphorylation of ScSwi6p (on Serine 160) by the ScClb6p-ScCdc28p kinase, facilitates nuclear export, by the karyopherin ScMsn5p, thereby resulting in down-regulation of MBF and SBF-dependent transcription (Queralt and Igual, 2003; Gaymonet et al., 2004). Intriguingly, however, the ScMsn5p-mediated export of ScSwi6p is specifically required for SBF function, but not MBF, again highlighting different regulatory properties (Queralt and Igual, 2003).

Nevertheless, the recruitment of ScSwi6p to the nucleus cannot be solely responsible for activating transcription. The SBF complex is bound to target gene promoters in early G1 phase, yet transcription occurs only in late G1 (Harrington and Andrews, 1996; Koch et al., 1996). This suggests a subsequent activation event must occur, presumably related to the functions of the ScWhi5p and ScStb1p proteins mentioned above.

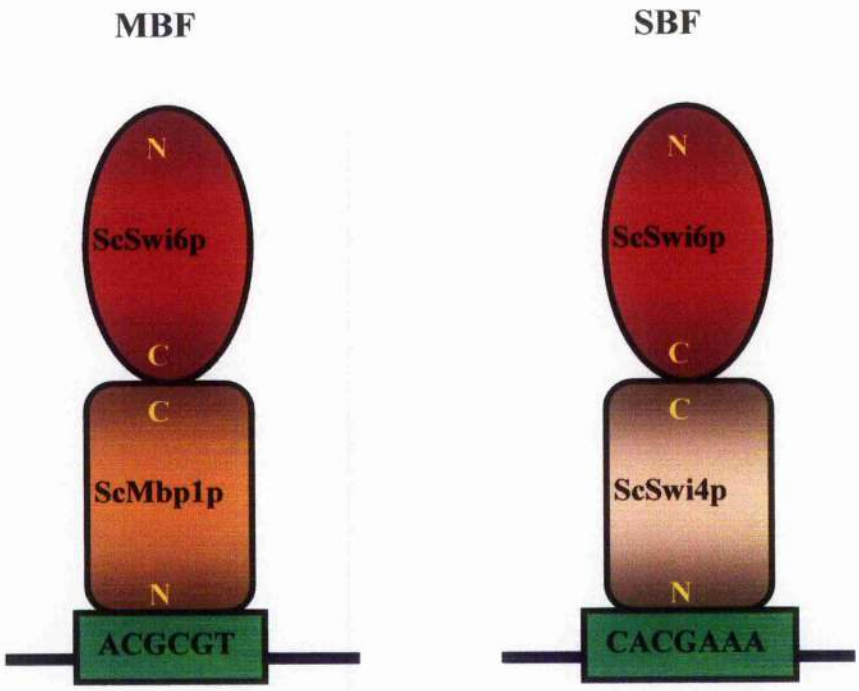


Figure 1.8 The MBF and SBF DNA-binding complexes
Schematic representations of the MBF and SBF DNA-binding complexes bound to the MCB and SCB sequence elements, respectively. The N and C-termini of each protein are indicated in yellow.

The *SWI4* gene is itself transcribed in a cell cycle-regulated manner, peaking in late M-G1 phase, thereby contributing to the temporal activation of SBF in late G1 (McKay et al., 2001). An additional level of control over the SBF complex has also been proposed through regulation of ScSwi4p DNA-binding by C-terminally mediated auto-inhibition (Baetz and Andrews, 1999).

1.6 G1-S phase-specific transcription in fission yeast

In fission yeast, passage of START and entry into S phase requires the activity of the SpCdc2p kinase in association with the G1 cyclins, SpPuc1p, SpCig1p and SpCig2p (Fisher and Nurse, 1995). Transcription of several genes that are essential for S phase onset and progression is cell cycle-regulated, peaking in late G1 (e.g. *cdc18*$^{+}$ and *mik1*$^{+}$).

In contrast to the situation in budding yeast, where the G1-S phase-specific transcription programme is controlled by two distinct transcription factors (SBF and MBF), fission yeast contains only one such factor. The DSC1 complex is responsible for controlling the late G1-S phase-specific transcription programme during mitotic cell division. Following isolation of a MCB-specific DNA-binding activity in budding yeast (MBF - Section 1.5.3) a similar DNA-binding activity was identified in fission yeast, bound specifically to the MCB elements present in the promoter of the *cdc22*$^{+}$ gene (Lowndes et al., 1992a; Maqbool et al., 2003). The *cdc22*$^{+}$ gene, which encodes the large subunit of ribonucleotide reductase, is periodically expressed in late G1-S and contains seven copies of the MCB element in its promoter (Gordon and Fantes, 1986; Fernandez-Sarabia et al., 1993; Maqbool et al., 2003). In addition, the *cdc10*$^{+}$ gene product was shown to be a component of this factor, shortly followed by the discovery that ScSwi6p was also a component of the budding yeast MBF complex (Lowndes et al., 1992a; Lowndes et al., 1992b). The amino acid sequence homology between the ScSwi4/6p and SpCdc10p proteins, allied to their roles as START-specific transcription factors, indicates that the mechanisms governing late G1-specific transcriptional control might be conserved in these distantly related yeasts (Merrill et al., 1992).

1.6.1 The DSC1 complex

Similar to the MBF system of budding yeast, the DSC1 complex regulates transcription of several genes required for entry into, and onset of, S phase by binding to the MCB elements common to their promoters (MacNeill and Nurse, 1997). The transcription of at least 10 genes (*cdc22*$^{+}$, *cdc18*$^{+}$, *cig2*$^{+}$, *cdt1*$^{+}$, *rad21*$^{+}$, *suc22*$^{+}$, *rad11*$^{+}$, *ste9*$^{+}$, *mik1*$^{+}$ and *cdt2*$^{+}$) required for S phase and DNA replication are under the control of DSC1 during the mitotic cell cycle (Lowndes et al., 1992; Fernandez-Sarabia et al., 1993; Kelly et al., 1993; Connolly and Beach, 1994; Hofmann and Beach, 1994; Birkenbihl and Subramani, 1995; Harris et al., 1996; Parker et al., 1997; Tournier and Millar, 2000; Ayte et al., 2001; Ng et al., 2001; Yoshida et al., 2003; Maqbool et al., 2003). DSC1 is composed of at least four gene products: SpCdc10p, SpRes1p, SpRes2p and SpRep2p and has been widely studied during the mitotic cell cycle (Lowndes et al., 1992; Caligiuri & Beach, 1993; Tanaka et al., 1992; Zhu et al., 1994; Miyamoto et al., 1994; Nakashima et al., 1995). In DSC1, the SpRes1p and SpRes2p proteins confer sequence-specific DNA-binding activity and interact with the regulatory SpCdc10p and co-activating SpRep2p subunits.

1.6.2 A 'DSC1-like' complex in meiosis

A 'DSC1-like' complex has been recently identified, which regulates transcription in the meiotic cycle and is composed of at least SpRes2p, SpCdc10p and SpRep1p (Cunliffe et al., 2004). In addition, microarray analysis has identified ~100 genes that are transcribed specifically during late G1-S phase in the meiotic cycle, which also have promoter regions enriched for the MCB motif (Mata et al., 2002).

In this specialised cell cycle, the chromosomes of a diploid cell are replicated forming homologous pairs (as opposed to sister chromatids in mitosis). During the first cellular division, cohesion is maintained between homologous chromosome pairs, allowing the exchange of genetic material – known as homologous recombination (Alberts et al., 1994). Consequently, the first nuclear and cellular division is equational. Each cell, now with diploid genomic content, subsequently undergoes a second meiotic division, without intervening DNA replication. Four haploid cells (spores in yeast, gametes in mammals) are ultimately produced from a single diploid cell (Alberts et al., 1994). In fission yeast, meiosis is initiated after two haploid cells (of opposite mating type) arrest in G1 phase (pre-START), as a result of nutrient depletion (Figure 1.1). These cells then conjugate, forming a transient diploid, which undergoes meiotic division to produce a four-spored ascus. This specialised form of cell division requires re-programming of the mitotic cell, such that it retains many of the factors required for DNA replication, whilst additionally producing many specialised meiotic gene products (Lee and Amon, 2001). The 'DSC1-like' complex regulates transcription of several genes during meiosis, some of which also function in mitotic S phase (e.g. *cdc22*$^{+}$ and *cdc18*$^{+}$), whilst others

(e.g. the *rec*$^{+}$ gene products involved in homologous recombination) have roles that are exclusively meiotic (Cunliffe et al., 2004). The genes encoding the SpRes1p, SpRes2p, SpRep1p and SpRep2p proteins were originally identified by genetic suppressor analyses, designed to isolate components of the regulatory pathway required for the *cdc10*-dependent passage of START.

1.6.3 SpCdc10p

The *cdc10*$^{+}$ gene was first identified as one of two essential genes, whose functions were required for passage of START and commitment to the mitotic cell cycle in fission yeast (Nurse et al., 1976; Nurse and Bisset, 1981). Mutants in *cdc10*$^{+}$ arrest in G1 and retain the ability to conjugate, indicating they cannot traverse START until after the execution of SpCdc10p function (Nurse and Bisset, 1981). The *cdc10*$^{+}$ gene was subsequently cloned by rescue of mutant function, and encodes a protein of 85 kDa (Aves et al., 1985). SpCdc10p shares striking sequence similarity with the ScSwi6p protein of budding yeast, specifically within the centrally located ankyrin motifs, which are characteristic of the SpCdc10p/ScSwip family of proteins (Aves et al., 1985; Breeden and Nasmyth, 1987b - Figures 1.7 and 1.9).

Similar to the role of ScSwi6p in SBF/MBF, SpCdc10p has no intrinsic DNA-binding activity, instead functioning as a regulatory subunit (Zhu et al., 1994; McInerny et al., 1995). Mutational analyses of SpCdc10p indicate that the C-terminal region and the ankyrin motifs are important for normal function (Reymond and Simanis, 1993). The majority of *cdc10* ts mutations locate within the ankyrin-repeat region indicating it has an important, although as yet undefined, role (Reymond et al., 1992). The C-terminus of SpCdc10p most likely mediates the regulatory function, demonstrated by the properties of the *cdc10-C4* mutant (McInerny et al., 1995). The SpCdc10-C4p protein product is truncated at the C-terminus, which results in loss of the final 61 amino acids. In *cdc10-C4* mutant cells, genes that are under DSC1 control show a loss in periodicity, becoming constitutively transcribed throughout the cell cycle (McInerny et al., 1995). By analogy with ScSwi6p, the C-terminus of SpCdc10p may be responsible for heterodimerization with the MCB-specific Resp DNA-binding proteins.

1.6.4 SpRes1p and SpRes2p

Three different groups cloned the *res1*$^{+}$ gene independently, as a multi-copy suppressor of the temperature sensitive *cdc10-129* mutant (Tanaka et al., 1992; Marks et al., 1992; Caligiuri and Beach, 1993). Concomitantly, SpRes1p was biochemically identified as a component of the DSC1 band-shift activity present in crude fission yeast protein extracts (Caligiuri and Beach, 1993). The *res2*$^{+}$ gene was then cloned, as a multi-copy suppressor of the Δ*res1* mutant and independently in a genetic screen designed to identify DNA-binding partners of SpCdc10p (Miyamoto et al., 1994; Zhu et al., 1994). The demonstration that SpRes2p, together with

SpCdc10p, bound specifically to MCB elements both *in vitro* and *in vivo*, confirmed that, like SpRes1p, it was a DNA-binding partner of SpCdc10p (Zhu et al., 1994).

1.6.5 Functional redundancy between SpRes1p and SpRes2p

The molecular mechanisms by which DSC1-dependent transcription is regulated, either in the mitotic or meiotic cycles, is not fully understood (Section 1.6.8). The situation is further complicated by the apparently distinct yet overlapping functions of the two DNA-binding proteins, SpRes1p and SpRes2p. Both SpRes1p and SpRes2p function during the mitotic cycle and display some overlap with respect to these roles. However, SpRes2p has an additional function in the meiotic cycle, which cannot be assumed by SpRes1p.

Cells simultaneously deleted for both *res1*$^+$ and *res2*$^+$ are inviable, and overexpression of either gene can rescue the conditional lethality of the *cdc10-129* mutant (Miyamoto et al., 1994). Furthermore, overexpression of *res2*$^+$ rescues the lethality of cells deleted for *res1*$^+$, demonstrating that SpRes2p can compensate for the loss of SpRes1p mitotic function (Miyamoto et al., 1994). Despite a shared mitotic role, the phenotypes of the Δ*res1* and Δ*res2* mutants indicate that only SpRes2p functions during meiosis and, moreover, that SpRes1p and SpRes2p act predominantly in the mitotic and meiotic cycles, respectively (Tanaka et al., 1992; Caligiuri and Beach, 1993; Miyamoto et al., 1994; Zhu et al., 1994).

Cells deleted for *res2*$^+$ are viable and display no obvious growth defects during mitotic division, yet they are defective in pre-meiotic DNA synthesis and undergo an abnormal meiosis (Miyamoto et al., 1994; Zhu et al, 1994). In contrast, cells deleted for *res1*$^+$ show no obvious defects in the meiotic cycle, yet when dividing mitotically they display severe heat and cold sensitivities, arresting pre-START at the restrictive temperature (Tanaka et al., 1992). Furthermore, *res1*$^+$ cannot suppress the meiotic defect of Δ*res2* cells (Miyamoto et al., 1994; Sturm and Okayama, 1996). In fact, SpRes1p levels decrease during meiosis and SpRes1p is thought to be a repressor of the sexual pathway (Caligiuri and Beach, 1993; Ayte et al., 1997). SpRes2p levels increase during conjugation, consistent with its additional meiotic function (Miyamoto et al., 1994; Zhu et al., 1994; Ayte et al., 1997).

1.6.6 Structure of SpRes1p and SpRes2p

The apparent contradiction that SpRes1p and SpRes2p have distinct roles yet are functionally overlapping, has been explained by analysis of their functional domains.

The *res1*$^+$ and *res2*$^+$ genes encode 72 kDa and 73 kDa proteins, respectively, which display the modular architecture typical of transcription factors. Consistent with an overlapping function, SpRes1p and SpRes2p share striking sequence similarity with each other in their N-termini, which include their DNA-binding domains, and the centrally located ankyrin-repeat motifs that are characteristic of the SpCdc10p/ScSwi6p family of proteins (Tanaka et al., 1992; Caligiuri and Beach, 1993; Miyamoto et al., 1994; Zhu et al., 1994; Ayte et al., 1995; Zhu et al., 1997). This sequence similarity extends to include the ScSwi4p/ScMbp1p proteins, which share a common architectural design, affirming their place as members of this family of transcription factors (Section 1.5.5 and Figure 1.5). As described in Section 1.5.6, the X-ray structure of the ScMbp1DNA-binding domain has been solved (Xu et al., 1997; Taylor et al., 1997). In contrast, no structural data beyond the amino acid sequence is available for any of the fission yeast members of this family. The DNA-binding domain of ScMbp1p, folds into a wHTH motif, which makes sequence specific contacts in the major groove of the DNA (Section 1.5.6). Based on the significant sequence homology displayed within this region between family members, it has been proposed that SpRes1p and SpRes2p will also bind DNA using this structural motif (Xu et al., 1997; Taylor et al., 1997). In agreement with this, a mutation within the SpRes1p DNA-binding domain (Glu 56 → Lys), within the putative recognition helix, results in enhanced DNA-binding and, moreover, renders SpRes1p independent of SpCdc10p (Caligiuri and Beach, 1993; Figure 1.5).

Unlike ScMbp1p and ScSwi4p, the SpRes1p and SpRes2p proteins are believed to bind to DNA as dimers (Ayte et al., 1997; Zhu et al., 1997). This resembles the situation for the E2F/DP family of proteins, which are responsible for regulating the analogous program of G1-S phase-specific transcription in mammalian cells (Zheng et al., 1999). Although not conserved in primary structure with any of the ScSwip/SpResp family proteins, the E2F family of transcription factors also adopt a wHTH fold to bind DNA (Zheng et al., 1999; Section 1.7).

Similarly to the ScSwi6p-ScSwi4p/ScMbp1p interaction in budding yeast SBF/MBF, SpRes1p and SpRes2p interact with SpCdc10p via their C-termini (Ayte et al., 1995; Zhu et al., 1997). The distinct functional specificities of SpRes1p and SpRes2p, however, are determined by their C-termini, in which they display greatest heterogeneity (Figure 1.9; Sturm and Okayama, 1996; Zhu et al., 1997; Whitehall et al., 1999). A domain within the C-terminus of SpRes2p confers an as yet undefined meiotic-specific function (Sturm and Okayama, 1996; Zhu et al., 1997; Whitehall et al., 1999). In addition, two discrete domains in SpRes2p confer a requirement for the transcriptional co-activator subunit, SpRep2p (Sturm and Okayama, 1996; Tahara et al., 1998).

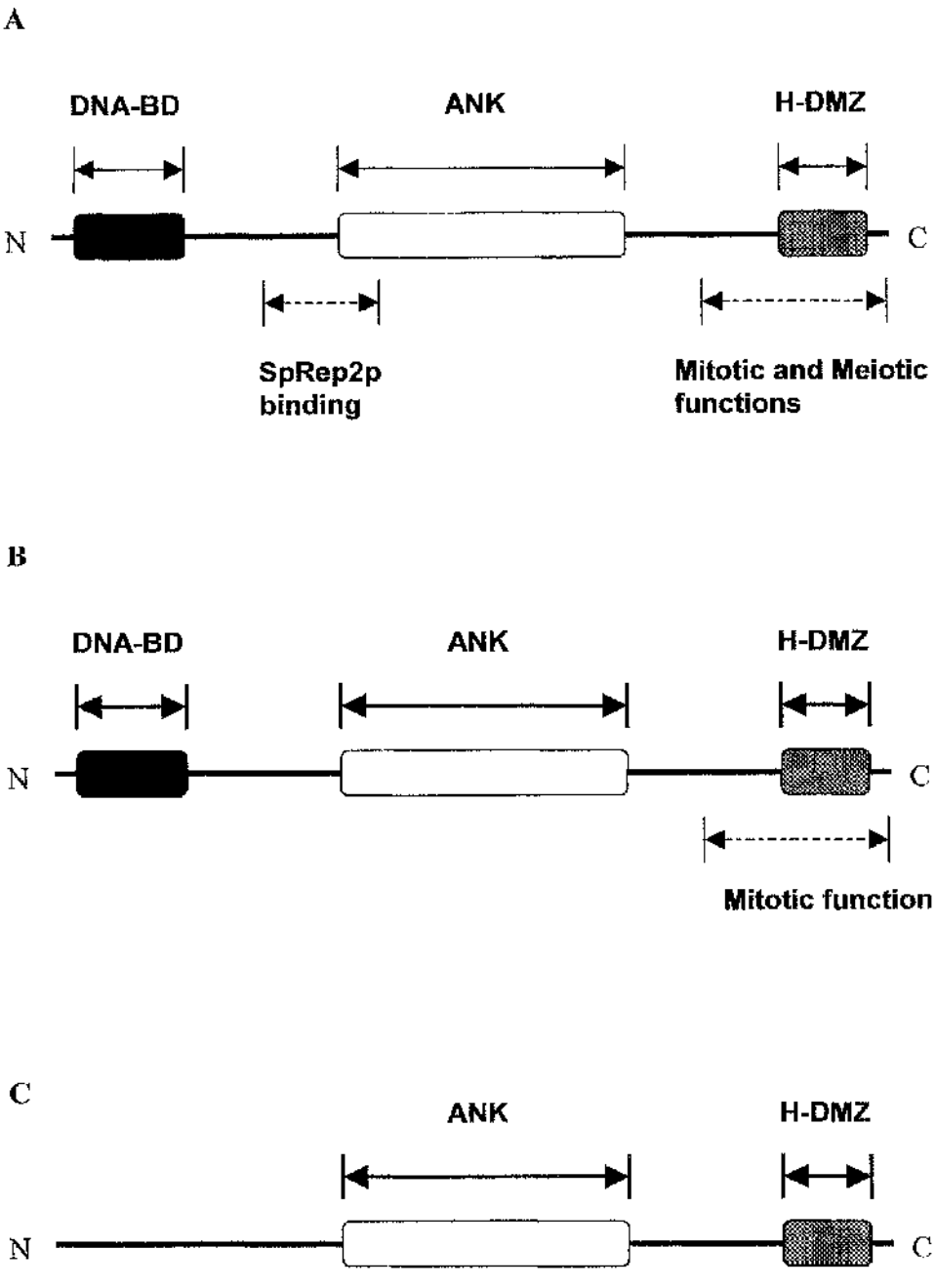
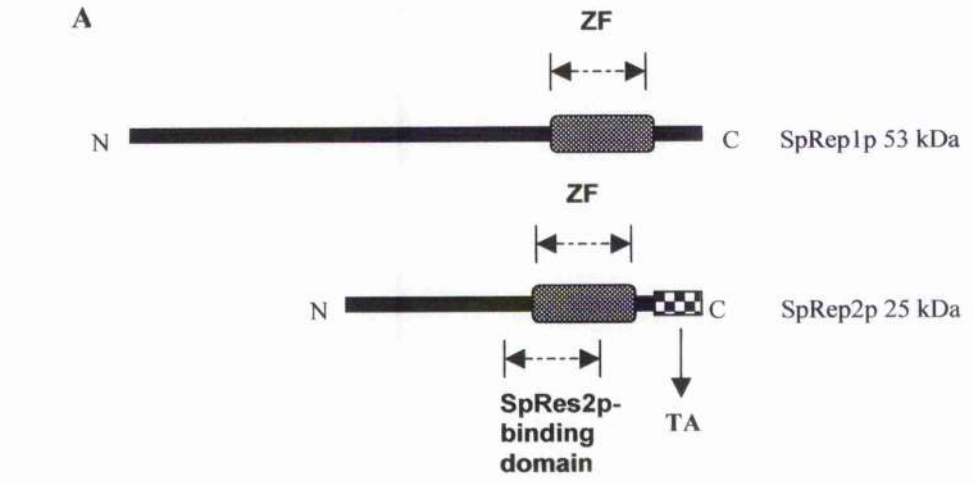


Figure 1.9 Domain architecture of SpRes2p, SpRes1p and SpCdc10p
A schematic representation of the major domains in **A** SpRes2p - 73 kDa, **B** SpRes1p - 72 kDa and **C** SpCdc10p - 85 kDa. DNA-BD = DNA-binding domain, ANK = Ankyrin repeat domain, H-DMZ = Heterodimerization domain (note that the N-terminal domain of SpCdc10p does not bind to DNA - adapted from Sturm and Okayama, 1996).

1.6.7 SpRep1p and SpRep2p

The *rep1*$^{+}$ gene encodes a 53 kDa protein containing a C terminal zinc finger motif, that is essential for its function (Sugiyama et al., 1994). Considerably smaller than SpRep1p, the 25 kDa SpRep2p protein also contains a C terminal zinc finger motif, sharing significant homology with SpRep1p in this region (Sugiyama et al., 1994; Nakashima et al., 1995; Figure 1.10). In contrast to SpRes1p and SpRes2p, which are both functional during mitosis, SpRes1p and SpRes2p have distinct meiotic and mitotic functions, respectively (Sugiyama et al., 1994; Nakashima et al., 1995). However, resembling the scenario for SpRes1p and SpRes2p, genetic suppressor analyses indicate that the *rep1*$^{+}$ and *rep2*$^{+}$ gene products also display some functional overlap.

Overexpression of either *rep1*$^{+}$ or *rep2*$^{+}$ can rescue the conditional lethality of the *cdc10-129* mutant, the property by which they were cloned (Sugiyama et al., 1994; Nakashima et al., 1995). Overexpression of *rep1*$^{+}$ rescues the cold sensitive lethality of cells deleted for *rep2*$^{+}$, demonstrating that SpRep1p can compensate for loss of SpRep2p mitotic function (Sugiyama et al., 1994). The *rep1*$^{+}$ gene is not expressed in the mitotic cycle; consequently, deletion of *rep1*$^{+}$ has no effect on mitotically dividing cells (Sugiyama et al., 1994). In the mitotic cycle Δ*rep2* cells (though viable at 30°C) are cold sensitive and arrest at START at ≤ 18°C. Conversely, Δ*rep2* cells show only slight meiotic defects, with the majority of cells successfully completing meiosis. In contrast, Δ*rep1* mutants are unable to initiate pre-meiotic DNA synthesis and, moreover, are deficient in induction of the *cdc22*$^{+}$ transcript (Sugiyama et al., 1994; Ding and Smith, 1998; Cunliffe et al., 2004). This behaviour is consistent with a suggested role for SpRep1p as an activator of DSC1-dependent transcription in the meiotic cycle. In agreement with this proposal is the demonstration that SpRep2p acts as a transcriptional activator subunit for SpRes2p (Nakashima et al., 1995; Tahara et al., 1998). In Δ*rep2* cells, DSC1-regulated genes are transcribed at lower levels compared to wild type, although the periodicity is maintained (Baum et al., 1997). Furthermore, SpRep2p contains centrally located SpRes2p binding and C-terminal activation domains, both of which are essential for its function (Nakashima et al., 1995; Tahara et al., 1998). Thus, SpRep1p and SpRep2p are likely to be activators of DSC1 in the meiotic and mitotic cycles, respectively.



B

```
           1        10        20        30        40        50        60        70
SpRep1p  MDSDRCLTDEISLNTLSSTFASDNNLRRKENFLKSRYPSKFDLNVSMLTKSDDVGKTPPSIFDSPSNPGF
SpRep2p  ..........................................................................

                80        90        100       110       120       130       140
SpRep1p  SRSHQMCSDKNKSPFLFDKIRDEPVSVDPVKLIINNRNKRKINRQKSRLQGLYRSDANGLQPLNSENVKM
SpRep2p  ..........................................................................

               150       160       170       180       190       200       210
SpRep1p  KKSTALSLTSSPLNSWKTDFKTPPKANVVCISLVIKGDGCASLLYEDLNQISNSCPEVAPNRQNALFSDE
SpRep2p  ..........................................................................

               220       230       240       250       260       270       280
SpRep1p  TLTTMFYSGVTEDEGSCNNLLQSSFGDDLDLGMQRSATWAPGYNYPSKFDSIPFASASPKIKAAFPFDPN
SpRep2p  ......................MHE.ADIPLSKPCLN..NP.PTYP.WSSPILSSHANPSLCDIVSSPSS

               290       300       310       320       330       340       350
SpRep1p  VYALNTNDGPVTSTDNNCDQLNTRMQSWNNGFNYSNLNGDIQYPAVTPKFFQQEGRALMVSDCNFGNEEI
SpRep2p  VSSFASSDD.FPFMNAYCLPI.....QQNHQFGSPVAASPNQQPLVESQNRRNVTYASLVIGKLGIAQLI

               360       370       380       390       400       410
SpRep1p  AYGGIPMRVCHSDS.TLCDARIAAKQALNGKRQYDTSTSKAELSTPPSKRRHIDDADLVFHSSPLLSRSR
SpRep2p  RQQTDPPQIHRKQDKGLMARVLSRSKKQEEH..ENHSSDARDAIKSALRRRMRRREGRVQKALRPTPN.

           420       430       440       450       460       470
SpRep1p  FICCYCTKPFLSISKLQEHPSSCSHVERLFGPAPNRLYDDGDGFLGSSPCSDW..
SpRep2p  LICSKCNTTFNHSTALMMHEATC......LNQVPFKLSD....FFVEDVIDDWLF
```

Figure 1.10

A. Domain architecture of SpRep1p and SpRep2p

Schematic representation of the major domains of SpRep1p and SpRep2p (adapted from Nakashima et al., 1995). TA = Transcriptional activation domain, ZF = Zinc finger motif. **B. Alignment of SpRep1p and SpRep2p**

Sequence alignment between the SpRep1p and SpRes2p proteins. The residue numbers are annotated for SpRep1p. Identical residues are shown as white characters boxed in red with similarity denoted by red characters boxed in white. This figure was generated using the ESPript program (Gouet et al., 1999).

1.6.8 Regulation of DSC1-dependent gene expression

Based largely on genetic evidence (Section 1.6.5), original models proposed that fission yeast contained two distinct yet functionally overlapping SpCdc10p-containing transcription factors, analogous to the situation in budding yeast (Miyamoto et al., 1994; Zhu et al., 1994). According to this model, a SpRes1p-SpCdc10p complex functioned during mitosis, whilst a SpRes2p-SpCdc10p complex acted in meiosis. However, substantial evidence accumulated arguing that both SpRes1p and SpRes2p functioned together with SpCdc10p in a single mitotic DSC1 complex.

Analysis of the DSC1 band-shift activity produced from fission yeast mitotic cell extracts showed it to contain at least SpRes1p, SpRes2p and SpCdc10p (Ayte et al., 1997; Zhu et al., 1997). Furthermore, SpRes1p and SpRes2p can heterodimerize *in vitro*, dependent on SpCdc10p (Zhu et al., 1997) and the DSC1 band-shift activity is lost in cells deleted for either *res1*$^+$ or res2$^+$, suggesting both subunits are required for the complex to form (Ayte et al., 1997; Zhu et al., 1997).

The existence of a SpRes1p-SpRes2p-SpCdc10p containing complex was confirmed by the demonstration that all three proteins were present in the *in vivo* DSC1 complex throughout the mitotic cycle (Whitehall et al., 1999). Therefore, a DSC1 complex composed of at least SpRes1p, SpRes2p and SpCdc10p, controls the mitotic G1-S phase-specific transcription programme in fission yeast (Whitehall et al., 1999). The stoichiometry of the complex is unknown, although it is most likely heterotetrameric consisting of one molecule each of SpRes1p and SpRes2p, each of which bind an SpCdc10p molecule (Ayte et al., 1997; Reymond et al., 1993; Figure 1.11).

The levels of SpCdc10p, SpRes1p and SpRes2p remain constant throughout the cell cycle and thus it is likely that DSC1 is regulated through post-translational modifications (Whitehall et al., 1999). In contrast to ScSwi6p (Section 1.5.8), SpCdc10p is predominantly nuclear throughout the cell cycle, and so regulation of sub-cellular localisation is unlikely to contribute to DSC1 regulation (Reymond et al., 1993; Wuarin et al., 2002). In budding yeast, ScCdc28p kinase activity has been shown to play a role in activating SBF/MBF-regulated transcription (Section 1.5.8; Wijnen et al., 2002). Similarly, in fission yeast the SpCdc2p kinase is required for passage of START (Nurse and Bisset, 1981). However, to date, there is no evidence for the direct involvement of SpCdc2p in activating DSC1-dependent transcription (Baum et al., 1997; Whitehall et al., 1999). The SpPas1p-SpPef1p kinase apparently plays a role in the activation of SpRes2p, by an unknown mechanism, although the SpPas1p-SpPef1p kinase may directly phosphorylate SpRes2p (Tanaka and Okayama, 2000).

Repression of DSC1-dependent transcription is mediated, at least in part, by SpCig2p-SpCdc2p kinase activity (Ayte et al., 2001). The SpCig2p-SpCdc2p kinase complex binds to SpRes2p and phosphorylates SpRes1p, causing repression of DSC1-dependent transcription,

although the mechanics of this are unknown (Ayte et al., 2001). Intriguingly, SpCig2p itself is a transcriptional target of DSC1. Thus a negative feedback loop appears to operate, whereby transcription of SpCig2p leads to increasing SpCig2p-SpCdc2p kinase activity that in turn down-regulates DSC1-dependent transcription (Ayte et al., 2001). This model of SpRes2p mediated repression is consistent with the expression pattern of MCB-regulated genes in cells either deleted for or overexpressing *res2*$^{+}$. Overexpression of SpRes2p leads to repression of DSC1-dependent transcription, whilst Δ*res2* cells show increased levels of transcription (Baum et al., 1997). In contrast, overexpression of SpRes1p leads to increased levels of DSC1-dependent transcription, whilst transcription is reduced in cells deleted for *res1*$^{+}$ (Baum et al., 1997).

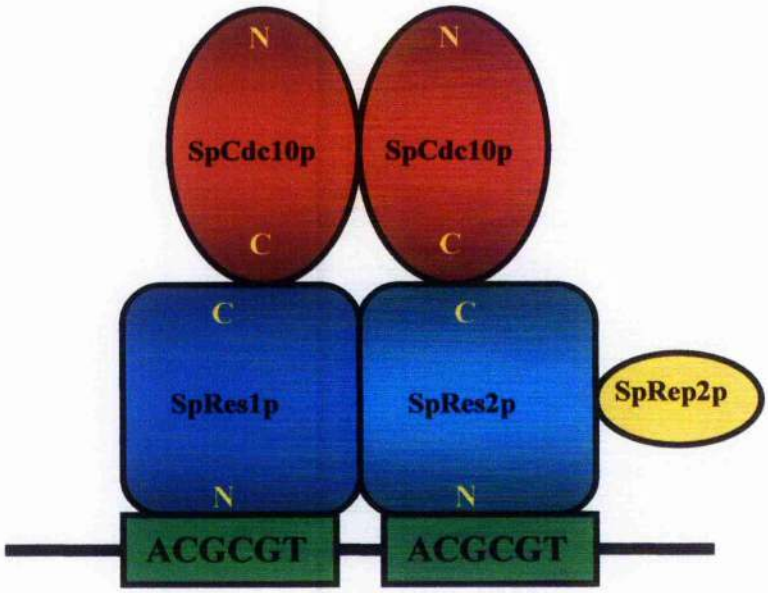


Figure 1.11 The DSC1 DNA-binding complex
A schematic representation of the fission yeast mitotic DSC1 DNA-binding complex bound to the MCB sequence element. The N and C-termini of each protein are indicated in yellow.

1.7 G1-S phase-specific transcription in mammalian cells

The conservation in both structure and function between the MBF/SBF and DSC1 systems of budding and fission yeasts, suggested that a homologous G1-S transcriptional control system also operates in higher eukaryotes.

In mammalian cells a G1-S phase-specific transcription programme is also required for passage of the Restriction point and entry into S phase, controlled by the E2F family of transcription factors (Dyson, 1998). Similar to yeasts, several of the genes that are activated by the E2F system encode functions required for DNA synthesis and replication (e.g. dihydrofolate reductase, thymidine kinase and the HsCdc6p, HsOrc1p and HsMcmp proteins) and key cell cycle regulators such as cyclin A, cyclin E and HsCdk1p (Ren et al., 2002). Furthermore, the TTTCGCGC consensus sequence element that is common to their promoters (and bound by the E2F proteins) resembles the MCB and SCB elements (Dyson, 1998). Moreover, the E2F proteins utilise a similar wHTH structural fold for DNA-binding (Zheng et al., 1999). Despite these tantalizing parallels, the E2F proteins display no significant amino acid sequence homology with any of the yeast proteins and so although they are responsible for regulating an analogous process in higher eukaryotes they are not true structural and functional homologues of the yeast factors.

1.7.1 The E2F and DP proteins

The functional E2F complex acts as a heterodimer comprising one subunit derived from the E2F family and one derived from the DP family. In mammalian cells, seven E2F (E2F-1 to E2F-7) and two DP (DP-1 and DP-2) proteins have been discovered to date (Trimarchi and Lees, 2001; DeBruin et al., 2003). The proteins they encode contain highly conserved DNA-binding and dimerisation domains. In addition, the C-termini of E2Fs 1-5 contain transactivation and pocket protein binding domains that are absent in E2F-6/7 and the DP proteins (Trimarchi and Lees, 2002; DeBruin et al., 2003). It is thought that the DP subunits activate transcription indirectly by enhancing the activity of the E2F subunit. Little is known about the specific properties of the individual E2Fs, although they have been categorised into three different subtypes based on both structural and functional properties (Dyson, 1998).

E2Fs 1-3: the activating E2Fs, these proteins activate transcription from E2F responsive genes, thereby driving proliferation. These proteins are under cell cycle control with levels peaking at G1-S boundary.

E2Fs-4 and -5: In contrast these proteins are constitutively expressed and are particularly prevalent in G0 cells. They appear to bind to promoters in an inactive conformation, associated

with pocket proteins and thus are believed to be responsible for transcriptional repression. Release of the pocket proteins resumes their transcriptional activation potential.

E2Fs-6 and -7: Both of these proteins act as transcriptional repressors. They each lack the C-terminal transactivation and pocket-binding domains and are therefore insensitive to the pRB family proteins. However, in contrast to E2F-6 (and all the other E2F and DP family members) E2F-7 contains two DNA-binding domains arranged in tandem and lacks the dimerisation domain, suggesting it can bind DNA in the absence of a DP partner (DeBruin et al., 2003; Logan et al., 2004). E2F-6 is proposed to act as a transcriptional repressor by binding to promoters, thereby preventing access by other 'active' E2F complexes and can recruit members of the polycomb chromatin re-modelling family to repress chromatin (Stevens and LaThangue, 2003).

1.7.2 Regulation of E2F-dependent gene expression

The molecular mechanisms by which E2F-dependent transcription is regulated are complex and not fully understood. The most prominent regulator of E2F-dependent transcription is the pRB pocket protein, which binds to and negatively regulates E2F. *RB* was the first tumour suppressor gene to be cloned and disruption in pRB function is involved in the majority of all human cancers (Weinberg, 1995).

The ability of pRB to bind E2F is dependent on its phosphorylation state, which is controlled by the actions of the Cdks, thereby linking E2F activity to the cell cycle. Each member of the pRB family contains multiple Cdk phosphorylation sites and the hyperphosphorylated forms have very low affinity for E2F. pRBs are thought to inactivate E2F by physically blocking the C-terminal transactivation domain. pRB can also recruit chromatin-remodelling machinery to promoters (e.g. SWI/SNF proteins, histone deacetylases, histone methyltransferases), thereby repressing transcription through changes in chromatin structure (Stevens and LaThangue, 2003).

During G0 and early G1 phases, pRB is hypophosphorylated and binds to E2F. In mammalian cells both positive and negative regulatory growth signals can act during G1, prior to the Restriction point (Sherr, 1996). Mitogenic growth factors initiate a signalling cascade, which ultimately leads to activation of the HsCdk4p/6p-cyclin D activities. The activated HsCdk4p/6p-cyclin D kinase complexes then phosphorylate pRB, causing its release from E2F. Consequently, E2F activates the expression of several genes whose functions are required for passage of the Restriction point and entry into S phase (Dyson, 1998).

Paradoxically, as a result of its temporal association with pRB family proteins, E2F bound to promoters can mediate both transcriptional activation and repression. Attempts to explain these conflicting roles have led to classification of E2F complexes in three distinct

groups (Dyson, 1998). Firstly, activator complexes correspond to those in which the activation domain of promoter bound E2F is free to stimulate transcription. Secondly, inhibited complexes represent those in which the activation domain of promoter bound E2F is blocked by pRB family proteins. Finally, an additional category corresponds to the repressor complexes in which pRB family proteins recruit and assemble chromatin-remodelling machinery, which represses chromatin.

Precisely how E2F activates transcription is unknown, although it can also interact with chromatin-remodelling machinery, presumably to reverse the effects of the pRB-mediated repression highlighted above (e.g. CBP, recruitment of HAT activity to promoter) and may also directly contact the general RNA polymerase II transcription machinery (Stevens and LaThangue, 2003).

In addition to pRB-mediated negative regulation, the decrease in E2F activity that is required for cells to exit S phase appears to be distinct from pRB-mediated phosphorylation. HsCdk2p-cyclin A can both bind to E2F-1 and phosphorylate DP-1, thereby inhibiting DNA-binding activity. As cyclin A is a transcriptional target of the E2F complex, this negative feedback is reminiscent of the action of SpCig2p upon DSC1 activity (Stevens and LaThangue, 2003; Section 1.6.8).

1.8 Aims of this study

The aims of this study were to produce a stable, reproducible and active source of recombinant protein of the components of the fission yeast DSC1 complex, in yields that would allow more extensive biophysical, biochemical and functional analyses to be undertaken.

The cloning of DSC1 component genes and their subsequent overexpression in *E. coli* has the potential to produce a reproducible source of protein in yields sufficient (5-10 mg l^{-1} culture) for this purpose, in contrast to their naturally low abundance in fission yeast. Chapter 3 describes the cloning and expression of DSC1 components as His-tagged fusion proteins in *E. coli*. Chapter 4 describes assaying the effects of His-tags on the activity of each component in fission yeast. In Chapter 5 DNA-binding activities are also assayed *in vitro*, following solubilisation of the recombinant proteins. In addition, GST-fusions of Res1p and Res2p are produced to further investigate the DNA-binding properties *in vitro*. Finally, in Chapter 6, attempts to purify the individual recombinant components of DSC1 and obtain preliminary structural data are described.

Chapter 2

Materials and methods

2.1 Molecular biology materials

2.1.1 Chemicals

All chemicals used were of the highest grade available commercially and distilled water was of Millipore-Q quality.

Specialised chemicals such as ampicillin, kanamycin, chloramphenicol, protease inhibitors (phenylmethanesulfonylfluoride – PMSF, chymostatin, pepstatin, antipain, leupeptin, aprotonin) and NZ amine, were from Sigma. Sephadex G-50, dATP (10μ Ci μl^{-1}) and dIdC (1 mg ml^{-1}) were supplied by Amersham Pharmacia Biotech. Dithiothreitol (DTT) was obtained from Melford laboratories Ltd, Suffolk.

2.1.2 Enzymes and kits

All restriction enzymes, T4 DNA ligase, calf intestinal alkaline phosphatase, dNTPs, *Taq* DNA polymerase and T4 polynucleotide kinase were obtained from Promega. $Vent_R$ DNA polymerase and Mung Bean Nuclease were obtained from New England Biolabs. The QIAprep® Spin Miniprep Kit and the QIAquick® Gel Extraction Kit were supplied by QIAGEN.

2.1.3 Molecular weight markers

DNA molecular weight marker X (0.07-12.2 kbp) was supplied by Roche.

2.1.4 Oligonucleotides

Oligonucleotides were designed in the laboratory as required and synthesised by MWG-AG Biotech or DNA Technology A/S. All oligonucleotides used in this thesis are listed in Appendix II. The annotation GO refers to the Glasgow lab oligo collection number.

2.1.5 Bacterial media

All strains of *Escherichia coli* were grown in Luria Broth (LB; 10 g Bacto-tryptone, 5 g Bacto yeast extract and 10 g NaCl per litre, pH 7.5) or NZY^+ broth (5 g NaCl, 2 g $MgSO_4.7H_2O$, 5 g yeast extract, 10 g NZ amine per litre, pH 7.5). LB plates were made by adding 7.5 g bacto agar to 500 ml LB. All media were autoclaved before use and supplemented with ampicillin (50 μg ml^{-1}), kanamycin (30 μg ml^{-1}) or chloramphenicol (34 μg ml^{-1}) where appropriate.

2.1.6 Bacterial strains

All bacterial strains used in this thesis are listed in Appendix I. The annotation GB refers to the Glasgow lab bacteria collection number.

***E. coli* DH5α:** a recombination deficient strain used for the propagation and storage of plasmid DNA. **Genotype:** *sup*E44 Δ*lac*U169 (ϕ80*lac*ZΔM15) *hsd*R17 *rec*A1 *end*A1 *hyr*A96 *thi*1 *rel*A1.
***E. coli* BL21 CodonPlus(DE3)-RIL (Stratagene):** an all-purpose strain for high-level protein expression and easy induction. **Genotype:** F^- *omp*T *hsdS*(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *end*A Hte[*arg*U *ile*Y *leu*W Cam^r].

2.1.7 Bacterial plasmid vectors

The pET vectors were purchased from Novagen, pCR2.1 was supplied by Invitrogen and pGEX-KG was obtained from a laboratory stock (Appendix III).

pET-28c: for expression of His-tagged recombinant proteins; the six histidine residues are linked to the N-terminus of the cloned gene (kanamycin resistant).

pET-14b: for expression of His-tagged recombinant proteins; the six histidine residues are linked to the N-terminus of the cloned gene (ampicillin resistant).

pCR2.1: for one-step cloning of a polymerase chain reaction product with polydeoxyadenosine 3' overhangs (ampicillin resistant).

pGEX-KG: for expression of GST-tagged recombinant proteins; the GST tag is linked to the N-terminus of the cloned gene (ampicillin resistant; Guan and Dixon, 1991).

2.2 Molecular biology methods

2.2.1 Polymerase chain reaction (PCR)

Polymerase chain reactions were performed using either *Taq* DNA polymerase or $Vent_R$ DNA polymerase.

2.2.2 PCR using *Taq* DNA polymerase

Standard PCR reactions were carried out in an MWG-Biotech Primus Thermal Cycler in 0.5ml thin walled PCR tubes with *Taq* DNA polymerase. A typical 100 µl reaction volume contained 10 x *Taq* buffer at a final concentration of 1 x, 200 µM each of dNTPs, 100 ng each of the appropriate primers, 50 ng of DNA template and 2 U of *Taq* DNA polymerase. Annealing temperatures for each reaction were determined empirically.

Typical cycle parameters were as follows:

30 cycles {
- Denaturation: 95°C for 2 min
- Denaturation: 95°C for 15 sec
- Annealing: 50°C for 30 sec (temperature adapted as appropriate for primers)
- Extension: 72°C for 1 min (time adapted as appropriate-1 min per kb of product)

Extension: 72°C for 5 min

2.2.3 PCR using $Vent_R$ DNA polymerase

PCR was performed as above with the following modifications. A typical 100 µl reaction volume contained 10 x reaction buffer (100 mM KCl, 100 mM $(NH_4)_2SO_4$, 200 mM Tris-HCl (pH 8.8), 20 mM $MgSO_4$, 1% Triton X-100) at a final concentration of 1 x, 200 µM of each dNTP, 100 ng each of the appropriate primers, 50 ng of DNA template and 2 U of $Vent_R$ DNA polymerase. Typical cycle parameters were as follows:

20 cycles {
- 95°C for 1 min
- 95°C for 30 sec
- 58°C for 30 sec (temperature adapted as appropriate for primers)
- 68°C for 2 min (time adapted as appropriate-1 min per kb of product)

68°C for 3 min

The yield and purity of the PCR product obtained was analysed by agarose gel electrophoresis (2% (w/v) agarose) and stored at -20°C before and after agarose gel purification.

2.2.4 Agarose gel electrophoresis

The appropriate amount (1 g 100 ml^{-1}) of agarose was dissolved in 1 x TBE (45 mM Tris-borate, 1 mM EDTA) and 2-3 µl ethidium bromide added to give slab agarose gels of the required percentage, routinely 1-2%. Samples for analysis were diluted 5 fold by the addition of 6 x loading buffer (0.25% (w/v) bromophenol blue, 0.25% (w/v) xylene cyanol FF, 0.15% (w/v) Ficoll) before being loaded on the agarose gel. These were run at 100 V in 1 x TBE for between 40 min - 1 h until the dye front was about 1 cm from the bottom of the gel. Gels were then viewed using a UV transilluminator and photographed using E. A. S. Y imaging software.

2.2.5 Purification of DNA from bacterial cultures

DNA was purified from bacterial cultures using the QIAprep® Spin Miniprep Kit supplied by QIAGEN. The kit was used as per the manufacturer's instructions. Briefly, a 5 ml culture containing the appropriate antibiotic was inoculated with a single colony from an LB antibiotic plate. This was incubated with shaking at 37°C for not more than 16 h. Cells were pelleted at 10,000 rpm for 5 min and resuspended in 250 µl Buffer P1. Cells were lysed by the addition of 250 µl Buffer P2 and incubated for 5 min. 350 µl of Buffer N3 was added and the mixture centrifuged at 10,000 rpm for 10 min in a benchtop centrifuge to pellet the cell debris. Supernatant was applied to a QIAprep® column and centrifuged briefly to allow the DNA to bind to the membrane of the spin column. The column was washed twice, firstly with 0.5 ml Buffer PB and then 0.75 ml Buffer PE, before the DNA was eluted in 50 µl nuclease-free water. The yield and purity of DNA obtained was analysed by agarose gel electrophoresis.

2.2.6 Extraction of DNA from an agarose gel

DNA was purified from agarose gels using the QIAquick® gel extraction kit (QIAGEN) as described in the manufacturer's instructions.

Briefly, DNA was excised from an agarose gel using a sterile scalpel blade. The agarose slice was solubilized at 50°C in the appropriate volume of Buffer QG (as supplied). This mixture was applied to a spin column and centrifuged for 1 min to allow the DNA to bind the column. The column was washed with 750 µl of Buffer PE containing 80% (v/v) ethanol. Residual ethanol was removed by centrifuging the column for a further 1 min, before eluting the DNA in 50 µl Buffer EB (1 mM Tris-HCl, pH 8.5) or nuclease-free water.

The yield and purity of DNA obtained was analysed by agarose gel electrophoresis.

2.2.7 Restriction digestion

Plasmid DNA and PCR products were routinely digested, prior to ligation, as follows: 5 µl DNA was digested with 1 µl enzyme for a single digest (or 1 µl of each enzyme in double digests), together with 2 µl of the appropriate enzyme buffer in a final volume of 20 µl. This was incubated at 37°C for 30 min or as otherwise instructed for the enzyme (exceptions were *Sma*I, which required incubation at 25°C for 30 min and *Nco*I in which a second aliquot of enzyme was added after 15 min due to the short half life of the enzyme).

2.2.8 Dephosphorylation of digested plasmid

Digested vector was dephosphorylated by the addition of 1 U of calf intestinal alkaline phosphatase and incubated at 37°C for a further 30 min. This was routinely carried out to prevent self-ligation of digested vector.

2.2.9 Ligations

Purified plasmid DNA was ligated with purified insert DNA at various ratios (1:3, 1:5, 1:7). A typical ligation reaction was set up as follows: 1 µl of digested, dephosphorylated and gel purified vector was mixed with 5 µl of digested and gel purified insert, 10 x T4 DNA ligase buffer (at a final concentration of 1 x) and 1 U of T4 DNA ligase. Routinely, ligations were incubated overnight at 4°C before being transformed into *E.coli* DH5α competent cells following the standard protocol (Section 2.2.13). The resulting colonies were then screened by restriction digestion analysis (Section 2.2.7) to confirm the presence of insert DNA.

2.2.10 Cloning of blunt ended PCR products

The TA Cloning® Kit (Invitrogen) was used to clone blunt-ended PCR products as described in the manufacturer's instructions.

2.2.11 Site-directed mutagenesis by overlap extension

The plasmid GB 130 (Appendix I), containing a ~2,300 bp fragment representing the recombinant *cdc10*^{1} gene, was used as a template in the PCR reactions, which were adapted from a method described in Sambrook and Russell (2001). Four primers (GO 546, GO 459, GO 547 and GO 548 - Appendix II) were used to introduce a site-specific mutation by overlap extension. The primers GO 546 and GO 459 were complementary to either end of the *cdc10*$^{+}$ gene and allowed introduction of *Nde*I and *Bam*HI sites, respectively, to the 5' and 3' ends of the PCR fragment. The primers GO 547 and GO 548 were complementary to the *cdc10*$^{+}$ gene and were designed to introduce a single base pair G to C substitution at position 426 relative to the ATG. This introduces a silent mutation Tyrosine (TAG) to Tyrosine (TAC) and removes the artificial internal stop codon. The primer pairs GO 547 and GO 548 were necessarily complementary to each other and could not be used in the same PCR reaction. Therefore, mutation had to be carried out in two stages.

Stage 1: Primer pairs GO 546/GO 548 and GO 459/GO 547 were used in separate PCR reactions to produce 2 different sized fragments. The PCR reaction in this step was carried out as described in Section 2.2.3.

Stage 2: The 2 fragments, each containing the G to C substitution at 426, were gel purified and joined together to generate the full-length cDNA in a fusion PCR reaction using primers GO 546 and GO 459.

The parameters of the fusion PCR reaction were as follows:

94°C for 1 min

10 cycles { 94°C for 10 sec; 50°C for 1 min; 68°C for 3.5 min }

10 cycles { 94°C for 10 sec; 58°C for 1 min; 68°C for 3.5 min }

20 cycles { 94°C for 1 min; 58°C for 1 min; 68°C for 3.5 min }

68°C for 7 min

2.2.12 Production of competent bacterial cells

Competent cells were made using the rubidium chloride method.

The appropriate bacterial *E. coli* strain was streaked onto a minimal LB plate and grown overnight at 37°C. A single colony was used to inoculate an overnight 5 ml LB culture. This was subcultured into 100 ml LB and grown at 37°C with shaking until the culture reached an optical density (OD) of 0.48 at 550 nm. The culture was then chilled on ice for 5 min, before centrifugation at 3000 rpm for 10 min at 4°C. The pellet was resuspended in 40 ml Buffer 1 (100 mM rubidium chloride, 10 mM calcium chloride, 50 mM manganese chloride, 15% (v/v) glycerol in 30 mM potassium acetate, pH 5.8). Cells were centrifuged as before and the pellet resuspended in 4 ml Buffer 2 (10 mM rubidium chloride, 75 mM calcium chloride, 15% (v/v) glycerol in 10 mM 3-[N-morpholino] propanesulfonic acid (MOPS), pH 6.5). The cells were divided into 100 µl aliquots and stored at -70°C.

2.2.13 Transformation of competent bacteria

In electro-transformation a BIORAD *E. coli* pulser was used with a method adapted from the BIORAD manual.

E. coli DH5α cells were routinely used for the propagation and harvesting of recombinant plasmid DNA, and were transformed using the following method. To 50 µl of competent bacteria 1-2 µl of the appropriate plasmid DNA was added. The cell-DNA suspension was transferred to a pre-chilled 0.2 cm electroporation cuvette and pulsed at 2.5 kV. 1 ml of SOC (2% Bacto tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) was added to the cells, which were then transferred to microfuge tubes and incubated at 37°C with shaking for 45 min. This mix was plated on an LB-agar plate containing the appropriate antibiotic and incubated overnight at 37°C.

E. coli BL21 CodonPlus(DE3)-RIL cells were routinely used for the expression of the desired recombinant proteins and transformed using the following method. To 50 µl of competent bacteria, 1-10 ng of DNA was added in a pre-chilled microfuge tube. The mixture was chilled for 15 min before heat shocking at 42°C for 90 sec and returned to ice. After 2 min 450 µl of SOC was added and this was incubated at 37°C with shaking for 45 min. This mix was plated on an LB-agar plate containing the appropriate antibiotic and incubated overnight at 37°C.

2.2.14 Electophoretic mobility shift assay (EMSA)

Electophoretic mobility shift assay (EMSA) was as described by Ng et al (2001). This technique is also commonly known as the band-shift assay and is referred to using both names in this thesis.

2.2.15 Probe preparation

DNA fragments were 5' end-labelled using T4 polynucleotide kinase (PNK). To 5-10 µg of DNA in 5 µl of dH_2O on ice was added 10 U of T4 PNK in 1 x T4 PNK buffer, followed by the addition of 10 µCi of [γ-^{32}P] dATP (10µ Ci µl^{-1}) to give a total volume of 8 µl. The reaction was incubated at 37°C for 1 h, and the DNA was purified using a Sephadex G-50 column. Sephadex G-50 was prepared by adding two volumes of TE and autoclaving. The plunger was removed from a 1 ml syringe (Plastipak) and a small wad of siliconized glass wool was used to plug the end before the end was placed inside a microfuge tube. Sephadex G-50 was added to the syringe; the syringe and microfuge tube were placed inside a 50 ml centrifuge tube and then centrifuged at 3000 rpm for 5 min to remove the TE. This was repeated until ~0.8 ml Sephadex G-50 remained in the syringe. A fresh microfuge tube was then placed at the bottom of the syringe and the radiolabelled probe added to the Sephadex G-50 column. This was spun again at 3000 rpm for 5 min and the syringe was monitored with a Geiger counter to confirm that unincorporated nucleotides had been removed from the probe. The purified probe was transferred to a fresh screw-top microfuge tube and stored at -70°C.

2.2.16 EMSA assay

Each mobility shift assay was performed by adding 10 µl of sample buffer (1 M Tris-HCl pH 7.5, 1 M KCl, 50% glycerol, 100 mM DTT, 100 mM protease inhibitors: chymostatin, pepstatin, antipain, leupeptin, aprotonin, 100 mM PMSF, 1 M $MgCl_2$) to 20 µg of protein (or as otherwise indicated), 1 µg of dIdC (1 mg ml^{-1}) and 1-2 µl of labelled probe. If required, non-specific or specific competitor DNA was also added. The samples were incubated for 5 min on ice before the addition of each reagent. Analysis of the formation of protein-DNA complexes was achieved by electrophoresis of samples on a 10% acrylamide gel in 1 x TBE buffer (40 mM Tris-borate, 1 mM EDTA pH 8) for 1.5-2 h at 180V. The gel was dried for 1.5 h at 80°C and exposed to autoradiography film at -70°C.

2.3 Protein biochemistry materials

2.3.1 Chemicals

Ultra pure imidazole and zinc chloride were purchased from MERCK, BDH. Triton X-100 was bought from Fisons, Loughborough. Isopropyl β-thiogalactopyranoside (IPTG) was obtained from Melford Laboratories Ltd, Suffolk. Glutathione Sepharose 4B was supplied from Amersham Pharmacia Biotech. Glutathione (reduced), NaF and acrylamide:bisacrylamide solution were supplied by Sigma. Bradfords Reagent (BIO-RAD protein assay reagent) was purchased from BIO-RAD.

2.3.2 Molecular weight markers

The relative molecular mass (subunit M_r) of proteins separated by SDS-PAGE was determined by comparison with low M_r markers (Amersham Pharmacia Biotech, Low molecular weight calibration kit for electrophoresis).

2.3.3 Phototgraphic materials

Nitrocellulose and Hyperfilm were purchased from Amersham Pharmacia Biotech. The X-Omat 100 processor was supplied by Kodak.

2.4 Protein biochemistry methods

2.4.1 Growth of bacterial cultures for protein induction

A single colony of transformed BL21 CodonPlus(DE3)-RIL cells was picked from an LB-agar antibiotic plate and grown at 37°C with shaking in 5 ml growth media plus the appropriate

antibiotic for 16 h. A 1 ml aliquot was subcultured into 50 ml growth media plus antibiotic and incubated at 37°C with shaking until the OD_{600} was between 0.5-1.0. IPTG, at a final concentration of 0.2 mM-1 mM (0.2 mM for 22°C and 15°C inductions or 1 mM for 30°C and 37°C inductions), was then added and the cultures induced at the required temperature for either 3 h (30°C and 37°C inductions) or 16 h (22°C and 15°C inductions). During 3 h inductions samples were taken at zero time and then at 1 h intervals to check for overexpression. During 16 h inductions samples were taken at zero time and at 16 h. All samples were pelleted by centrifugation at 13,000 rpm for 1 min and the pellets resuspended in Laemmli sample buffer (2% (w/v) SDS, 10% (w/v) sucrose, 62.5 mM Tris-HCl, pH 6.8, Pyronin Y dye) using 10 µl per 0.1 OD_{600} units. Cells were harvested by centrifugation at 3,000 rpm for 15 min and the supernatant discarded. The pellet was then resuspended in 3 ml of the appropriate buffer for purification and stored at -20°C.

2.4.2 Large-scale protein induction

A single colony was picked from an LB-agar antibiotic plate and grown at 37°C with shaking in 10 ml growth media plus antibiotic for 16 h. This was then subcultured into 500 ml growth media plus antibiotic and grown at 37°C with shaking until the OD_{600} was between 0.5-1.0. Protein expression was induced by the addition of 0.2-1 mM IPTG (Section 2.4.1). Induction was again carried out for either 3 h or 16 h depending on the temperature used. Samples at zero time and at the end of induction were kept for analysis on SDS-PAGE to check that expression had occurred as before. Cells were harvested by centrifugation at 8,250 rpm for 15 min in a Beckman J2-21 centrifuge using a JA-14 rotor and the supernatant discarded. The pellet was then resuspended in 20 ml of the appropriate buffer for purification and stored at -20°C.

2.4.3 Dialysis of protein samples

Visking tubing was prepared by boiling in 10 mM sodium bicarbonate, pH 8.0, 1 mM EDTA for 10 min. This was then rinsed in distilled water before being stored in 100% (v/v) ethanol. Tubing was thoroughly rinsed in distilled water before use.

Dialysis of protein took place at 4°C for several hours using multiple changes of dialysis buffer.

2.4.4 Concentration of protein samples

After dialysis, protein samples were concentrated by centrifugation in a Centricon Plus-20 centrifugal concentrator filter (Amicon), with cut off 30,000 Da, according to the manufacturer's instructions.

2.4.5 Determination of protein concentration

The method of Bradford (1976) was routinely employed. A standard curve was produced using known concentrations of BSA. The absorbance of the unknown samples were taken then measured at 595 nm and their concentration extrapolated from the standard curve.

2.4.6 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

The solutions required for SDS-PAGE were as follows:

Acrylamide solution

29.2% (w/v) acrylamide/0.8% (w/v) bis-acrylamide.

Resolving gel buffer

10% (w/v) acrylamide, 0.5 M Tris-HCl, pH 8.8, 0.1% (w/v) sodium dodecyl sulphate (SDS), 0.1% (w/v) ammonium persulphate, 0.1% (v/v) N, N, N^1, N^1-tetramethlethylene diamine (TEMED).

Stacking gel buffer

4.5% (w/v) acrylamide, 0.06 M Tris-HCl, pH 6.8, 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulphate, 0.1% (v/v) TEMED.

Running buffer

25 mM Tris-HCl, pH 8.3, 0.25 M glycine, 1% (w/v) SDS.

Proteins were resolved under denaturing conditions using the method of Laemmli (1970). Samples for analysis on SDS-PAGE were resuspended in Laemmli sample buffer to which 1 M DTT was added (at a final concentration of 150 mM) prior to boiling for 5 min to ensure that all proteins were denatured. A 10 µl aliquot of each sample was loaded on the gel along with 10 µl of a low molecular mass marker (Section 2.3.2). Gels were run using the Biorad Mini-Protean gel kit system at 400 V for 1 h or until the dye front was approximately 1 cm from the bottom of the gel. Gels were then stained in 0.1% (w/v) Coomassie Brilliant Blue, 10% (v/v) acetic acid, 50% (v/v) methanol for 1 h and destained in 10% (v/v) acetic acid, 10% (v/v) methanol overnight.

2.4.7 Immunoblotting using ECL ™ (enhanced chemiluminessence)

Anti-PentaHis-HRP conjugate antibody was supplied by QIAGEN and Anti-GST-HRP conjugate antibody was supplied by Amersham Pharmacia Biotech. Immunoblotting protocols were carried out as detailed in the manufacturer's instructions.

2.4.8 Solutions used in immunoblotting

Transfer buffer (10 x) per litre

25 mM Tris-HCl, pH 7.2, 192 mM glycine, 0.02% (w/v) SDS, 20% (v/v) methanol.

2.4.9 Solutions for Anti-PentaHis-HRP conjugate detection

Blocking buffer: 10 mM Tris-HCl pH 7.5, 0.1% (w/v) Blocking reagent, 0.1% (v/v) Tween-20.

TBS buffer: 10 mM Tris-HCl pH 7.5.

TBS-Tween/Triton buffer: 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 0.05% (v/v) Tween-20, 0.2% (v/v) Triton X-100.

Anti-PentaHis-HRP conjugate antibody: 20 mM Tris-HCl pH 7.5, 1% (w/v) Non-fat milk, 0.1% (v/v) Tween-20, 1:2500 dilution of Anti-PentaHis-HRP conjugate antibody.

2.4.10 Solutions for Anti-GST-HRP conjugate detection

Blocking buffer: 20 mM Tris-HCl pH 7.5, 5% (w/v) Blocking reagent, 1% (v/v) Tween-20.

TBS buffer: as above (Section 2.4.8).

TBS-Tween/Triton buffer: as above (Section 2.4.8).

Anti-GST-HRP conjugate antibody: 20 mM Tris-HCl pH 7.5, 0.1% (v/v) Tween-20, 1:5000 dilution of Anti-GST-HRP conjugate antibody.

2.4.11 Immunoblotting protocol

SDS-PAGE analysis was performed as described in Section 2.4.6. Proteins were then electrophoretically transferred to nitrocellulose using the Biorad Mini-Protean gel kit system at 50 mA for 2 h in 1 x transfer buffer. Staining of the nitrocellulose with the non-fixative dye Ponceau S was employed to check the efficiency of protein transfer. This was washed off in distilled water and the non-specific binding sites were blocked by immersing the nitrocellulose in blocking buffer for 1 h at room temperature with shaking. The membrane was washed twice for 10 min in TBS-Tween/Triton buffer then incubated with diluted HRP conjugate antibody solution at room temperature for 1 h. The membrane was again washed in TBS-Tween/Triton

buffer for 10 min and given a final wash in TBS buffer before detection. The detection step was carried out as described in the Amersham protocol for ECL detection. In the dark room, autoradiography film was placed onto the membrane and exposed for 30 sec initially before developing the film using a Kodak X-omat 100 processor.

2.4.12 Solubilisation of expressed fusion proteins

Fusion proteins were expressed as described previously (Section 2.4.1-2.4.2). The pellets were then resuspended in the appropriate volume of lysis buffer and disrupted by 3 passes through an automatic French pressure cell (pre-cooled on ice) at 750 or 950 psi for small or large-scale preparations, respectively. Protease inhibitors were routinely added to prevent degradation of protein. A 100 µl aliquot of this whole cell extract sample was retained before the remainder was centrifuged at 4°C in a Beckman J2-21 centrifuge using a JA-17 or JA-14 rotor for small or large preparations, respectively. 100 µl aliquots of the supernatant and pellet, resuspended in the original starting volume of lysis buffer, were retained. An equal volume of Laemmli sample buffer was added to these samples and the solubility of the recombinant protein viewed by SDS-PAGE.

2.4.13 Purification of GST-tagged proteins

GST-tagged proteins were purified using the BioCAD® SPRINT™, Perfusion Chromatography® System (PE Biosystems). A GSTrapFF 5 ml (Glutathione sepharose) column was routinely used.

2.4.14 Preparation of bacterial cell extracts

GST fusion proteins were expressed as described (Section 2.4.1-2.4.2).

The pellet from a 500 ml bacterial cell culture was resuspended in 20 ml 1 x PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3) and the cells disrupted under high pressure using a French Pressure cell at 950 psi. Typically 3 passes were made and protease inhibitors added to prevent degradation of protein. The supernatant was clarified by centrifugation at 10,000 rpm for 15 min at 4°C in a Beckman J2-21 centrifuge using a JA-14 rotor.

2.4.15 Column purification of GST-tagged proteins

The column was prepared by equilibrating with 5 column volumes of binding buffer (1 x PBS) prior to use. Clarified supernatant was loaded onto the column in 5 ml increments. After extensive washing of the column in binding buffer, bound GST-tagged protein was eluted in elution buffer (10 mM reduced glutathione in 50 mM Tris-HCl, pH 8.0) and 2 ml fractions

collected. Elution of the protein was monitored by measuring the absorbance at 280 nm. Aliquots of the appropriate fractions were then mixed with an equal volume of Laemmli sample buffer and analyzed using SDS-PAGE.

2.4.16 Purification of His-tagged proteins

His-tagged proteins were purified using the BioCAD® SPRINT™, Perfusion Chromatography® System (PE Biosystems). A POROS® MC (metal chelate) column (4.6 mm/100 mm) was routinely used.

2.4.17 Preparation of bacterial cell extracts

His-tagged fusion proteins were expressed as described (Section 2.4.1-2.4.2). The pellet from a 500 ml bacterial cell culture was resuspended in 20 ml starting buffer (100 mM NaCl, 0.5 mM imidazole in 50mM KH_2PO_4, pH 7.5) and the cells disrupted under high pressure using a French Pressure cell at 950 psi. Typically 3 passes were made and protease inhibitors added to prevent degradation of protein. The supernatant was clarified by centrifugation at 10,000 rpm for 15 min at 4°C in a Beckman J2-21 centrifuge using a JA-14 rotor.

2.4.18 Column purification of His-tagged proteins

The column was prepared by loading the imidodiacetate binding sites with zinc ions (0.1 M $ZnCl_2$, pH 4.5-5) for 25 column volumes and then washing with distilled water followed by 0.5 M NaCl to remove any excess metal ions. The column was then washed with 5 column volumes of elution buffer (100 mM NaCl, 500 mM imidazole in 50 mM KH_2PO_4, pH 7.5). Finally the column was equilibrated with starting buffer (100 mM NaCl, 0.5 mM imidazole in 50 mM KH_2PO_4, pH 7.5) before use.

Clarified supernatant was loaded onto the column in 5 ml increments. After extensive washing of the column in starting buffer bound His-tagged protein was eluted in a linear imidazole gradient (0.5 mM – 500 mM) and 2 ml fractions collected. Elution of protein was monitored by measuring the absorbance at 280 nm. After all the protein eluted the column was regenerated by washing with 15 column volumes of stripping buffer (50 mM EDTA, 1 M NaCl). Aliquots of the appropriate fractions were then mixed with an equal volume of Laemmli sample buffer and analyzed using SDS-PAGE.

2.4.19 Solubilization of insoluble proteins from inclusion bodies

The following protocol was adapted from the Protein Folding Kit (Novagen).

After induction in *E. coli* and overexpression of the desired protein (Section 2.4.1), the cells were pelleted as normal and resuspended in 0.1-culture volume of Inclusion body wash

buffer (10 mM EDTA, 1% (v/v) Triton X-100 in 20 mM Tris-HCl, pH 7.5). Cells were disrupted by passage through a French press (Section 2.4.12), and insoluble material pelleted by centrifugation at 10,000 rpm for 10 min at 4°C. The pellet was then washed a further twice in Inclusion body wash buffer.

Inclusion bodies were resuspended in 50 mM KH_2PO_4, pH 7.5 (or 150 mM NaF, pH 7.5 for Circular dichroism experiments) at a concentration of 10-20 mg ml^{-1} and supplemented with the appropriate volume of 30% (v/v) N-laurylsarcosine to give the desired concentration of detergent in the buffer, typically 0.1-3% (complete solubilisation was generally achieved with 0.2% (v/v) detergent). This was incubated at room temperature for 30 min with agitation before clarifying any insoluble material by centrifugation at 10,000 rpm for 15 min at 4°C. Samples of the supernatant were taken for analysis using SDS-PAGE.

2.4.20 Mass spectrometric data

Mass spectra were recorded in collaboration with Dr. A. Pitt, Sir Henry Wellcome Functional Genomics Facility, University of Glasgow.

Sample preparation

Gel bands were broken into approximately 1 mm cubes and then destained and dehydrated by sequential washing with 300 µl of: 25 mM ammonium bicarbonate, 50% 25 mM ammonium bicarbonate in acetonitrile and 100% acetonitrile, followed by drying in a centrifugal evaporator (Univap, Uniscience). Gel slices were re-hydrated in 20 µl of 20 ng ml^{-1} trypsin (in 25 mM ammonium bicarbonate) for 15 min. Additional 25 mM bicarbonate was then added to ensure the gel slice was covered and the digest incubated at 37°C overnight. The supernatant was removed, acidified by the addition on 1 µl of 5% formic acid and used directly for mass spectrometry or frozen at -20°C until required. 10 µl aliquots of the samples was subjected to nLC-MSMS on a QStar Pulsar i electrospray mass spectrometer fitted with a nanospray source (Protana), using a 20 micron i.d., 10 micron orifice distally coated electrospray tip mounted in a ProADP2 adaptor (New Objective) connected to a nanoflow LC system (LC Packings) with the minimum length of 20 micron i.d. fused silica capillary possible. nLC was performed using a 0.3 x 5 mm reversed phase trap (PepMap C18, Dionex) and a 75 micron x 15 cm PepMap C18 reversed phase column (Dionex) using standard methodology, with a loading pump flow of 30 µl min^{-1} and a main flow of 200 nl min^{-1}. Peptides were trapped on the C18 trap and desalted for 5 min before being separated using a 5-40% acetonitrile gradient over 15 min. All solutions contained 0.5% formic acid. Mass spectrometric analysis was performed in IDA mode (AnalystQS software, Applied Biosystems), selecting the 4 most intense ions for MSMS analysis. A survey scan of 400-1500 Da was collected for 3 sec followed by 3 sec MSMS scans

of 50-2000 Da using the standard rolling collision energy settings. IDA was triggered for ions with charge states 2-4 above a threshold of 10 counts. Masses were then added to the exclusion list for 3 min. Peaks were extracted using the Mascot script (BioAnalyst, Applied Biosystems) and automatically exported to the Mascot (Matrix Science) search engine. Data was searched against the database using the MASCOT Daemon, with a peptide tolerance of 1.0 Da, an MSMS ion tolerance of 0.5 Da allowing for 1 missed cleave and variable methionine oxidation.

2.4.21 Circular dichroism

Circular dichroism (CD) spectra were recorded at 20°C on a JASCO J-810 spectropolarimeter (Jasco UK). All spectra were recorded in collaboration with Dr. S. Kelly and Mr. T. Jess in the Scottish Circular Dichroism Facility, University of Glasgow. For CD experiments, purified protein samples were routinely prepared in 150 mM NaF, pH 7.5 (Section 2.4.19).

2.5 Fission yeast materials

2.5.1 Fission yeast media

Media used for the propagation of fission yeast were as described by Moreno et al (1991).

2.5.2 Fission yeast strains

All fission yeast strains used in this thesis are listed in Appendix I. The annotation GG refers to the Glasgow lab fission yeast collection number.

2.5.3 Fission yeast plasmid vectors

The pREP1 (GB 27) and pREP3X (GB 28) vectors, used for the expression of the His-tagged components of the DSC1 complex in fission yeast, were obtained from laboratory stocks (Appendix III).

pREP1: for the expression of genes under the control of the *nmt1*$^{+}$ (no message in thiamine) promoter (Maundrell, 1993).

pREP3X: for the expression of genes under the control of the *nmt1*$^{+}$ (no message in thiamine) promoter. Derived from the pREP1 vector it contains different restriction sites in the polylinker (Forsburg, 1993).

2.6 Fission yeast methods

2.6.1 Fission yeast cell culture

To resuscitate yeast strains from -70°C glycerol stocks, (maintained in 25% glycerol 75% YE) cells were streaked onto complete rich medium (YE) plates (Moreno et al., 1991). The cells were then grown at the permissive temperature (25°C or as otherwise indicated for mutant strains) for 2-3 nights and checked microscopically to ensure no contamination had occurred, and that the cells were growing normally. A few individual colonies were then picked and streaked onto a YE master-plate, which was incubated overnight at 25°C. The master plate was used to replica-plate colonies, using a velvet cloth, onto selective medium to confirm the yeast strain genotype. Strains containing plasmids or being used in experiments were propagated in EMM supplemented with the appropriate amino acids at a concentration of 100 μg ml^{-1} (Moreno et al., 1991). Temperature sensitive mutants were incubated at the restrictive temperature of 21°C to display their mutant phenotype. Cell number per ml of liquid culture was determined from a sample added to Isoton (Becton Dickinson); following sonication cells were counted electronically with a Z2 Coulter Counter. A cell count of 1-2 x 10^6 cells ml^{-1} indicated cells were at exponential phase of growth.

2.6.2 Production of competent fission yeast cells

500 ml of EMM plus appropriate supplements was inoculated with a 10 ml overnight culture. This culture was grown at 25°C until a density of approximately 1 x 10^7 cells ml^{-1} was obtained. The cells were harvested in 50 ml centrifuge tubes by spinning at 3000 rpm for 5 min in a chilled desktop centrifuge. As much supernatant as possible was removed and the cells gently resuspended in a total of 100 ml of ice-cold 1 M sorbitol. The cells were centrifuged as before and resuspended in 50 ml ice-cold 1 M sorbitol. This was repeated and the cells resuspended in 20 ml ice-cold 1 M sorbitol. Finally the cells were centrifuged and resuspended in 1-2 ml ice-cold 1 M sorbitol before being frozen at -70°C in 50 μl aliquots.

2.6.3 Transformation of competent fission yeast

Aliquots of cells were thawed at room temperature and immediately stored on ice. To 50 μl of competent cells 1-2 μl of the appropriate plasmid DNA was added. The cell-DNA suspension was transferred to a pre-chilled 0.2 cm electroporation cuvette and pulsed at 2.5 kV. 1 ml of 1 M sorbitol was then immediately added to the cuvette, the cell suspension returned to the microfuge tube and placed on ice. The transformation mixture was then plated out onto selective EMM, left to air dry, and incubated at permissive temperature for 3-4 days.

2.6.4 Induction of gene overexpression

Several fission yeast strains containing the pREP1 or pREP3X vectors with the appropriate insert were used to overexpress components of the DSC1 complex. Δ*res1* cells (GG 146) were transformed with the pREP vector alone (GG 796), *His-res1*$^{+}$ (GG 790), *His-res2*$^{+}$ (GG 802), *His-rep1*$^{+}$ (GG 811) or *His-rep2*$^{+}$ (GG 817) and Δ*res2* cells (GG 156) were transformed with the pREP vector alone (GG 801) or containing *His-res2*$^{+}$ (GG 808); Appendix I. The pREP vector series uses the *nmt1*$^{+}$ (no message in thiamine) promoter to control expression of the inserted genes. Strains were streaked from glycerol stocks onto EMM plates containing the appropriate amino acids for nutritional selection of the plasmid, with and without thiamine at a concentration of 5 μg μl^{-1} (*nmt1*$^{+}$ promoter 'off') and growth was compared at 30°C and 21°C. For protein extraction, 5ml cultures were grown to saturation at 30°C with the appropriate amino acids with and without 5 μg μl^{-1} thiamine. These were used to inoculate 200 ml cultures supplemented with the appropriate amino acids and 5 μg μl^{-1} thiamine (for *nmt1*$^{+}$ promoter 'off' experiments) or without thiamine (for *nmt1*$^{+}$ promoter 'on' experiments), and then grown for 16 h. Protein was then extracted for use in EMSA experiments as detailed (Section 2.6.5).

2.6.5 Protein extraction from fission yeast

Protein extraction from fission yeast was as described by Ng et al. (2001).

200 ml cultures of fission yeast cells, in mid-exponential stage of growth, were prepared and harvested by centrifugation at 5000 rpm for 10 min in screw-cap centrifuge tubes. The cell pellet was resuspended in 200 μl of ice-cold lysis buffer (50 mM KCl, 50 mM Tris-HCl pH 8, 25% glycerol, 2 mM DTT, 0.1% Triton X-100, 5 μg of protease inhibitors: chymostatin, pepstatin, antipain, leupeptin, aprotonin, 0.2 mM PMSF) in 2 ml screw-capped microfuge tubes. The cells were pelleted at 13,000 rpm for 1 min at 4°C in a high-speed microcentrifuge and again resuspended in 200 μl of lysis buffer. Acid washed glass beads (425-600 micron, Sigma) were added to just beneath the meniscus and the tubes were chilled on ice for 2-3 min before being disrupted using a Ribolyser (Hybaid Ltd, UK) with 1 burst at 40 sec, setting 4. The cell debris was pelleted by centrifugation at 13,000 rpm for 5 min at 4°C and the protein supernatants were transferred to a fresh chilled microfuge tube, and clarified by centrifugation at 13,000 rpm for 30 min at 4°C. Supernatants were transferred to a fresh chilled microfuge tube and 5 μl was removed to determine protein concentration, which was estimated using Bradford's reagent (Section 2.4.5) and the remainder of the protein sample was snap frozen on solid CO_2, and stored at -70°C.

Chapter 3

Molecular cloning and expression of individual components of the fission yeast DSC1 complex

3.1 Introduction

In both budding and fission yeasts the decision to enter the mitotic cell cycle is made at a point in late G1 phase known as START (Forsburg and Nurse, 1991). Passage through START and entry into S phase requires CDK activity and the activated transcription of genes that are essential for S phase. In budding yeast the transcriptional activation of these genes is dependent on the SBF and MBF transcription factor complexes and in fission yeast the MBF counterpart, DSC1 (Johnston and Lowndes, 1992).

The SBF complex is composed of an ScSwi6p-ScSwi4p heterodimer, which binds to the 5' CACGAAA 3' sequence in the promoters of several genes to regulate their transcription (Breeden and Nasmyth, 1987a; Andrews and Herskowitz, 1989a; Taba et al., 1991). Genes regulated by SBF include the G1 cyclins *CLN1*, *CLN2*, *PCL1* and *PCL2* (Johnston and Lowndes, 1992; Ogas et al., 1991; Measday et al., 1994; Koch and Nasmyth, 1994) and the *HO* endonuclease gene required for the mating type switch (Nasmyth, 1983; Breeden and Nasmyth, 1987a).

The MBF complex recognises promoters containing the MCB sequence element 5' ACGCGT 3' (McIntosh et al., 1991; Lowndes et al., 1991) and is composed of ScSwi6p in combination with ScMbp1p (Lowndes et al., 1992b; Dirick et al., 1992; Moll et al., 1992; Koch et al., 1993). Targets for MBF regulation in budding yeast include many genes required for DNA synthesis and two late G1 cyclins *CLB5* and *CLB6* (Johnston and Lowndes, 1992; McIntosh et al., 1991; Epstein and Cross, 1992; Schwob and Nasmyth, 1993). More recently a genome wide analysis of the SBF and MBF binding patterns has implicated these two complexes in the control of many more genes, some of which are transcription factors themselves, implying a greater gene regulatory role for these complexes than first thought (Iyer et al., 2001; Horak et al., 2002). In both SBF and MBF, DNA binding is mediated by, ScSwi4p and ScMbp1p, respectively (Primig et al., 1992; Dirick et al., 1992; Koch et al., 1993).

In the distantly related fission yeast the transcriptional activation of at least ten genes (*cdc22*$^{+}$, *cdc18*$^{+}$, *cig2*$^{+}$, *cdt1*$^{+}$, *rad21*$^{+}$, *suc22*$^{+}$, *rad11*$^{+}$, *ste9*$^{+}$, *mik1*$^{+}$ and *cdt2*$^{+}$) required for S phase and DNA replication is controlled by the DSC1 complex (Lowndes et al., 1992a; Fernandez-Sarabia et al., 1993; Kelly et al., 1993; Connolly and Beach, 1994; Hofmann and Beach, 1994; Birkenbihl and Subramani, 1995; Harris et al., 1996; Parker et al., 1997; Tournier and Millar, 2000; Ayte et al., 2001; Ng et al., 2001; Yoshida et al., 2003). Similar to MBF in budding yeast, this complex binds to MCB elements in the promoters of these genes. The SpRes1p, SpRes2p, SpCdc10p and SpRep2p proteins are all components of this complex (Lowndes et al., 1992a; Tanaka et al., 1992; Caligiuri and Beach, 1993; Miyamoto et al., 1994; Zhu et al., 1994; Nakashima et al., 1995) with a role in the meiotic cycle proposed for SpRep1p in place of SpRep2p (Sugiyama et al., 1994; Cunliffe et al., 2004). At least three of these

proteins, namely SpRes1p, SpRes2p and SpCdc10p, share close homology to their budding yeast counterparts.

A comparison of SBF/MBF/DSC1 proteins at the level of primary structure revealed a significant degree of homology specifically in their N and C-termini and in two centrally located ankyrin repeats. In ScSwi4p and ScMbp1p the N-terminus harbours their DNA binding capability whilst their C-termini are required for heteromeric complex formation with the C-terminus of ScSwi6p. Similarly, in SpRes1p and SpRes2p, the N-terminus contains the DNA binding domain and their C-termini bind to SpCdc10p. Common to all six proteins are the two centrally located ankyrin-repeat motifs (Bork, 1993; Sedgwick and Smerdon, 1999). The ankyrin repeats, believed to be involved in protein-protein interactions, have been reported to be required for interaction with cyclin-CDK complexes (Siegmund and Nasmyth, 1996; Foord et al., 1999). Both SpCdc10p and ScSwi6p also share a putative leucine zipper domain in their C termini (Reymond and Simanis, 1993; Sidorova and Breeden, 1993).

Much work has been carried out on these genes and their protein products, both genetic and biochemical, in terms of characterising their biological function (the functional properties of these proteins will be discussed in Chapters 4 and 5 of this thesis). In recent years significant progress has been made toward elucidating the structure and function of the ScSwi4p-ScSwi6p-ScMbp1p family of transcription factors from budding yeast. In contrast, little or no work has been carried out to study the 3-D structure of the fission yeast DSC1 components and so in comparison to their budding yeast homologues, these proteins remain poorly characterised at the atomic level.

To date, the 3-D structure of the DNA-binding domain of SpMbp1p has been determined by two independent groups at 2.1 Å resolution using X-ray crystallography (Taylor et al., 1997; Xu et al., 1997), and also more recently using NMR spectroscopy (Nair et al., 2003). X-ray crystallography was also employed to obtain a detailed structural analysis of the central ankyrin domain of ScSwi6p at 2.1 Å resolution (Foord et al., 1999).

Detailed structural analysis of a protein using methods such as those highlighted above requires highly pure, soluble and active preparation in large amounts (5 -10 mg). Whilst protein purification from the native source is the ideal situation, it can often involve lengthy protocols, using several types of chromatography in order to achieve purification to near homogeneity (a pre-requisite for detailed biophysical analysis). Furthermore, because many proteins of biological interest (and especially cell cycle transcription factors) are present in naturally tiny amounts, they are typically recovered in very low yields and are therefore not amenable for many structural and functional analyses. It is for these reasons that the production of a protein of interest is often carried out in a heterologous organism whose molecular mechanisms have been specifically engineered to produce a protein in amounts much greater than that available from the native source. Indeed, the analysis of all the budding yeast components of MBF have

been possible as a result of expression as recombinant proteins in *E. coli*. The pET system (Novagen) is a widely used prokaryotic system for cloning and expressing recombinant proteins in *E. coli*. Target genes are cloned into pET vectors under the control of the tightly regulated T7 RNA polymerase promoter and maintained in hosts that lack a T7 RNA polymerase gene. Once established in a non-expression host, the plasmid is then transformed into a strain engineered to carry a chromosomal copy of the T7 RNA polymerase gene under *lacUV5* control. Expression of the protein is then chemically induced by addition of IPTG.

The use of this prokaryotic system is not without limitations, as many of the post-translational modifications carried out in eukaryotes are unavailable in this host. Alternatively, many researchers therefore exploit a eukaryotic host for recombinant protein expression (the methylotrophic yeast *Pichia pastoris* and baculovirus systems are now widely used). Nevertheless, it is important to note that the relative success of each system used for heterologous expression must be determined empirically for each individual protein (many proteins that do not express at all, or to low levels, in the eukaryotic systems have often been successfully expressed in *E. coli* and vice versa). As precedent for this project, *E. coli*, as host for heterologous protein expression, was successfully applied permitting detailed structural information about the budding yeast MBF components.

The aim of this chapter was to clone the components of the fission yeast DSC1 transcription factor complex and to express them as His-tagged fusion proteins in *E. coli*. The cloning and expression of these components using the pET system has the potential to provide a reproducible source of recombinant proteins in amounts suitable for use in future functional and biophysical studies.

3.2 Cloning the *res1*$^+$, *res2*$^+$, *rep1*$^+$ and *rep2*$^+$ components of DSC1

The plasmids used for the cloning of the DSC1 components are listed in Appendix I – GB 121 (*rep1*$^+$), GB 23 (*rep2*$^+$), GB 164 (*res2*$^+$) and GB 160 (*res1*$^+$).

Previously in our laboratory, PCR of the cDNAs encoding the components of the fission yeast DSC1 complex was performed using primers that permitted the addition of N-terminal *Nde*I and C-terminal *Bam*HI restriction sites. These sites were not present in the open reading frames of the genes being cloned. The fidelity of the PCR reactions and the integrity of the restriction sites at the ends of each gene were confirmed by DNA sequencing after cloning into the vector pBC KS$^+$ (MWG-Biotech – using oligos listed in Appendix II). The presence of the restriction ends, specifically the N-terminal *Nde*I site, allowed cloning into the pET-28c vector in the correct reading frame with an N-terminal 6-histidine tag (His-tag). The inclusion of the His-tag provides a means of conveniently purifying the fusion protein by immobilised metal ion affinity chromatography (IMAC). This vector also contains the neomycin phosphotransferase gene, conferring kanamycin antibiotic resistance to the cloned plasmid.

The plasmids containing the His-tagged cDNAs of the DSC1 components were digested with *Nde*I and *Bam*HI, as was the pET-28c vector to generate compatible cohesive ends for ligation. The vector was also treated with calf intestinal alkaline phosphatase to remove the phosphate groups exposed by digestion in order to prevent self-ligation (Methods 2.2.7 and 2.2.8). The products of the restriction digestions were analysed on a 1.5% (w/v) agarose gel and bands of the appropriate size excised and purified using the QIAquick® Gel Extraction Kit (Methods 2.2.4 and 2.2.6). 5 µl samples were analysed on a 1.5% (w/v) agarose gel to determine the amount of DNA retrieved and thus to empirically determine the vector:insert ratio to be used in the subsequent ligation. Ligation reactions were carried out as described (Methods 2.2.9) and transformed into *E. coli* DH5α cells the following day (Methods 2.2.13). Transformants were then plated on LB-agar plates supplemented with the antibiotic kanamycin (30 µg ml^{-1}) and incubated at 37°C overnight. Colonies were selected and cultured overnight at 37°C in 5 ml LB media supplemented with kanamycin (30 µg ml^{-1}).

Plasmid DNA was purified from the overnight cultures as described in Methods section 2.2.5 and analysed by restriction enzyme mapping to confirm the presence of an insert of the predicted size (Figures 3.1 and 3.2). Potential positive clones were sequenced (MWG-Biotech - using oligos listed in Appendix II) to confirm the presence of insert and ensure that the genes had been cloned in-frame with the N-terminal His-tag. After confirmation by restriction mapping and DNA sequencing, a positive clone for each gene was stored at -70°C in the lab bacterial collection: GB 177 (*rep1*$^+$), GB 191 (*rep2*$^+$), GB 178 (*res2*$^+$) and GB 201 (*res1*$^+$) - Appendix *I*. These clones were then used for the subsequent overexpression of heterologous protein.

Figure 3.1 Restriction digestion analyses of the *rep1*$^{+}$-pET-28c and *rep2*$^{+}$-pET-28c plasmids

Clones containing insert were analysed by restriction mapping. Samples of each digest were analysed on a 1.5% (w/v) agarose gel stained with ethidium bromide and viewed under UV illumination. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs.

Digestion of the *rep1*$^{+}$-pET-28c plasmid: **1** *Nde*I-*Bam*HI **2** *Eco*RI **3** *Eco*RV.

Digestion of the *rep2*$^{+}$-pET-28c plasmid: **4** *Nde*I-*Bam*HI.

The restriction enzymes used in each digest and the predicted fragment sizes are summarised for each clone in Tables 3.1 and 3.2.

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Nde*I-*Bam*HI**	**2**	**1420, 5330**
***Eco*RI**	**3**	**370, 660, 5720**
***Eco*RV**	**2**	**2300, 4450**

Table 3.1 *rep1*$^{+}$-pET-28c

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Nde*I-*Bam*HI**	**2**	**660, 5330**

Table 3.2 *rep2*$^{+}$-pET-28c

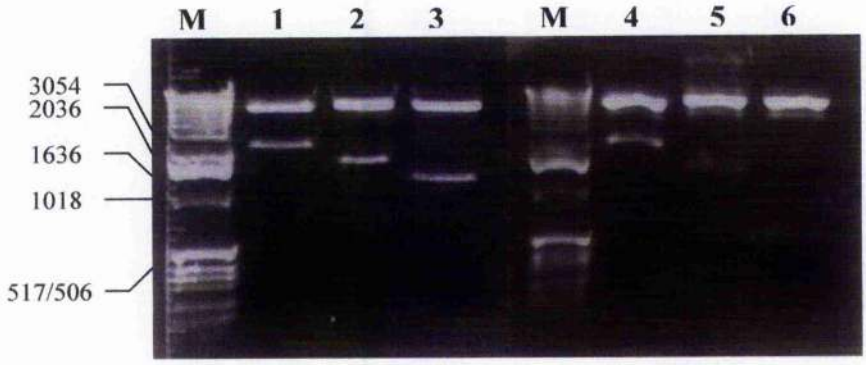


Figure 3.2 Restriction digestion analyses of the *res2*$^{+}$-pET-28c and *res1*$^{+}$-pET-28c plasmids

Clones containing insert were analysed by restriction mapping. Samples of each digest were analysed on a 1.5% (w/v) agarose gel stained with ethidium bromide and viewed under UV illumination. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs.

Digestion of the *res2*$^{+}$-pET-28c plasmid: **1** *Nde*I-*Bam*HI **2** *Eco*RI **3** *Nco*I

Digestion of the *res1*$^{+}$-pET-28c plasmid: **4** *Nde*I-*Bam*HI **5** *Hind*III **6** *Xba*I

The restriction enzymes used in each digest and the predicted fragment sizes are summarised for each clone in Tables 3.3 and 3.4.

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Nde*I-*Bam*HI**	**2**	**1970, 5330**
***Eco*RI**	**2**	**1530, 5770**
***Nco*I**	**2**	**1200, 6100**

Table 3.3 *res2*$^{+}$-pET-28c

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Nde*I-*Bam*HI**	**2**	**1910, 5330**
***Hind*III**	**2**	**1370, 5870**
***Xba*I**	**2**	**570, 6670**

Table 3.4 *res1*$^{+}$-pET-28c

3.3 Cloning the *cdc10*$^+$ component of DSC1

The recombinant plasmid containing the *cdc10*$^+$ gene cloned into the *Nde*I-*Bam*HI sites of pET-28c (previously cloned in our laboratory) was found to produce a truncated protein product upon overexpression. Subsequent DNA sequencing analysis identified an artificial internal stop codon as the cause of this truncation. Site-directed mutagenesis was therefore carried out to remove the internal stop site and is outlined below (details of the site-directed mutagenesis of *cdc10*$^+$ are described in Methods 2.2.11). The site-directed mutagenesis invokes a G to C single base pair substitution at position 426 relative to the ATG. This is a silent mutation changing TAG (Stop) to TAC (Tyr) since the wild type *cdc10*$^+$ gene contains a TAT (Tyr) at the corresponding codon. The incorporation of C rather than T, with respect to wild type, at base pair 426 was necessary to remove an internal *Nde*I site.

3.3.1 Site-directed mutagenesis of *cdc10*$^+$

The plasmid pET-28c containing a 2,304 bp fragment representing the *cdc10*$^+$ gene cloned in at *Nde*I-*Bam*HI was obtained from a laboratory stock (GB 130: Appendix I) and used as template for site-directed mutagenesis. Figure 3.3 provides a schematic representation of the mutagenesis reaction and the primers used.

Initially, two separate PCR reactions were carried out using primer pairs GO 546/GO 548 and GO 547/GO 459, to generate fragments of approximately 500 bp and 2000 bp, respectively. The products obtained from these PCR reactions are shown in Figure 3.4. Both products of the predicted size, containing the G to C substitution at position 426 (as introduced by the mutagenic primers GO 547 and GO 548) were gel purified (Methods 2.2.6) and analysed by electrophoresis on a 1.5% (w/v) agarose gel (data not shown). 0.5 µl of each purified PCR product was then used as template in a single fusion PCR reaction using the primers GO 546 and GO 459. These primers are complementary to either end of the *cdc10*$^+$ gene in GB130 and incorporate the N-terminal *Nde*I and C-terminal *Bam*HI sites, respectively. All PCR reactions were carried out using the Vent$_R$ DNA polymerase (New England Biolabs). This enzyme has extensive 3' to 5' exonuclease proofreading activity, thus helping to minimise errors in base mis-incorporation.

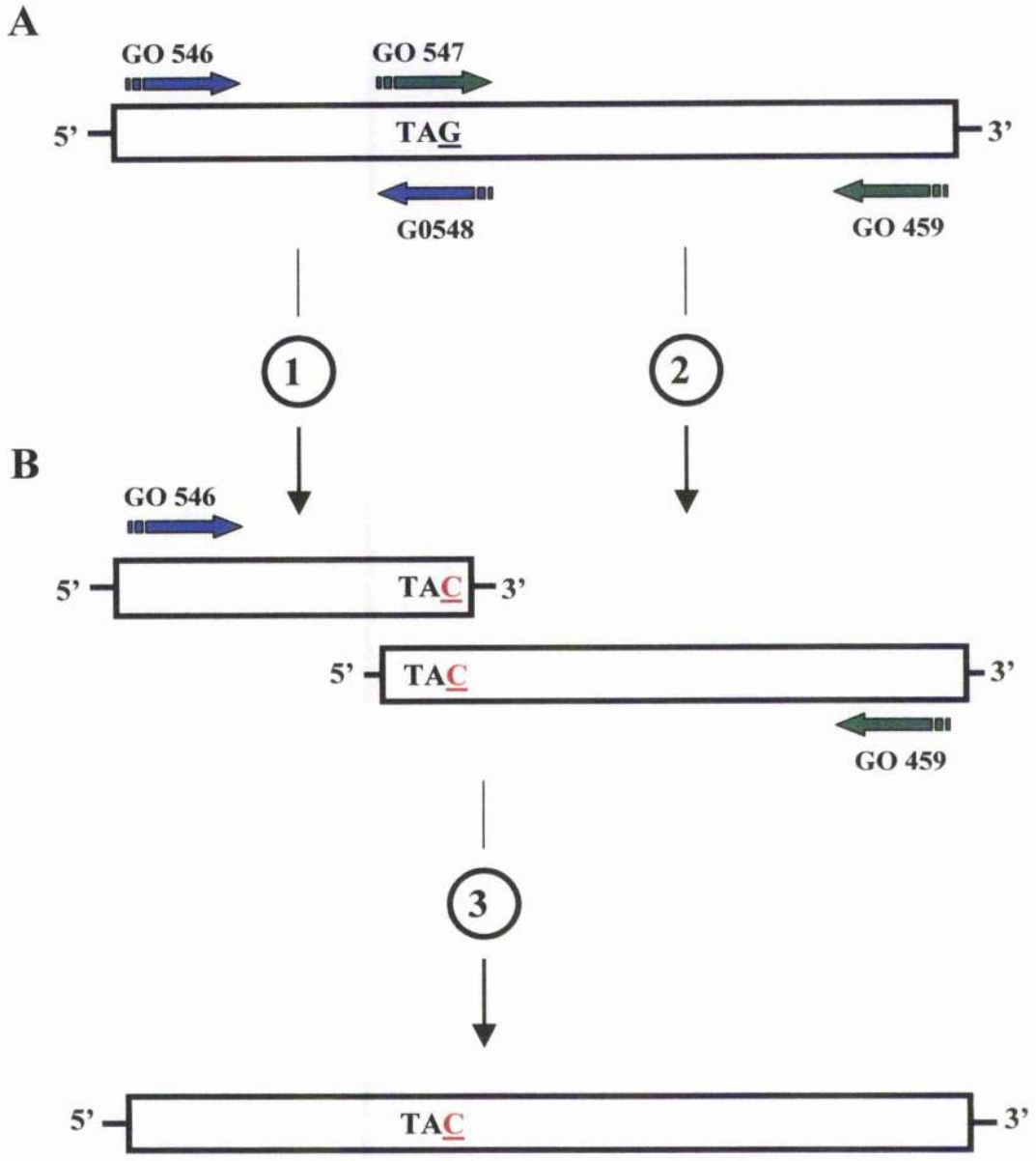


Figure 3.3 PCR strategy for the site-directed mutagenesis of *cdc10*$^{+}$

A schematic representation of the primers used in the site-directed mutagenesis of *cdc10*$^{+}$. **A** In separate PCR reactions, 1 and 2, two fragments of the target gene were amplified. PCR 1 uses primers GO 546 and GO 548, whereas PCR 2 uses primers GO 547 and GO 459. **B** The two overlapping fragments were then used as a template in a fusion PCR reaction, 3, with primers GO 546 and GO 459 to produce the full-length mutant DNA.

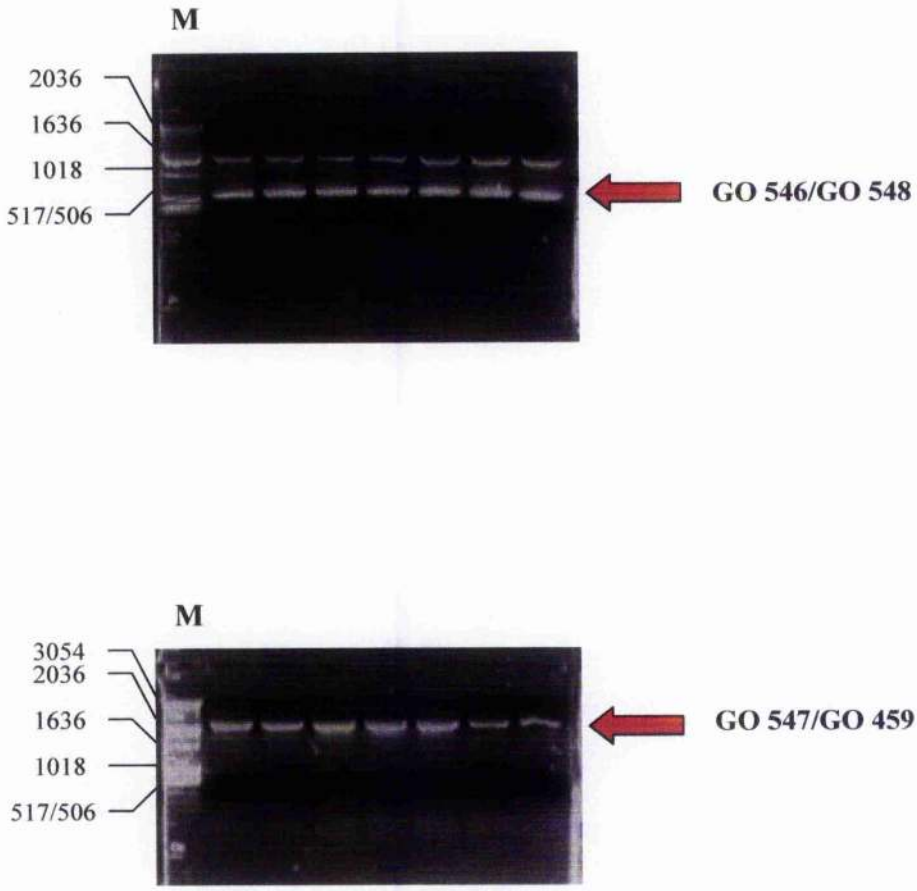


Figure 3.4 Site-directed mutagenesis of *cdc10*$^{+}$ by PCR amplification

PCR products were analysed on 1.5% (w/v) agarose gels stained with ethidium bromide. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs. Arrows indicate PCR products of the predicted size, from reactions using primer pairs GO 546/GO 548 and GO 547/GO 459. The presence of contaminating PCR products in the GO 546/GO 548 reaction was due to non-specific priming in early cycles.

3.3.2 Cloning the silently mutated *cdc10*$^{+}$ gene into pCR2.1®

The TA Cloning® Kit (Invitrogen) was used to clone the blunt-ended PCR product from the fusion PCR reaction (Methods 2.2.10).

Cloning of PCR products into the pCR2.1® vector is a one-step method that utilises the non-template dependent activity of *Taq* polymerase to add single deoxyadenosine (A) to the 3' ends of PCR products. The linearised pCR2.1® vector contains single 3' deoxythymidine (T) residues, thus allowing efficient ligation of a PCR insert with vector. $Vent_R$ DNA polymerase was used in the PCR reactions to improve fidelity as it has a much lower rate of base mis-incorporation than *Taq* polymerase due to intrinsic proofreading activity. However, $Vent_R$ does not leave single 3' A overhangs, and so a necessary extra final step at the end of the fusion PCR reaction was to incubate the reaction mix with *Taq* at 72°C for 10 min.

3.3.3 Ligation, transformation and identification of clones

After treatment with *Taq* a PCR product corresponding to the approximate size of the *cdc10*$^{+}$ gene was gel purified, an aliquot analysed on an agarose gel (data not shown) and used in a ligation reaction with the linearised pCR2.1® vector. The cloning of the *cdc10*$^{+}$ fusion PCR product into the pCR2.1® vector and transformation was as described in Methods section 2.2.10.

Transformants were plated on LB-agar plates containing 1.6 mg μl^{-1} X-Gal, 50 $\mu g\ ml^{-1}$ ampicillin and incubated at 37°C overnight. Colonies were selected according to the blue/white screening method and white colonies picked and cultured overnight in LB media supplemented with ampicillin (50 $\mu g\ ml^{-1}$) at 37°C. Plasmid DNA was purified from the overnight cultures as described in Methods section 2.2.5 and analysed by restriction enzyme mapping to confirm the presence of an insert of the correct size (Figure 3.5). Potential positive clones were sequenced (MWG-Biotech - using oligos listed in Appendix II) to confirm the presence of insert and that the site-directed mutagenesis was successful. After confirmation by restriction mapping and DNA sequencing, a positive clone was stored at -70°C in the lab bacterial collection as GB 313 (Appendix I).

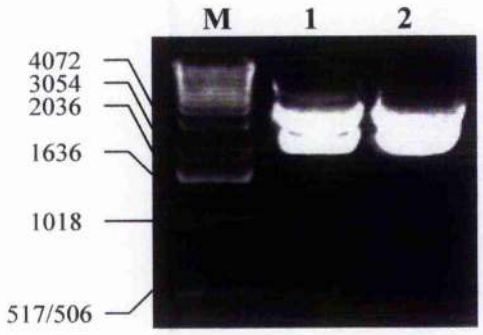


Figure 3.5 Restriction digestion analysis of the recombinant *cdc10*$^{+}$-pCR2.1® plasmid

Clones containing insert were analysed by restriction mapping. Samples of each digest were analysed on a 1.5% (w/v) agarose gel stained with ethidium bromide and viewed under UV illumination. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs.

Digestion of the *cdc10*$^{+}$-pCR2.1® plasmid: **1** *Nde*I-*Bam*HI **2** *Eco*RI.

The restriction enzymes used in each digest and the predicted fragment sizes are summarised for each clone in Table 3.5.

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Nde*I-*Bam*HI**	**2**	**2300, 3930**
***Eco*RI**	**2**	**2400, 3830**

Table 3.5 *cdc10*$^{+}$-pCR2.1®

3.3.4 Cloning the silently mutated *cdc10*$^{+}$ gene into pET-14b

The plasmid GB 313, containing the *cdc10*$^{+}$ fusion PCR product, was digested with *Nde*I and *Bam*HI as was the pET-14b vector to generate cohesive ends for ligation. The vector was also treated with calf intestinal alkaline phosphatase, to remove the phosphate groups exposed by digestion and so prevent self-ligation (Methods 2.2.7 and 2.2.8). The products of the restriction digests were analysed on a 1.5 % (w/v) agarose gel and bands of the appropriate size excised and purified using the QIAquick® Gel Extraction Kit (Methods 2.2.4 and 2.2.6). 5 µl samples were analysed on a 1.5 % (w/v) agarose gel to determine the amount of DNA retrieved and thus to empirically determine the vector:insert ratio to be used in the subsequent ligation reaction. Ligation reactions were carried out as described (Methods 2.2.9) and transformed into *E. coli* DH5α cells the following day (Methods 2.2.13).

Transformants were then plated on LB-agar plates supplemented with the antibiotic ampicillin (50 µg ml^{-1}) and incubated at 37°C overnight. Colonies were selected and cultured overnight at 37°C in 5 ml LB media supplemented with ampicillin (50 µg ml^{-1}).

Plasmid DNA was purified from the overnight cultures as described in Methods section 2.2.5 and analysed by restriction enzyme mapping to confirm the presence of an insert of the correct size (Figure 3.6). Potential positive clones were sequenced (MWG-Biotech - using oligos listed in Appendix II) to confirm the presence of insert and that the gene had been cloned in-frame with the N-terminal His-tag. After confirmation by restriction mapping and DNA sequencing, a positive clone was stored at -70°C in the lab bacterial collection as GB 314 (Appendix I). This clone was then used for the subsequent overexpression of heterologous protein.

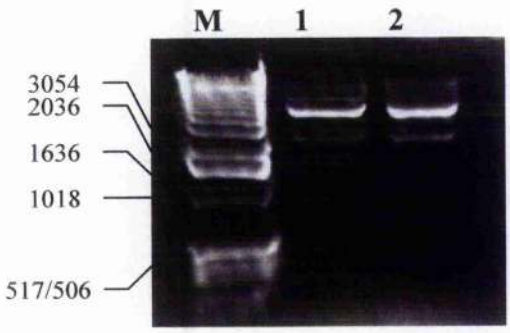


Figure 3.6 Restriction digestion analysis of the recombinant *cdc10*$^{+}$-pET-14b plasmid

Clones containing insert were analysed by restriction mapping. Samples of each digest were analysed on a 1.5% (w/v) agarose gel stained with ethidium bromide and viewed under UV illumination. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs.

Digestion of the *cdc10*$^{+}$-pET-14b plasmid: **1** *Nde*I-*Bam*HI **2** *Pst*I.

The restriction enzymes used and the expected fragment sizes produced from each digest are summarised for each clone in Table 3.6.

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Nde*I-*Bam*HI**	**2**	**2300, 4660**
***Pst*I**	**2**	**2250, 4710**

Table 3.6 *cdc10*$^{+}$-pET-14b

3.4 Overexpression of the recombinant proteins

The previously isolated clones, containing insert for each component of the DSC1 complex, were then transformed into the *E. coli* expression strain BL21 (DE3) CodonPlus-RIL and small scale protein inductions performed to test for successful overexpression of each protein (Methods 2.4.1).

An aliquot of a 5 ml overnight culture supplemented with either kanamycin (30μg/ml) or ampicillin (50 μg ml^{-1}), as appropriate, was subcultured into a fresh 50 ml culture and grown with shaking at 37°C until the optical density reached ~ 0.5 at 600 nm. Expression of protein was then induced by the addition of IPTG to a final concentration of 1 mM. Inductions were initially carried out at 37°C and 1 ml samples were removed at the point of induction (0 h) and at hourly intervals thereafter for 3 h. Samples were then pelleted by centrifugation, resuspended in Laemmli sample buffer and analysed by 10% (w/v) SDS-polyacrylamide gel electrophoresis and Coomassie blue staining. The results of each recombinant protein induction are shown in Figures 3.7-3.9. Immunoblot analysis was also carried out, using an antibody directed against the His-tag, to confirm that the N-terminus was intact (Figures 3.10-3.12).

All five proteins were successfully expressed at 37°C using this system, although His-Res1p and His-Rep2p were only observed by immunoblotting. The His-Res1p and His-Res2p proteins resolved with a M_r approximately 75,000, while their predicted values are 74,604 and 75,864, respectively. The His-Rep1p and His-Rep2p proteins resolved with a M_r approximately 55,000 and 30,000, while their predicted values are 54,781 and 26,831, respectively. The His-Cdc10p protein resolved with a M_r approximately 90,000 while the predicted value is 87,676. Once the proteins had been successfully expressed it was then necessary to test their solubility.

This was achieved by breaking open the bacterial cells, by either chemical or mechanical means, to release the protein into the extracellular environment. Insoluble proteins and cellular debris are typically removed by centrifugation with any soluble protein being retrieved from the supernatant. The discussion of the solubility of these recombinant proteins is delayed until Chapter 5 where it is discussed with relevance to the biological activity of the recombinant DSC1 components *in vitro*.

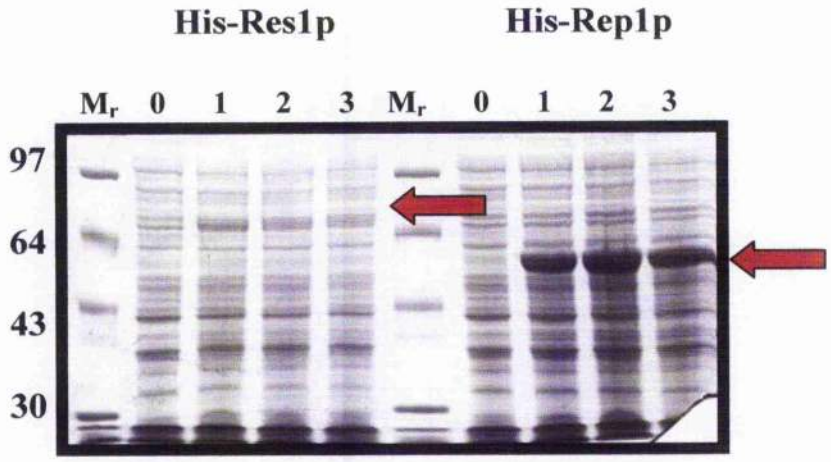


Figure 3.7 Overexpression of His-Res1p and His-Rep1p

His-Res1p and His-Rep1p were overexpressed at 37°C in BL21 (DE3) CodonPlus-RIL cells by induction with 1 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at hourly intervals thereafter for 3 h. Samples were centrifuged and pellets resuspended in Laemmli sample buffer (10 µl / 0.1 absorbance unit) then denatured by boiling for 5 min in the presence of DTT (150 mM). Samples were then analysed on a 10 % SDS-polyacrylamide gel stained with Coomassie brilliant blue. Expression of His-Res1p was not detected by Coomassie blue staining under these conditions (the anticipated position of His-Res1p is indicated by the arrow). Expression of His-Rep1p was observed in lanes 1-3 (indicated by the arrow). Molecular weight markers are shown (M_r) with sizes indicated in kDa.

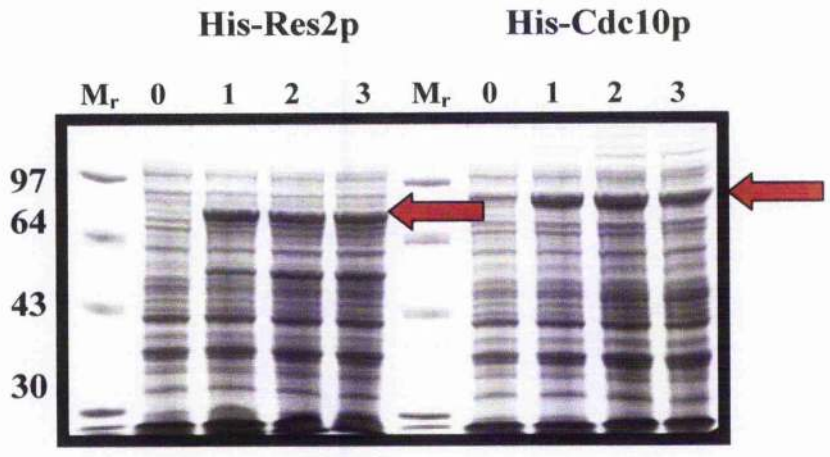


Figure 3.8 Overexpression of His-Res2p and His-Cdc10p

His-Res2p and His-Cdc10p were overexpressed at 37°C in BL21 CodonPlus (DE3)-RIL cells by induction with 1 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at hourly intervals thereafter for 3 h. Samples were centrifuged and pellets resuspended in Laemmli sample buffer (10 µl / 0.1 absorbance unit) then denatured by boiling for 5 min in the presence of DTT (150 mM). Samples were then analysed on a 10 % SDS-polyacrylamide gel stained with Coomassie brilliant blue. Expression of His-Res2p and His-Cdc10p was observed in lanes 1-3 in both cases (indicated by arrows). Molecular weight markers are shown (M_r) with sizes indicated in kDa.

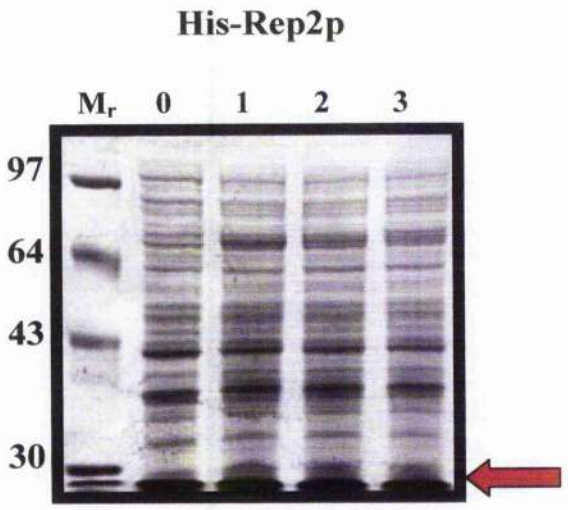


Figure 3.9 Overexpression of His-Rep2p

His-Rep2p was overexpressed at 37°C in BL21 CodonPlus (DE3)-RIL cells by induction with 1 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at hourly intervals thereafter for 3 h. Samples were centrifuged and pellets resuspended in Laemmli sample buffer (10 µl / 0.1 absorbance unit) then denatured by boiling for 5 min in the presence of DTT (150 mM). Samples were then analysed on a 10 % SDS-polyacrylamide gel stained with Coomassie brilliant blue. Expression of His-Rep2p was not detected by Coomassie blue staining under these conditions (the anticipated position of His-Rep2p is indicated by the arrow). Molecular weight markers are shown (M_r) with sizes indicated in kDa.

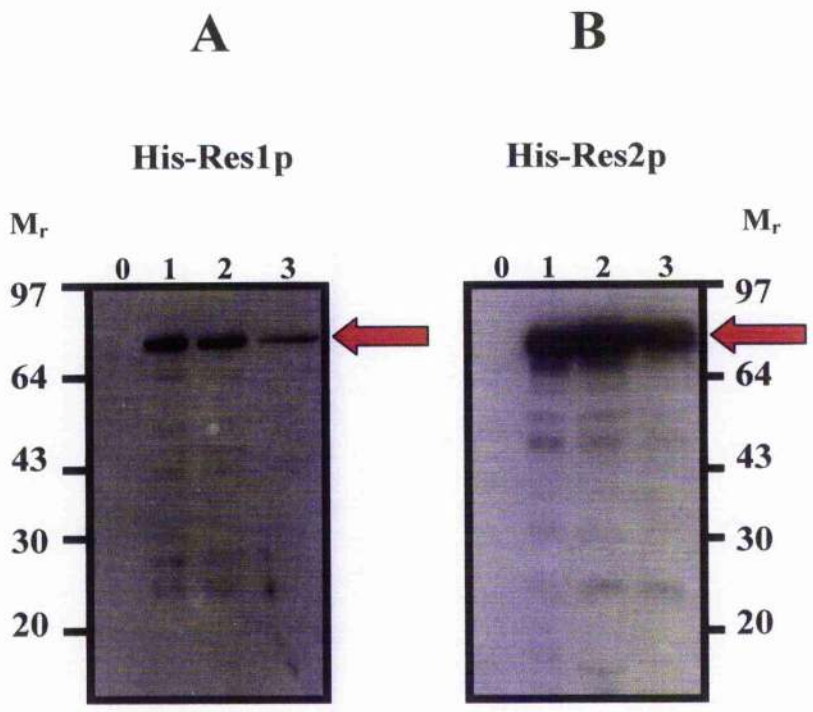


Figure 3.10 Immunoblot analysis of overexpressed His-Res1p and His-Res2p

A His-Res1p and **B** His-Res2p were overexpressed at 37°C in BL21 CodonPlus (DE3)-RIL cells by induction with 1 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at hourly intervals thereafter for 3 h. Immunoblotting was carried out as described (Methods 2.4.7-2.4.11) and blots were probed with Anti-PentaHis-HRP conjugate antibody (Qiagen) at 1:2500. Expression of His-Res1p and His-Res2p was observed in lanes 1-3 in both cases. Molecular weight markers are shown (M_r) with sizes indicated in kDa.

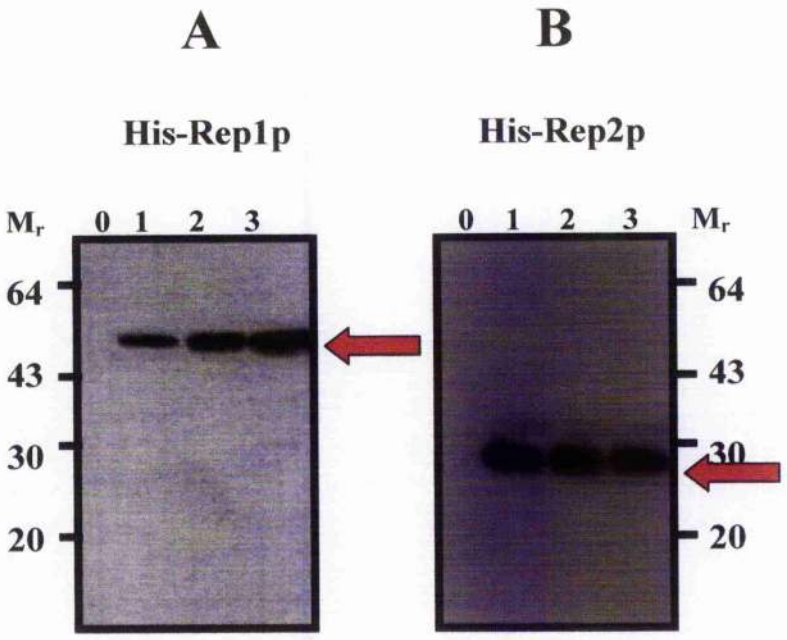


Figure 3.11 Immunoblot analysis of overexpressed His-Rep1p and His-Rep2p

A His-Rep1p and **B** His-Rep2p were overexpressed at 37°C in BL21 CodonPlus (DE3)-RIL cells by induction with 1 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at hourly intervals thereafter for 3 h. Immunoblotting was carried out as described (Methods 2.4.7-2.4.11) and blots were probed with Anti-PentaHis-HRP conjugate antibody (Qiagen) at 1:2500. Expression of His-Rep1p and His-Rep2p was observed in lanes 1-3 in both cases. Molecular weight markers are shown (M_r) with sizes indicated in kDa.

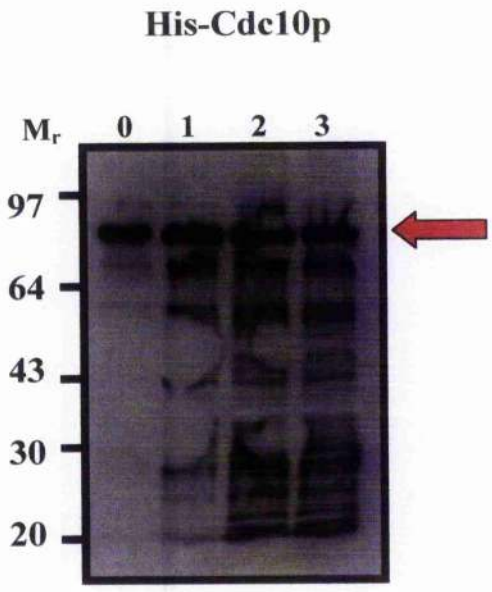


Figure 3.12 Immunoblot analysis of overexpressed His-Cdc10p

His-Cdc10p was overexpressed at 37°C in BL21 CodonPlus (DE3)-RIL cells by induction with 1 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at hourly intervals thereafter for 3 h. Immunoblotting was carried out as described (Methods 2.4.7-2.4.11) and blots were probed with Anti-PentaHis-HRP conjugate antibody (Qiagen) at 1:2500. Expression of His-Cdc10p was observed in lanes 1-3 as indicated by the arrow (basal expression was observed in lane 0 h - see text for discussion). Molecular weight markers are shown (M_r) with sizes indicated in kDa.

3.5 Discussion

In this chapter the cloning strategy has been described which allowed the successful and reproducible overexpression of the individual components of the fission yeast DSC1 complex as analysed by SDS-PAGE, followed by Coomassie blue staining and immunoblotting.

High-level expression of the Res2p, Rep1p and Cdc10p components of DSC1 as His-tagged fusion proteins was achieved as visualised by SDS-PAGE and Coomassie blue staining (Figures 3.7-3.8). Upon immunoblot analysis of His-Cdc10p some basal expression was seen at 0 h (Figure 3.12). High-level basal expression can often result in plasmid toxicity causing prevention of growth and eventually leading to plasmid loss. However, the low levels of basal expression seen with His-Cdc10p had no such adverse effects. The overexpression of the His-Res1p and His-Rep2p components was achieved at lower levels, such that they were detectable only by immunoblotting (Figures 3.7, 3.10 and 3.9, 3.11, respectively). Attempts to optimise the expression of all the recombinant proteins made use of alternative *E. coli* host strains, each with subtle molecular differences. Expression of each recombinant protein was tested in *E. coli* BL21 (DE3) pLysS cells. These cells contain a plasmid that provides a source of T7 lysozyme, which binds to T7 RNA polymerase, serving to minimise basal transcription. Furthermore, with specific regard to overexpression of His-Res1p and His-Rep2p, the BL21Star™ (DE3) strain was also tested. The BL21Star™ (DE3) strain carries a mutated RNase enzyme that lacks the ability to degrade mRNA. Consequently, mRNAs expressed in this strain are typically more stable, resulting in increased protein expression. However, despite the use of these and other host strains, optimal conditions for each recombinant protein were achieved using the BL21 (DE3) CodonPlus-RIL cells. This observation was attributed to the fact that BL21 (DE3) CodonPlus-RIL cells contain extra copies of several tRNA species that are 'rare' to *E. coli*. Most amino acids are encoded by more than one codon and each organism has a bias for particular codons for certain amino acids. Therefore, in each cell of that organism, the tRNA population will reflect the codon bias in the mRNA. As fission yeast proteins were being expressed in *E. coli*, differences in codon usage between the two organisms could result in aberrant translation. The tRNAs that are required and abundant in fission yeast might be absent or in short supply in *E. coli*. This can lead to inefficient translation and incorporation of the wrong amino acid as a result of 'rare' tRNA depletion in *E. coli*. Use of an *E. coli* strain which has extra copies of these 'rare' tRNAs engineered on a plasmid can often be used to overcome such problems.

The expression of soluble recombinant proteins suggests that they are capable of folding into their native or near-native state in *E. coli*. Production of soluble protein in this case also indicates that the presence of the N-terminal His-tag does not affect folding or expression in this host.

The solubility of the recombinant proteins produced in this chapter is discussed in Chapter 5 where their activities are analysed by *in vitro* assay. The following Chapter 4 discusses the use of fission yeast, the organism from which these proteins are derived, as a host to assay the effects of the N-terminal His-tag on their biological activity *in vivo*.

Chapter 4

Biological activity assays of recombinant DSC1 proteins *in vivo*

4.1 Introduction

Chapter 3 described the cloning of the fission yeast genes of DSC1 components *cdc10*$^{+}$, *res1*$^{+}$, *res2*$^{+}$, *rep1*$^{+}$ and *rep2*$^{+}$ into the bacterial expression vector pET-28c (pET-14b for *cdc10*$^{+}$), in-frame with an N-terminal His-tag. Overexpression of each protein in *E. coli* has the potential to produce sufficient amounts to allow detailed structural analyses, with the addition of the His-tag to facilitate purification by means of affinity chromatography.

The function of a protein is dictated by its 3-D architecture. Even a single amino acid change in the native sequence can dramatically alter protein activity, exemplified in this context by the *sct1-b1* mutant of *res1*$^{+}$, and various point mutants of *cdc10*$^{+}$ (Caligiuri and Beach, 1993; Reymond et al., 1992). Therefore, as the N-terminal His-tag is naturally absent from these proteins, its presence may affect their function. In this chapter therefore, a study was undertaken to examine possible effects of the N-terminal His-tag on the biological activity of these proteins when produced *in vivo*.

Current models predict that the fission yeast mitotic DSC1 complex is composed of the SpCdc10p, SpRes1p, SpRes2p and SpRep2p proteins (SpRep2p is replaced by SpRep1p in the meiotic cycle). SpCdc10p was the first component of DSC1 to be identified (Lowndes et al., 1992). The identification of the *cdc10*$^{+}$ gene product as a physical component of the DSC1 complex, coupled to the availability of conditional lethal *cdc10* mutants (Reymond et al., 1992), allowed the identification of further DSC1 components, through yeast genetic suppressor analysis.

SpRes1p and SpRes2p were first identified by their genetic interaction with *cdc10*$^{+}$. Both proteins were subsequently shown to interact with each other. The *res1*$^{+}$ gene was identified as a high copy suppressor of the *cdc10-129* mutant and is required for passage of START (Tanaka et al., 1992). In addition, the *res1*$^{+}$ gene could almost completely suppress a Δ*cdc10* deletion mutant. In contrast, the *cdc10*$^{+}$ gene could not rescue cells deleted for *res1*$^{+}$. Furthermore, the putative SpRes1p protein showed significant homology to both SpCdc10p and ScSwi6p and in particular to the DNA binding domain of ScSwi4p. Moreover, the *res1*$^{+}$ gene was cloned independently as an extragenic suppressor of the *cdc10-129* mutant, and shown to be part of the DSC1 band-shift activity (Caligiuri and Beach, 1993).

Cells deleted for *res1*$^{+}$ can grow, albeit poorly, at 30°C but show severe heat and cold sensitivities, resulting in a lethal phenotype at 36°C and 21°C (Tanaka et al., 1992; Caligiuri and Beach, 1993). The observation that cells deleted for the *res1*$^{+}$ gene could proceed through the cell cycle at 30°C prompted the search for a gene redundant in function with *res1*$^{+}$ and ultimately to the identification of SpRes2p.

Cloned by its ability to rescue the heat and cold sensitivities of the Δ*res1* mutant, the *res2*$^{+}$ gene could also effectively suppress the *cdc10-129* mutant at restrictive temperature, but

unlike *res1*$^+$, was unable to rescue a Δ*cdc10* strain. In contrast to *res1*$^+$, deletion of *res2*$^+$ showed no apparent defects in mitotic growth, although defects in the meiotic cycle indicated *res2*$^+$ may have a major role (Miyamoto et al., 1994). The *res2*$^+$ gene was also cloned independently in a genetic screen designed to identify SpCdc10p binding proteins, and was subsequently shown to bind to MCB elements *in vitro*, in association with SpCdc10p (Zhu et al., 1994).

The *rep1*$^+$ gene was isolated as an extragenic suppressor of the *cdc10-129* mutant at restrictive temperature but, like *res2*$^+$, was unable to rescue a Δ*cdc10* mutant. In addition, *rep1*$^+$ could effectively suppress the heat and cold sensitivities of the Δ*res1* null mutant. Similar to cells deleted for *res2*$^+$, the Δ*rep1* mutant had no apparent mitotic defects, but instead was defective in the meiotic cycle (Sugiyama et al., 1994).

The fact that the *rep1*$^+$ gene was not expressed in mitosis but could rescue the *cdc10-129* and Δ*res1* mutants, led to the search for a *rep1*$^-$ like gene acting in mitosis. Consequently, the *rep2*$^+$ gene was cloned by its ability to rescue the *cdc10-129* mutant at restrictive temperature. In addition, *rep2*$^+$ could suppress the heat and cold sensitivities of the Δ*res1* mutant, whereas cells deleted for *rep2*$^+$ were viable at 30°C but showed cold sensitivity at 18°C (Nakashima et al., 1995).

Thus, a common property of these genes (with the exception of *cdc10*$^+$) is their ability to rescue the heat and cold sensitivities of a Δ*res1* mutant at 36°C and 21°C, and these properties were exploited to assay the biological activity of the recombinant His-tagged proteins. If the His-tag does not affect their function, the recombinant proteins should behave similarly to their wild type counterparts in rescuing a Δ*res1* mutant.

A further characteristic of cells deleted for either *res1*$^+$ or *res2*$^+$ is loss of DSC1 as detected *in vitro* by band-shift analysis (Zhu et al., 1997; Ayte et al., 1997). The loss of this band-shift activity in strains deleted for either *res1*$^+$ or *res2*$^+$ can be reconstituted by introduction of either the *res1*$^+$ or *res2*$^+$ gene on a plasmid, respectively. However, expression of the *res1*$^+$ or *res2*$^+$ gene in strains deleted for either *res2*$^+$ or *res1*$^+$, respectively, does not result in reappearance of this band-shift activity (Zhu et al., 1997).

These observations, together with the demonstration that the DSC1 complex could be super-shifted in band-shift analysis using antibodies specific to either SpCdc10p, SpRes1p or SpRes2p, was the first evidence to suggest that a single complex containing all three proteins existed (Zhu et al., 1997; Ayte et al., 1997). These observations revised initial proposals that suggested that, similar to the role of ScSwi6p in the SBF and MBF complexes of budding yeast, the SpCdc10p protein of fission yeast was a common subunit of two distinct DNA binding complexes. Earlier models proposed that a SpRes1p-SpCdc10p DSC1 complex acted primarily in the mitotic cycle, whilst a SpRes2p-SpCdc10p complex behaved similarly in the

meiotic cycle (Miyamoto et al., 1994; Zhu et al., 1994). The existence of a single fission yeast DSC1 complex was confirmed by the demonstration that a SpCdc10p-SpRes1p-SpRes2p heteromeric complex existed throughout the mitotic cell cycle (Whitehall et al., 1999).

In addition to the genetic suppressor activities mentioned above, the abilities of both His-Res1p and His-Res2p to participate in the DSC1 complex were assessed. In the case of His-Res1p and His-Res2p, the ability of these proteins to participate in a DSC1 complex produced *in vivo* can be tested. The SpRep1p protein, however, is not expressed during the mitotic cycle and so does not form a part of the DSC1 complex under these conditions (although it is postulated to form part of DSC1 in the meiotic cycle; Sugiyama et al., 1994; Cunliffe et al., 2004) thus precluding the use of this assay for His-Rep1p. The SpRep2p protein binds to SpRes2p *in vitro* and both SpRes1p and SpRes2p *in vivo* (Nakashima et al., 1995; Sturm and Okayama, 1996). Nevertheless, the DSC1 band-shift activity is not lost in a Δ*rep2* mutant background suggesting that the DSC1 complex can still form in the absence of SpRep2p (Baum et al., 1997).

To summarise, the ectopic expression of *His-res1*$^{+}$, *His-res2*$^{+}$, *His-rep1*$^{+}$ or *His-rep2*$^{+}$ in a Δ*res1* mutant of fission yeast, allows an *in vivo* test for gene function via rescue of a conditional-lethal mutant phenotype at restrictive temperature. With respect to *His-res1*$^{+}$ and *His-res2*$^{+}$, expression of these cDNAs in both Δ*res1* and Δ*res2* mutants, respectively, should allow reconstitution of the DSC1 band-shift activity. These methods were exploited to assay the biological activities of recombinant DSC1 components.

The aim of this chapter was to clone the recombinant components of the DSC1 complex into the fission yeast expression vector pREP and using these constructs, assay the effect of the His-tag on the biological activities of these proteins *in vivo*.

4.2 The fission yeast expression vector pREP

The fission yeast thiamine-repressible expression vectors pREP1 and pREP3X were used to clone the recombinant DSC1 components for the *in vivo* experiments (Appendix I- GB27 and GB28).

In pREP, genes are cloned downstream of the thiamine repressible *nmt1*$^{+}$ (*n*ot *m*ade in *t*hiamine) promoter (Maundrell, 1992). As the name suggests, transcription of genes under the control of this promoter is regulated by the presence of thiamine. In the absence of thiamine transcription is induced, whereas repression is achieved by addition of 5 μg μl^{-1} thiamine to the growth medium. Nevertheless, despite addition of thiamine, transcription from the *nmt1*$^{+}$ promoter is often not completely repressed. The ability to switch off transcription depends on the cloned gene. For example, many genes under *nmt1*$^{+}$ control retain their suppressor activity despite the presence of thiamine whereas, at this same dosage, transcription of other genes may be switched off (Maundrell, 1992; Forsburg, 1993). Thus, the effect of expressing a gene under *nmt1*$^{+}$ control, in either the presence or absence of thiamine, must be determined empirically.

The availability of suitable restriction sites within the multiple cloning region of the pREP3X vector allowed cloning of the *His-res1*$^{+}$, *His-rep1*$^{+}$ and *His-rep2*$^{+}$ cDNAs, following PCR amplification. In contrast, the *His-res2*$^{+}$ cDNA was cloned into the pREP1 vector in a modified directional cloning procedure, reflecting the suitability of restriction sites present within the insert and multiple cloning region of the vector. Both pREP1 and pREP3X are derived from the parent pREP vector series and differ in only two restriction sites within their multiple cloning regions (pREP3X contains a *Xho*I site in addition to a *Bal*I site in place of *Nde*I in pREP1).

4.3 Cloning recombinant *His-res1*$^+$, *His-res2*$^+$, *His-rep1*$^+$ and *His-rep2*$^+$ components of DSC1

The recombinant plasmids, containing the *His-res1*$^+$, *His-res2*$^+$, *His-rep1*$^+$ and *His-rep2*$^+$ cDNAs, were cloned (Chapter 3) and are listed in Appendix I - GB 177 (*His-rep1*$^+$), GB 191 (*His-rep2*$^+$), GB 178 (*His-res2*$^+$) and GB 201 (*His-res1*$^+$). These plasmids were used as templates in PCR reactions, using primers GO 458 and GO 459, designed to add 5' *Sal*I and 3' *Bam*HI restriction sites (upstream of the His-tag and immediately following the stop codon), respectively. Addition of these restriction sites was necessary to facilitate cloning into the pREP3X vector. In contrast, *His-res2*$^+$ was cloned into pREP1, due to the presence of a *Sal*I restriction site within the open reading frame of the *res2*$^+$ gene (Section 4.3.3).

4.3.1 Cloning *His-res1*$^+$, *His-rep1*$^+$ and *His-rep2*$^+$ cDNAs into pREP3X

Figure 4.1 provides a schematic representation of the PCR reactions and the primers used. PCR reactions were performed using the $Vent_R$ DNA polymerase. Conditions for PCR were essentially as described with the annealing temperature at 50°C (Methods 2.2.3). PCR products of the predicted size were gel purified (Methods 2.2.6) and analysed by electrophoresis on a 1.5 % (w/v) agarose gel (Figure 4.2).

4.3.2 Ligation, transformation and identification of clones

5 µl of each purified PCR product was digested with *Sal*I and *Bam*HI, as was the pREP3X vector, to generate compatible cohesive ends for ligation. The vector was also treated with calf intestinal alkaline phosphatase (Methods 2.2.7 and 2.2.8). The products of the restriction digestions were analysed on a 1.5 % (w/v) agarose gel and bands of the appropriate size excised and purified using the QIAquick® Gel Extraction Kit (Methods 2.2.4 and 2.2.6). 5 µl samples were analysed on a 1.5 % (w/v) agarose gel to assay the amount of DNA retrieved to determine the vector:insert ratio to be used in the ligation. Ligation reactions were carried out as described (Methods 2.2.9) and transformed into *E. coli* DH5α cells the following day (Methods 2.2.13).

Transformants were then plated on LB-agar plates supplemented with the antibiotic ampicillin (50 µg ml^{-1}) and incubated at 37°C overnight. Colonies were selected and cultured overnight at 37°C in 5 ml LB media supplemented with ampicillin (50 µg ml^{-1}).

Plasmid DNA was purified from the overnight cultures as described in Methods section 2.2.5 and analysed by restriction enzyme mapping to confirm the presence of an insert of the predicted size (Figures 4.3 - 4.5). Potential positive clones were sequenced (MWG-Biotech - using oligos listed in Appendix II) to confirm the presence of insert and fidelity of the PCR reactions. After confirmation by restriction mapping and DNA sequencing, a positive clone for

each gene was stored at -70°C in the lab bacterial collection: GB 202 (*His-rep1*$^{+}$), GB 343 (*His-rep2*$^{+}$) and GB 203 (*His-res1*$^{+}$) - Appendix I.

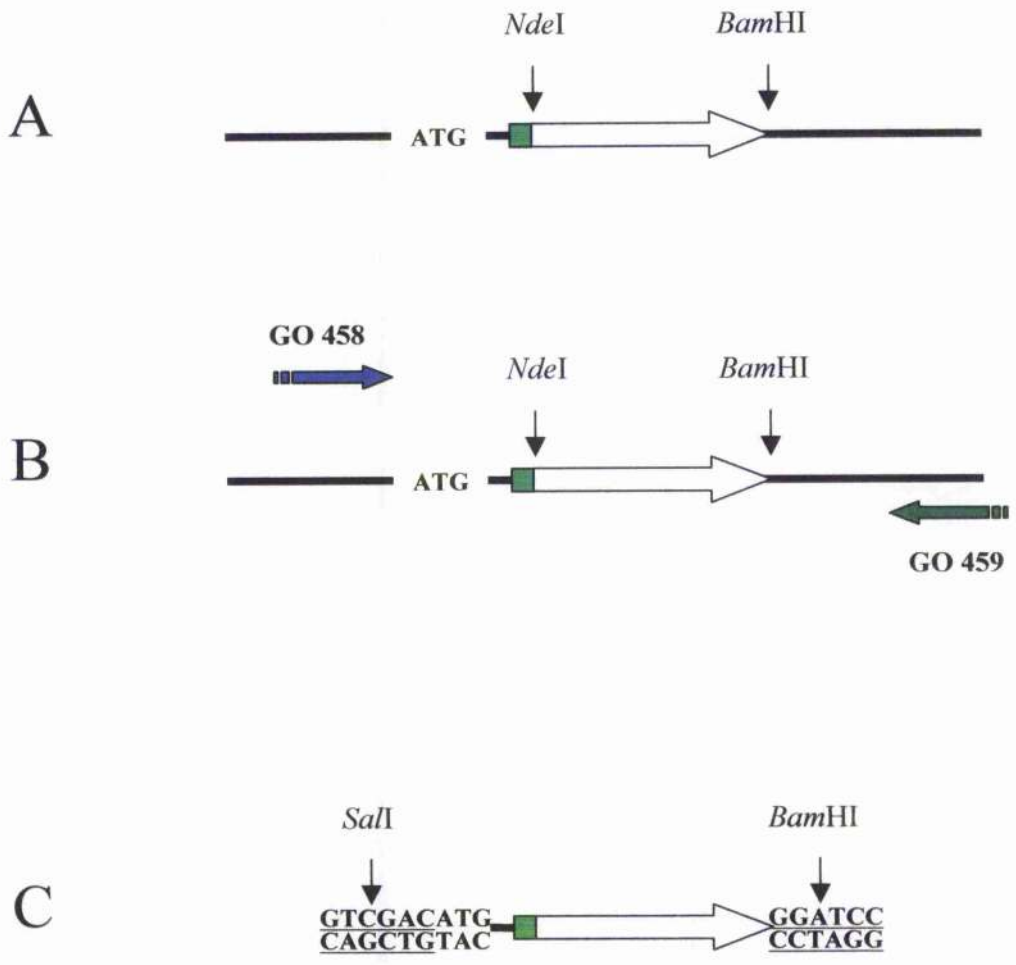


Figure 4.1 Schematic representation of the PCR reactions for pREP3X cloning

The white arrow represents the ORFs of the *res1*$^{+}$, *rep1*$^{+}$ and *rep2*$^{+}$ cDNAs cloned into pET28c at *Nde*I-*Bam*HI sites. The His-tag is boxed in green, the black lines represent pET vector flanking sequence and arrows indicate the positions of restriction sites.

A Restriction map representing the *Nde*I-*Bam*HI fragment of each cDNA cloned into the pET-28c vector (Chapter 3). The **ATG** supplied by the vector is indicated upstream of the His-tag.

B The positions of the primers used in the PCR reactions are indicated in relation to the corresponding sequence of the pET-28c vector. GO 458 adds a *Sal*I site (**GTCGAC**) upstream of the ATG and His-tag. GO 459 includes the *Bam*HI (**GGATCC**) site immediately downstream of the stop codon.

C A schematic representation of the DNA fragments produced following PCR.

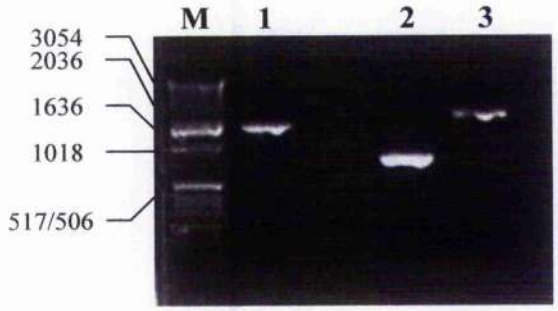


Figure 4.2 PCR amplification of *His-rep1*$^+$, *His-rep2*$^+$ and *His-res1*$^+$

PCR products were analysed on a 1.5% (w/v) agarose gel stained with ethidium bromide and viewed under UV illumination. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs. PCR products of the predicted size, from reactions using primer pairs GO458/GO459 are shown in lanes **1** *His-rep1*$^+$ **2** *His-rep2*$^+$ and **3** *His-res1*$^+$. The *His-rep1*$^+$ DNA sequence resolves at the 1636 bp marker (the *rep1*$^+$ cDNA is 1419 bp long), *His-rep2*$^+$ resolves between the 517 bp and 1018 bp markers (the *rep2*$^+$ cDNA is 660 bp long) and the *His-res1*$^+$ DNA sequence resolves between the 1636 bp and 2036 bp markers (the *res1*$^+$ cDNA is 1914 bp long).

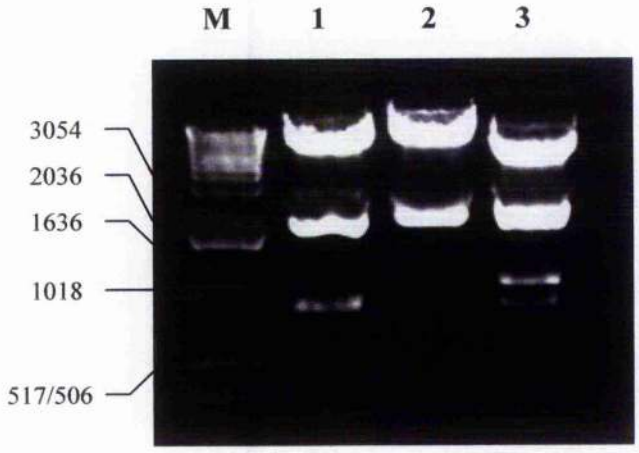


Figure 4.3 Restriction digestion analysis of the recombinant *His-res1*$^{+}$-pREP3X plasmid

Clones containing insert were analysed by restriction mapping. Samples of each digest were analysed on a 1.5% (w/v) agarose gel stained with ethidium bromide and viewed under UV illumination. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs.

Digestion of the recombinant *His-res1*$^{+}$-pREP3X plasmid: **1** *Nde*I-*Bam*HI **2** *Sal*I-*Bam*HI **3** *Hind*III.

The restriction enzymes used in each digest and the predicted fragment sizes are summarised in Table 4.1. Fragments marked (*) appear as a result of partial digest.

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Nde*I-*Bam*HI**	**4**	**7600, 700*, 1900, 800**
***Sal*I-*Bam*HI**	**2**	**8400, 1900**
***Hind*III**	**4**	**7600, 700*, 2200, 500**

Table 4.1 *His-res1*$^{+}$-pREP3X

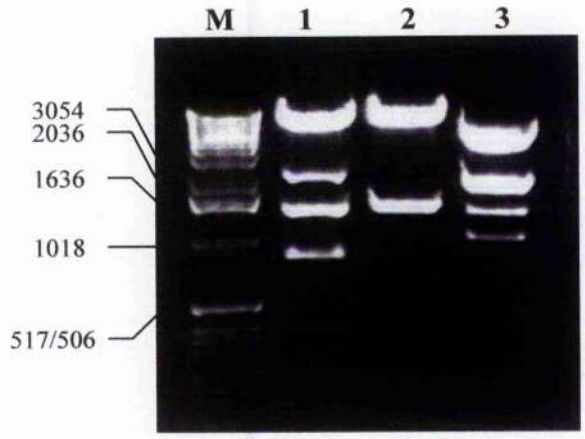


Figure 4.4 Restriction digestion analysis of the recombinant *His-rep1*$^{+}$-pREP3X plasmid

Clones containing insert were analysed by restriction mapping. Samples of each digest were analysed on a 1.5% (w/v) agarose gel stained with ethidium bromide and viewed under UV illumination. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs.

Digestion of the recombinant *His-rep1*$^{+}$-pREP3X plasmid: **1** *Nde*I-*Bam*HI **2** *Sal*I-*Bam*HI **3** *Hind*III.

The restriction enzymes used in each digest and the predicted fragment sizes are summarised in Table 4.2. Fragments marked (*) appear as a result of partial digest.

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Nde*I-*Bam*HI**	**4**	**7600, 200*, 1400, 800**
***Sal*I-*Bam*HI**	**2**	**8400, 1400**
***Hind*III**	**4**	**7500, 300*, 1200, 1100**

Table 4.2 *His-rep1*$^{+}$-pREP3X

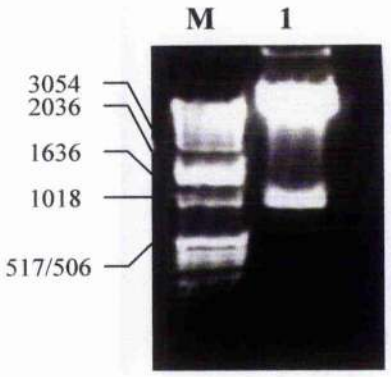


Figure 4.5 Restriction digestion analysis of the recombinant *His-rep2*$^{+}$-pREP3X plasmid

Clones containing insert were analysed by restriction mapping. Samples of each digest were analysed on a 1.5% (w/v) agarose gel stained with ethidium bromide and viewed under UV illumination. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs.

Digestion of the recombinant *His-rep2*$^{+}$-pREP3X plasmid: **1** *Sal*I-*Bam*HI.

The restriction enzymes used in each digest and the predicted fragment sizes are summarised in Table 4.3

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Sal*I-*Bam*HI**	**2**	**8400, 700**

Table 4.3 *His-rep2*$^{+}$ -pREP3X

4.3.3 Cloning recombinant *His-res2*$^+$ cDNA into pREP1

As mentioned, the pREP3X vector was unsuitable for cloning the *His-res2*$^+$ cDNA, owing to a *Sal*I restriction site within the open reading frame of the *res2*$^-$ gene. Alternatively, the *His-res2*$^+$ cDNA was cloned into the pREP1 vector. The pREP1 and pREP3X vectors are essentially identical differing only in the restriction sites available within their multiple cloning regions. The plasmid containing the *His-res2*$^+$ cDNA (GB 178) was digested with *Xba*I and *Bam*HI (which cut upstream of the His-tag and immediately following the stop codon, respectively). Importantly, the *Xba*I enzyme cleaves upstream of an ATG which is 5' to, and in-frame with, the His-tag. The inclusion of an ATG is necessary due to the destruction of the pREP1 vector supplied ATG, within the *Nde*I site, in the subsequent cloning procedure. The pREP1 vector was digested with *Nde*I and *Bam*HI and treated with calf intestinal alkaline phosphatase, in preparation for ligation. The products of the restriction digestions were analysed on a 1.5% (w/v) agarose gel and bands of the appropriate size excised and purified using the QIAquick® Gel Extraction Kit (Methods 2.2.4 and 2.2.6). 5 µl samples were analysed on a 1.5% (w/v) agarose gel, again to allow determination of the vector:insert ratio to be used in the subsequent ligation. The ligation reaction was carried out essentially as described (Methods 2.2.9), although in two steps owing to the incompatible *Xba*I and *Nde*I restriction ends of the insert and vector, respectively. An initial ligation reaction was carried out for 30 min, to allow the compatible *Bam*HI ends to ligate. 0.5 µl of Mung bean nuclease (an enzyme which degrades single stranded nucleotides) was then added to remove the *Xba*I and *Nde*I overhangs, creating blunt ends. Finally, 0.5 µl of T4 DNA ligase was added to ligate the blunt ends (Figure 4.6). Ligation reactions were then transformed into *E. coli* DH5α cells the following day (Methods 2.2.13).

Transformants were plated on LB-agar plates supplemented with the antibiotic ampicillin (50 µg ml^{-1}) and incubated at 37°C overnight. Colonies were selected and cultured overnight at 37°C in 5 ml LB media supplemented with ampicillin (50 µg ml^{-1}).

As before; following plasmid DNA purification, restriction enzyme mapping (Figure 4.7) and DNA sequencing (MWG-Biotech - using oligos listed in Appendix II); a positive clone was stored at -70°C in the lab bacterial collection: GB 195 (*res2*$^+$) - Appendix I.

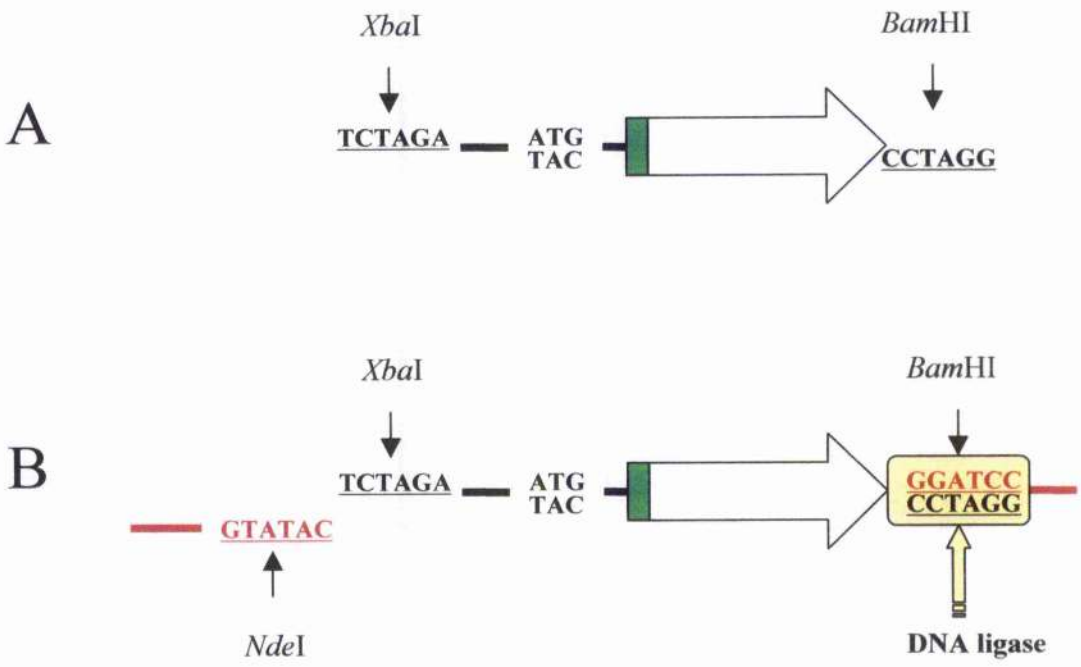


Figure 4.6 Cloning the *His-res2*$^{+}$ cDNA into pREP1 using Mung Bean nuclease

The white arrow represents the ORF of the *res2*$^{+}$ cDNA cloned into pET-28c at *Nde*I-*Bam*HI (Chapter3). The His-tag is boxed in green, the black lines represent pET vector flanking sequence and arrows indicate the positions of restriction sites. The red lines represent pREP1 vector sequence.

A The plasmid containing the *His-res2*$^{+}$ cDNA (GB 178) was digested with *Xba*I and *Bam*HI to excise the *His-res2*$^{+}$ cDNA. The *Xba*I enzyme cleaves upstream of the pET vector supplied ATG, which is in-frame with the His-tag.

B The pREP1 vector was digested with *Nde*I and *Bam*HI. An initial ligation reaction was then carried out for 30 min to allow the compatible *Bam*HI ends between the insert and vector to ligate.

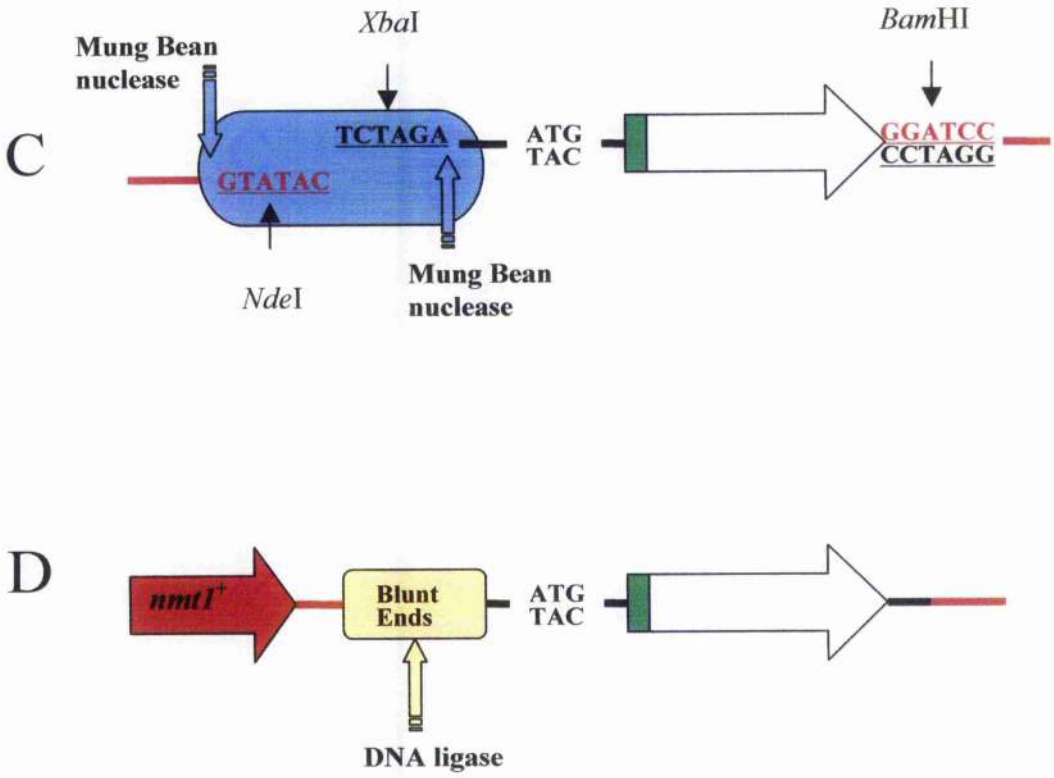


Figure 4.6 Cloning the *His-res2*$^{+}$ cDNA into pREP1 using Mung Bean nuclease

C Mung bean nuclease was then added to remove the *Xba*I and *Nde*I overhangs of the insert and vector, respectively. Degradation of the single stranded nucleotides creates blunt ends and destroys the pREP1 supplied ATG.

D Finally, the blunt ends are ligated by DNA ligase, placing the *His-res2*$^{+}$ cDNA under *nmt1*$^{+}$ control.

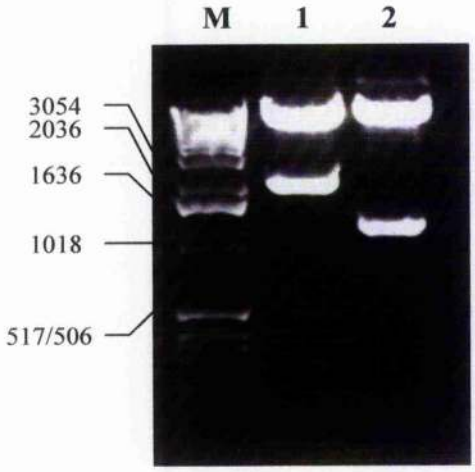


Figure 4.7 Restriction digestion analysis of the recombinant *His-res2*$^{+}$-pREP1 plasmid

Clones containing insert were analysed by restriction mapping. Samples of each digest were analysed on a 1.5% (w/v) agarose gel stained with ethidium bromide and viewed under UV illumination. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs.

Digestion of the recombinant *His-res2*$^{+}$-pREP1 plasmid: **1** *Nde*I-*Bam*HI **2** *Nco*I.

The restriction enzymes used in each digest and the predicted fragment sizes are summarised in Table 4.4

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Nde*I-*Bam*HI**	**2**	**7600, 1900**
***Nco*I**	**2**	**9100, 1200**

Table 4.4 *His-res2*$^{+}$-pREP1

4.4 Biological activity assays – rescuing Δ*res1* lethal phenotype

The lethal phenotype of the Δ*res1* mutant, at 21°C, can be rescued by expression of the full-length wild type *res1*$^{+}$, *res2*$^{+}$, *rep1*$^{+}$ or *rep2*$^{+}$ cDNAs (Tanaka et al., 1992; Miyamoto et al., 1994; Sugiyama et al., 1995; Nakashima et al., 1995). Accordingly, expression of the corresponding recombinant cDNAs, each having an N-terminal His-tag, would be expected to behave similarly, if the tag does not affect the function of the proteins. Therefore, following successful cloning of the recombinant *His-res1*$^{+}$, *His-res2*$^{+}$, *His-rep1*$^{+}$ and *His-rep2*$^{+}$ cDNAs into the pREP vector each construct was subsequently transformed into the fission yeast Δ*res1* strain (GG 146 - Appendix I and Methods 2.6.3) and stored at -70°C in the lab fission yeast collection. Δ*res1* cells were also transformed with the empty pREP vector to act as a negative control (GB 28 - Appendix I).

To assess the ability of the *His-res1*$^{+}$, *His-res2*$^{+}$, *His-rep1*$^{+}$ and *His-rep2*$^{+}$ cDNAs to rescue the cold sensitivity of a Δ*res1* mutant at 21°C, cells were streaked from -70°C glycerol stocks onto EMM plates containing adenine (for nutritional selection of the plasmid), in either the presence (*nmt1*$^{+}$ promoter 'off') or absence (*nmt1*$^{+}$ promoter 'on') of thiamine at a concentration of 5 μg μl^{-1} (Methods 2.6.4). Rescue of the lethal phenotype of the Δ*res1* mutant at 21°C was compared between strains transformed with *His-res1*$^{+}$-pREP3X, *His-res2*$^{+}$-pREP1, *His-rep1*$^{+}$-pREP3X or *His-rep2*$^{+}$-pREP3Xand the empty pREP vector. In all four cases the Δ*res1* mutant was viable at the restrictive temperature, whereas cells containing the empty vector were not, indicating the recombinant cDNAs could function in a similar fashion to their wild type counterparts (Figures 4.8 - 4.11).

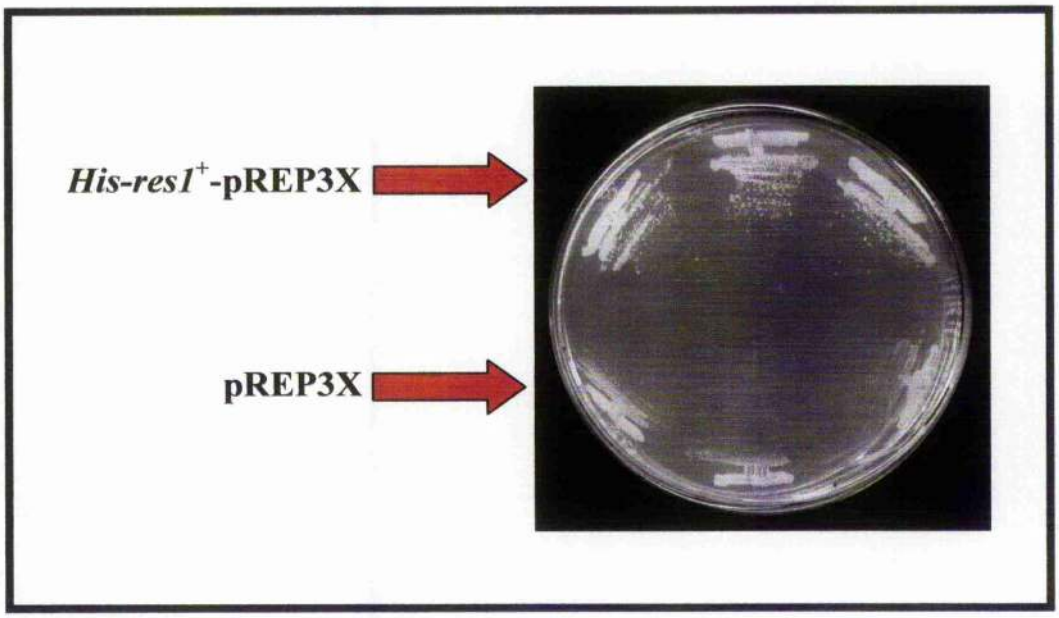


Figure 4.8 Expression of *His-res1*$^{+}$ rescues the growth defect of the Δ*res1* mutant at 21°C

The Δ*res1* strain (GG 146) was transformed with pREP3X (GG 796) or *His-res1*$^{+}$-pREP3X (GG 790). Transformants were allowed to grow at 30°C on EMM + thiamine (5 μg μl^{-1}), then streaked out onto the same medium and incubated at 21°C for 6-8 days.

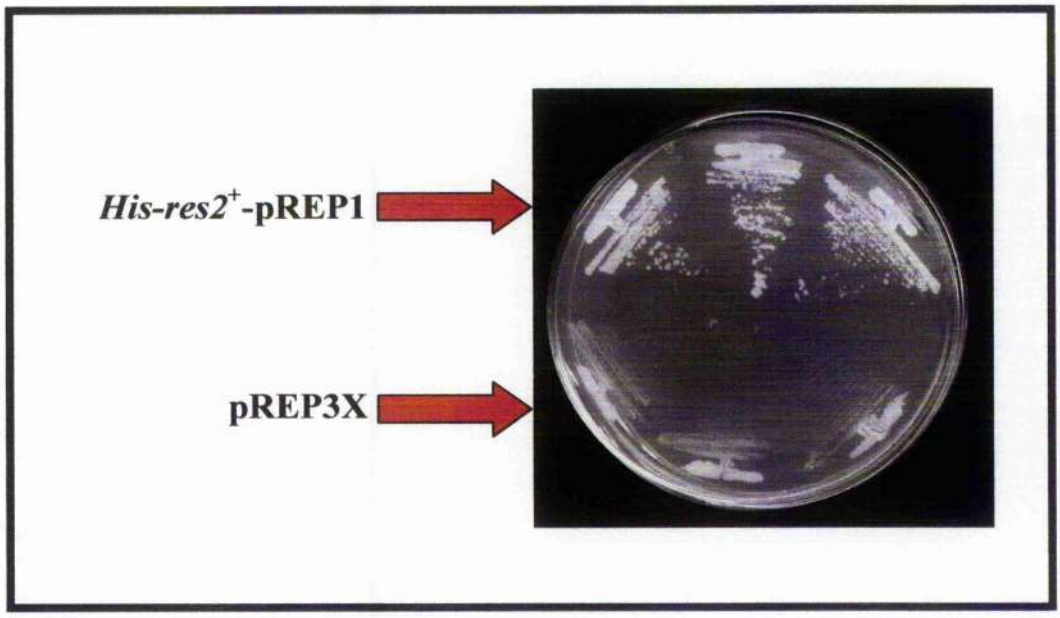


Figure 4.9 Expression of *His-res2*$^{+}$ rescues the growth defect of the Δ*res1* mutant at 21°C

The Δ*res1* strain (GG 146) was transformed with pREP3X (GG 796) or *His-res2*$^{+}$-pREP1 (GG 802). Transformants were allowed to grow at 30°C on EMM + thiamine (5 μg μl^{-1}), then streaked out onto the same medium and incubated at 21°C for 6-8 days.

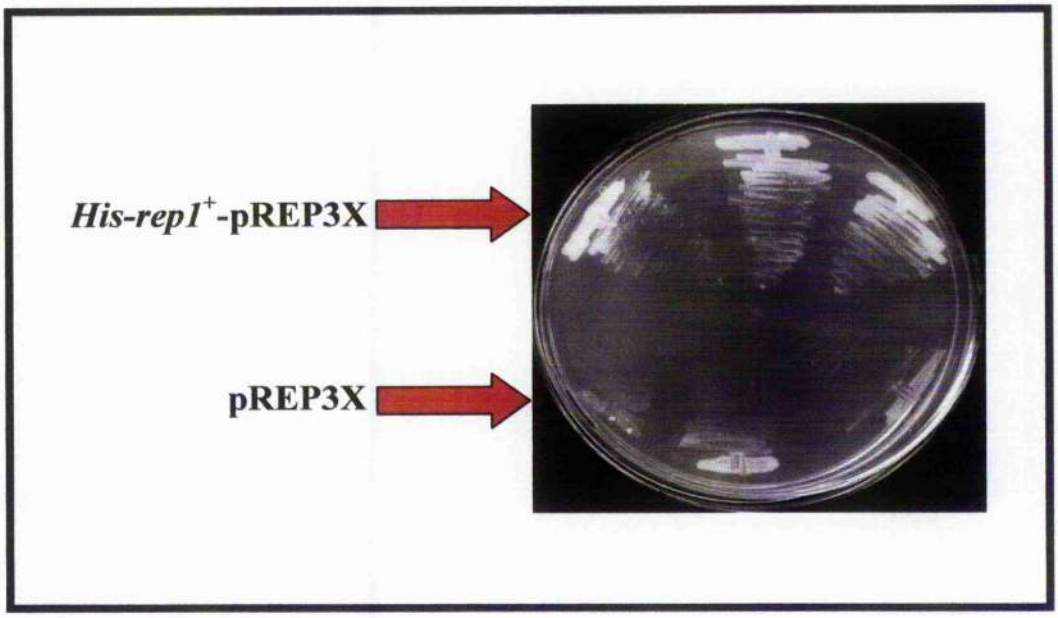


Figure 4.10 Expression of *His-rep1*$^{+}$ rescues the growth defect of the Δ*res1* mutant at 21°C
The Δ*res1* strain (GG 146) was transformed with pREP3X (GG 796) or *His-rep1*$^{+}$-pREP3X (GG 811). Transformants were allowed to grow at 30°C on EMM + thiamine (5 μg μl^{-1}), then streaked out onto the same medium and incubated at 21°C for 6-8 days.

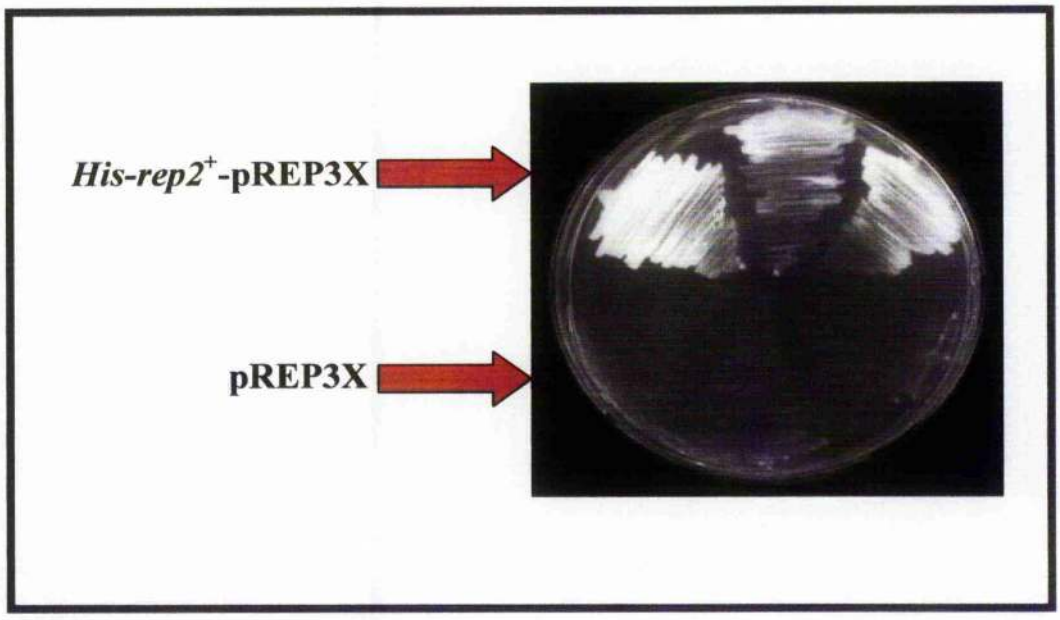


Figure 4.11 Expression of *His-rep2*$^{+}$ rescues the growth defect of the Δ*res1* mutant at 21°C

The Δ*res1* strain (GG 146) was transformed with pREP3X (GG 796) or *His-rep2*$^{+}$-pREP3X (GG 817). Transformants were allowed to grow at 30°C on EMM + thiamine (5 μg μl^{-1}) then streaked out onto the same medium and incubated at 21°C for 6-8 days.

4.5 Biological activity assays – rescuing DSC1 binding activity

The expression of either the wild type *res1*$^+$ or *res2*$^+$ cDNAs in plasmids has the ability to restore *in vitro* DSC1 DNA-binding activity to either Δ*res1* or Δ*res2* mutants, respectively, that is otherwise absent in these strains (Zhu et al., 1997). Expression of the recombinant *His-res1*$^+$ or *His-res2*$^+$ cDNAs in such deletion strains would be expected to give similar results, if the tags do not affect their function. Therefore, this assay was used as another way to test the activities of the recombinant His-Res1p and His-Res2p proteins.

The electrophoretic mobility-shift assay (EMSA), often referred to as the gel retardation or band-shift assay, allows detection of sequence specific DNA/protein interactions *in vitro*. This technique is based on the effect of bound protein(s) on the electrophoretic mobility of a radio-labelled DNA fragment. DNA molecules are naturally negatively charged and therefore migrate toward the positive electrode in an electric field. Furthermore, in an acrylamide gel, DNA molecules are separated according to their size, small molecules having the highest mobility and vice versa. In addition, protein bound to a particular DNA fragment will further retard its mobility. In this assay a radio-labelled DNA fragment (or 'probe'), of specific sequence, is incubated with a protein extract from the cells under investigation. Any protein(s) capable of binding to the DNA fragment will form a complex and the DNA/protein mix is then electrophoresed on a non-denaturing acrylamide gel. The radio-labelled DNA fragment is then visualised by autoradiography and unbound DNA, which runs proportional to its size, appears at the bottom of the gel. However, the mobility of any DNA bound by protein is retarded, causing a visible 'band-shift', relative to the unbound probe. Furthermore, following detection of such a band-shift activity, the specificity of interaction between the DNA fragment and the protein(s) can be analysed. For example, in a 'self-competition' experiment, addition of an identical (but non radio-labelled) DNA fragment will compete with the radio-labelled DNA for protein binding, if binding is specific to that sequence. Addition of increasing amounts of such 'cold-competitor' DNA, results in loss of the radioactive signal by titration, and so the band-shift is diminished and is no longer visible, following autoradiography.

A more stringent test of binding specificity is possible using a mutated version of the DNA fragment. Indeed, alteration of as few as one or two base pairs, can result in loss of binding, such is the specificity often observed by sequence specific DNA-binding proteins. Therefore, such tests of specificity allow distinction between complexes formed by specific and non-specific DNA-binding proteins.

The EMSA assay was used to originally identify the DSC1 complex in fission yeast (Lowndes et al., 1992). In these experiments the DNA fragments used as probes were derived from sections of the *cdc22*$^+$ promoter, that contain MCB sequences in two clusters, with one cluster named MCB1 and the other MCB2 (Maqbool et al, 2003). In the following assays a

DNA probe containing the MCB1 cluster was used, and a mutated MCB1m probe was used to test specificity. Using protein extracted from Δ*res1* cells, transformed with *His-res1*$^{+}$-pREP3X, *His-res2*$^{+}$-pREP1 or pREP3X, and Δ*res2* cells, transformed with *His-res1*$^{+}$-pREP3X, *His-res2*$^{+}$-pREP1 or pREP3X (Appendix I) the ability of His-Res1p or His-Res2p to reconstitute the DSC1 band-shift activity was assayed.

For protein extraction, 5 ml cultures were grown to mid-log phase at 30°C, with the appropriate amino acids, in the presence (for *nmt1*$^{+}$ promoter 'off' experiments) or absence (for *nmt1*$^{+}$ promoter 'on' experiments) of thiamine (5 μg μl^{-1}). Cultures (200 ml), in the presence or absence of thiamine (5 μg μl^{-1}), were then inoculated with the 5 ml pre-cultures and grown for 16 h. Protein was then extracted for use in EMSA experiments as detailed (Methods 2.6.5). Figures 4.12-4.14 show the results of these EMSA assays.

Upon expression of His-Res2p in Δ*res2* cells a 'DSC1-like' band-shift activity, comparable to the DSC1 complex formed in wild-type extracts, was present. Importantly, this band-shift was not observed in Δ*res2* cells alone or Δ*res2* cells containing empty vector. Furthermore, this complex bound specifically to the MCB1 element, as judged by self-competitor analysis and an inability to bind the mutated MCB1m sequence (Figures 4.12-4.13). However, this complex did not super-shift using an antibody directed against the His-tag (data not shown). In contrast to His-Res2p in Δ*res2* cells, a DSC1-like band-shift activity was not detected when His-Res1p was expressed in Δ*res1* cells, either in the absence or presence of thiamine (Figure 4.14).

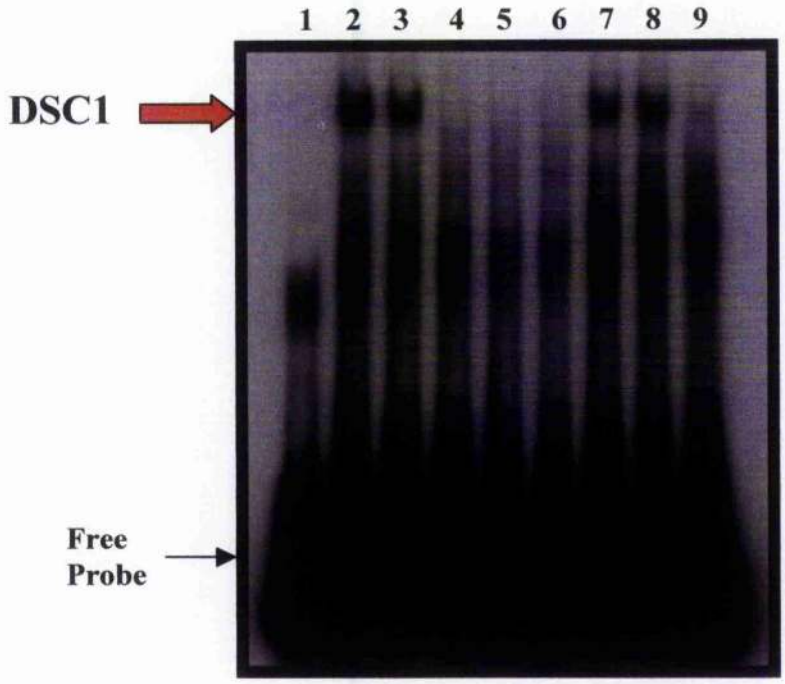


Figure 4.12 'DSC1-like' complex is detected in Δ*res2* cells expressing His-Res2p

EMSA analysis of protein extracts prepared from wild-type cells (GG 217), Δ*res2* cells (GG 156), Δ*res2* cells transformed with pREP3X (GG 801) or Δ*res2* cells transformed with *His-res2*$^{+}$-pREP1 (GG 808). Protein extracts were prepared and incubated with a radio-labelled probe derived from the *cdc22*$^{+}$ promoter (MCB1). To confirm the specificity of the retarded complex an excess of unlabelled MCB1 probe was added to lane 9, GG 808. The upper arrow indicates the position of the putative DSC1 complex and the lower arrow denotes unbound radio-labelled probe.

1 = free MCB1 probe

2 = wild-type protein extract + MCB1 probe

3 = wild-type protein extract + MCB1 probe

4 = Δ*res2* protein extract + MCB1 probe

5 = Δ*res2* : pREP3X protein extract + MCB1 probe

6 = Δ*res2* : pREP3X protein extract + MCB1 probe

7 = Δ*res2* : *His-res2*$^{+}$-pREP1 protein extract + MCB1 probe

8 = Δ*res2* : *His-res2*$^{+}$-pREP1 protein extract + MCB1 probe

9 = Δ*res2* : *His-res2*$^{+}$-pREP1 protein extract + excess unlabelled MCB1 probe

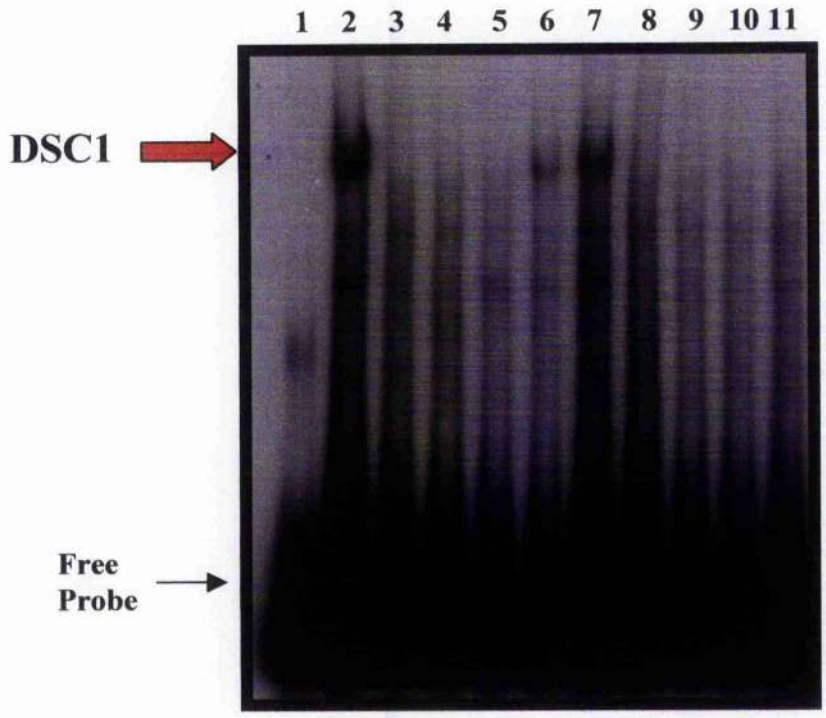


Figure 4.13 'DSC1-like' complex is detected in Δ*res2* cells expressing His-Res2p

EMSA analysis of protein extracts prepared from wild-type cells (GG 217), Δ*res2* cells (GG 156), Δ*res2* cells transformed with pREP3X (GG 801) or Δ*res2* cells transformed with *His-res2*$^{+}$-pREP1 (GG 808) grown in either the presence or absence of thiamine. Protein extracts were prepared and incubated with a radio-labelled probe derived from the *cdc22*$^{+}$ promoter (MCB1) or with a probe containing a mutated MCB1 (MCB1m) element to check the specificity of the retarded complex. The upper arrow indicates the position of the DSC1 complex and the lower arrow denotes unbound radio-labelled probe.

1 = free MCB1 probe

2 = wild-type protein extract + MCB1 probe

3 = wild-type protein extract + MCB1m probe

4 = Δ*res2* protein extract + MCB1 probe

5 = Δ*res2* : pREP3X protein extract + MCB1 probe

6 = Δ*res2* : *His-res2*$^{+}$-pREP1 protein extract (+ thiamine) + MCB1 probe

7 = Δ*res2* : *His-res2*$^{+}$-pREP1 protein extract + MCB1 probe

8 = Δ*res2* protein extract + MCB1m probe

9 = Δ*res2* : pREP3X protein extract protein extract + MCB1m probe

10 = Δ*res2* : *His-res2*$^{+}$-pREP1 protein extract (+ thiamine) + MCB1m probe

11 = Δ*res2* : *His-res2*$^{+}$-pREP1 protein extract + MCB1m probe

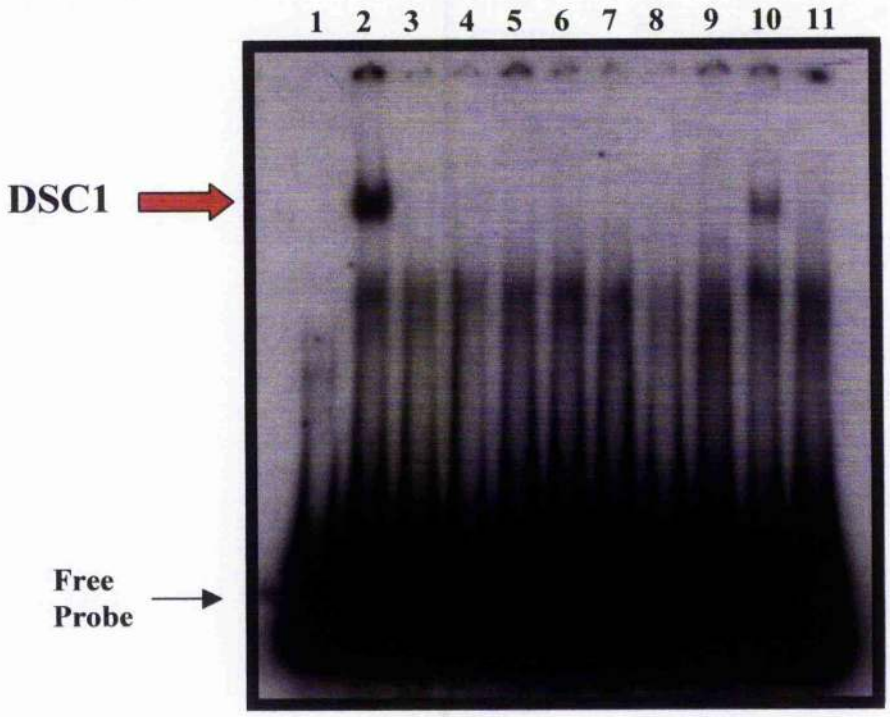


Figure 4.14 'DSC1-like' complex is not detected in Δ*res1* cells expressing His-Res1p

EMSA analysis of protein extracts prepared from wild-type cells (GG 217), Δ*res1* cells (GG 146), Δ*res1* cells transformed with pREP3X (GG 796) or Δ*res1* cells transformed with *His-res1*$^+$-pREP3X (GG 790) or *His-res2*$^+$-pREP1 (GG 802) grown in the presence or absence of thiamine. Protein extracts were prepared and incubated with a radio-labelled probe derived from the *cdc22*$^+$ promoter (MCB1). Protein extracts prepared from Δ*res2* cells (GG 156) and Δ*res2* cells transformed with *His-res2*$^+$-pREP1 (GG 808) grown in the presence of thiamine were incubated with the same probe and with an excess of unlabelled MCB1 probe for comparison. The upper arrow indicates the position of the DSC1 complex and the lower arrow denotes unbound radio-labelled probe.

1 = free MCB1 probe

2 = wild-type protein extract + MCB1 probe

3 = Δ*res1* protein extract + MCB1m probe

4 = Δ*res1* : pREP3X protein extract + MCB1 probe

5 = Δ*res1* : *His-res1*$^+$-pREP3X protein extract (+ thiamine) + MCB1 probe

6 = Δ*res1* : *His-res1*$^+$-pREP3X protein extract + MCB1 probe

7 = Δ*res1* : *His-res2*$^+$-pREP1 protein extract (+ thiamine) + MCB1 probe

8 = Δ*res1* : *His-res2*$^+$-pREP1 protein extract + MCB1 probe

9 = Δ*res2* protein extract protein extract + MCB1 probe

10 = Δ*res2* : *His-res2*$^+$-pREP1 protein extract (+ thiamine) + MCB1 probe

11 = Δ*res2* : *His-res2*$^+$-pREP1 protein extract (+ thiamine) + excess unlabelled MCB1 probe

4.6 Discussion

In this chapter a series of assays have been described, designed to assess the effect of an N-terminal His-tag on the biological activities of the SpRes1p, SpRes2p, SpRep1p and SpRep2p proteins of the fission yeast DSC1 complex. To this end, a cloning strategy was designed allowing directional cloning of the respective recombinant cDNAs into the fission yeast overexpression vector pREP.

The recombinant cDNAs, *His-res1*$^+$, *His-res2*$^+$, *His-rep1*$^+$ and *His-rep2*$^+$ were able to rescue the cold-sensitivity of a Δ*res1* mutant at 21°C (Figures 4.8-4.11). Notably, this rescue was most efficient when cells were grown in the presence of thiamine. As outlined in Section 4.2, genes under the control of the *nmt1*$^+$ promoter may still be expressed despite the presence of the transcriptional repressor thiamine. With respect to *His-res1*$^+$, *His-res2*$^+$, *His-rep1*$^+$ and *His-rep2*$^+$ suppression of the cold-sensitive lethality of a Δ*res1* mutant was possible at levels of transcription achieved in the presence of thiamine. This was not surprising given that the rescue of a Δ*res1* mutant was achieved by expression of the wild-type *res1*$^+$ cDNAs from the SV40 promoter (Tanaka et al., 1992). The level of transcription from the pREP vector, in the presence of thiamine, is comparable with that seen from an induced SV40 promoter (Forsburg, 1993).

With respect to His-Res2p, reconstitution of a 'DSC1-like' band-shift activity has been demonstrated upon expression of this protein in a Δ*res2* mutant (Figures 4.12-4.13). In comparison, the DSC1 band-shift activity was lost in protein extracts prepared from the Δ*res2* strain (Figures 4.12-4.13). This 'DSC1-like' band-shift activity was reduced in protein extracts prepared from Δ*res2* cells expressing His-Res2p in the presence of thiamine (Figure 4.13). Additionally, this DSC1-like complex bound specifically to the MCB1 element, was competed by addition of an identical non radio-labelled probe and did not bind to a mutated MCB1m sequence. Thus, from these results, it is likely that the 'DSC1-like' complex observed in Δ*res2* cells expressing His-Res2p is similar to the wild-type DSC1 complex.

In contrast, expression of His-Res2p in Δ*res1* cells did not reconstitute the DSC1 band-shift complex nor did expression of His-Res1p Δ*res2* cells (Figure 4.14 and data not shown). Despite expression of His-Res1p in the Δ*res1* mutant background, a 'DSC1-like' band-shift activity was not observed (Figure 4.14). Intriguingly, although the DSC1 complex was not detected *in vitro* when His-Res2p was expressed in Δ*res1* cells, both the heat and cold sensitive lethality of these cells was rescued by expression of this protein from the pREP vector (Figure 4.9 this study, and Zhu et al., 1997). Therefore, as suggested by Zhu and co-workers, detection of the DSC1 complex *in vitro* does not correlate with loss of transcriptional activity *per se*.

Nevertheless, the successful reconstitution of the DSC1 band-shift in Δ*res1* cells expressing SpRes1p has been reported (Caligiuri and Beach, 1993) and failure to do so in this study may reflect inclusion of the His-tag. The ability of the His-Res1p protein to efficiently

rescue the cold sensitive lethality of a *Δres1* mutant suggests that inclusion of the His-tag does not affect function, but rather some other property such as protein stability. For example, levels of the His-Res1p protein may be at a level sufficient to rescue the *Δres1* cells and form an active DSC1 complex *in vivo*, but this complex cannot be detected *in vitro*. Indeed, despite the efficient rescue of the cold-sensitive lethality of *Δres1* cells by all of the recombinant proteins (indicating they are all expressed and functional) none have been detected by immunoblotting using an antibody directed against the His-tag. In addition, a super-shift of the DSC1-like complex, reconstituted by expression of *His-res2*$^{+}$-pREP1in *Δres2* cells, was not detected using this same antibody.

The inability to detect the recombinant proteins, or to super-shift the 'DSC1-like' complex containing His-Res2p, may be a result of cleavage at the N-terminus and thus loss of the His-tag. The presence of a thrombin cleavage site immediately downstream of the His-tag supports this explanation. However, we are unaware of any fission yeast thrombin homologue(s), and given that the cleavage process is a highly specific sequence-dependent reaction, a degenerate protease would be unlikely to cleave at this site. Importantly, the use of a recombinant protein expression system in fission yeast using tags containing thrombin and enterokinase cleavage sites has been reported (Lu et al. 1997). Also, the function of both SpRes1p and SpRes2p requires an intact N-terminus (Tanaka et al., 1992; Ayte et al., 1995; Zhu et al., 1997).

Whilst the His-tag has no apparent effects on the function of these proteins, the incorporation of the tag may affect their stability and expression level, such that they are not detectable by immunoblotting. In addition, with respect to the super-shift experiment, the His-tag is at the N-terminus. Given that binding of both SpRes1p and SpRes2p to DNA is mediated by their N-termini, it is possible that the His-tag may be masked by interaction with the DNA and/or components of the DSC1 complex and so is occluded from binding to the monoclonal antibody.

Even so (for reasons outlined in Chapter 3), it is our intention to study the structures of these recombinant proteins when produced in *E. coli*. Differences in protein production between eukaryotes and prokaryotes, particularly in post-translational modification, are well documented (Sudbery, 1996). Such differences may be manifested in alterations of protein structure and activity. It is encouraging that the results described in this chapter indicate these recombinant proteins can function normally in fission yeast (in comparison to their wild type counterparts). As the presence of the His-tag appears to impart no obvious impediment to the function of these proteins *in vivo*, it is likely that its inclusion does not significantly alter their structure. Having assessed the effect of the His-tag on the function of these proteins *in vivo*, in the following chapter, the recombinant proteins, when produced in *E. coli*, are analysed to

assess their solubility and ultimately their biological activity. As discussed in Chapter 3, the expression of soluble recombinant proteins in *E. coli* suggests they are capable of folding into their native or near-native state, and that the tag does not disrupt folding or expression. Following production of soluble protein the biological activity of the recombinant protein may then be assayed.

Chapter 5

Biological activity assays of recombinant DSC1 proteins *in vitro*

5.1 Introduction

Chapter 4 focused on the analysis of the biological activity of recombinant DSC1 components, when produced *in vivo*. The results of these assays indicated that addition of the N-terminal His-tag (required for affinity purification) did not affect the cellular function of these proteins. Therefore, it is unlikely that the presence of this tag alters protein structure, to the detriment of function.

Detailed structural information is obtained through biophysical studies using techniques such as X-ray crystallography, nuclear magnetic resonance (NMR), circular dichroism (CD), surface plasmon resonance (SPR) and isothermal titration calorimetry (ITC). With the exception of SPR, these techniques require milligram quantities of protein. To obtain sufficient quantities of material for such studies, the fission yeast DSC1 components have been over-expressed as His-tagged fusion proteins, using *E. coli* as a heterologous host. Despite the apparent activity of the recombinant His-Res1p, His-Res2p, His-Rep1p and His-Rep2p proteins *in vivo*, it is still necessary to ensure that these proteins, when produced *in vitro*, are biologically active, as this will be the source for structural analysis.

To date, the fission yeast DSC1 components have been poorly characterised in terms of 3-dimensional structure. These studies have been limited to secondary structure predictions based on sequence alignments and homology with their budding yeast counterparts. Significant information has been obtained regarding the domain architecture of these proteins, revealing the regions that are important for their function. Much of this information has been obtained through the *in vivo* analysis of mutated, truncated and hybrid proteins. In addition to this, studies of *in vitro* translated and bacterially expressed proteins have also contributed to understanding about their functions.

Analysis of truncated and mutant SpCdc10p proteins *in vivo* has indicated that the central ankyrin repeats and C-terminal region are important for function. The C-terminal region of SpCdc10p has been implicated in protein-protein interactions with itself and/or SpRes1p (Reymond and Simanis, 1993). The majority of *cdc10* temperature sensitive mutants contain lesions that map to either the central ankyrin repeats, or within the C-terminus (Reymond et al., 1992), and the C-terminally truncated SpCdc10-C4p mutant, displays increased MCB-regulated gene expression *in vivo*, implying an important regulatory function for this region (McInerny et al., 1995). Together, these results define important roles for both the centrally located ankyrin repeats and the C-terminus of SpCdc10p.

The Cdc10p protein is unable by itself to bind to MCB-DNA *in vitro* (Zhu et al., 1994; Ayte et al., 1995). However, it is required for the MCB-specific DNA binding of both SpRes1p and SpRes2p. EMSA analysis using *in vitro* translated protein has revealed that SpRes1p cannot bind MCB-DNA in the absence of SpCdc10p, although SpRes2p does so very weakly

(Zhu et al., 1997). In contrast, bacterially expressed SpRes1p binds specifically to MCB-DNA on its own, although this was dependent on an N-terminal fusion partner (Ayte et al., 1995; Ayte et al., 1997). Both SpRes1p and SpRes2p share significant homology in their N-termini and in the centrally located ankyrin repeats, but diverge in their C-terminal region. Whilst SpRes1p and SpRes2p can functionally substitute for each other with respect to their mitotic roles, SpRes1p is unable to mimic the meiotic role of SpRes2p (Sturm and Okayama, 1996).

Analysis of hybrid proteins, generated by fusing together different domains of SpRes1p and SpRes2p, has revealed that the functional specificity of both SpRes1p and SpRes2p is dictated almost entirely by their C-terminal region (Sturm and Okayama, 1996; Zhu et al., 1997). The meiotic-specific function of SpRes2p lies within this C-terminal region in addition to a domain required for interaction with SpRep2p (Sturm and Okayama, 1996). Similarly, yeast two-hybrid analysis has revealed the domains responsible for the transcriptional activation and SpRes2p-binding activities of SpRep2p (Tahara et al., 1998). Binding assays have shown that bacterially expressed GST-Rep2p binds to *in vitro* translated SpRes2p (Nakashima et al., 1995; Sturm and Okayama, 1996). SpRes2p binds to bacterially expressed GST-Res2p, GST-Cdc10p and GST-Res1p *in vitro*, whereas SpRes1p binds only GST-Res2p and GST-Cdc10p, but not GST-Res1p (Ayte et al., 1997). Notably, however, there have been no reported studies on the SpRep1p protein, either *in vivo* or *in vitro*, and as such the functional properties and domain architecture of this protein remain poorly characterised.

Using the EMSA technique, the ability (and therefore the biological activity) of bacterially produced His-Res1p and His-Res2p to bind to MCB elements *in vitro* can be assayed. The biological activity of bacterially produced His-Cdc10p may be similarly assayed, following successful co-expression with either His-Res1p or His-Res2p. In contrast, assay of the biological activity of both SpRep1 and SpRep2p is not so facile. The absence of any GST-fusions of these DSC1 components precludes binding assays with bacterially expressed His-Rep2p or His-Rep1p. The ability of either protein to participate in an *in vitro* band-shift complex with either His-Res1p and/or His-Res2p may be assessed. Indeed, following expression and/or co-expression of the respective DSC1 components it might be possible to reconstitute a recombinant DSC1 complex *in vitro*. An *in vitro* reconstituted complex has the potential to provide an ideal model system for structural and functional studies.

In order to assess the biological activity of the recombinant DSC1 components they must firstly be solubilised following successful overexpression. For experimental analysis of a protein it is essential that, following release from the cell, it be released into a solvent in which it is stable. Upon release from its native environment, a protein is exposed to foreign conditions and agents that may cause irreversible damage. The nature of the solvent is therefore critical. Exposure to proteolytic enzymes, oxidising agents or extremes of temperature and pH can all have adverse effects on the native structure of a protein. Also, due to the multiple acid-base

groups present in the amino acid side-chains, the structure and thus solubility of a protein will also depend on the ionic composition/polarity of the solvent. As a result, proteins are routinely dissolved in buffer solutions, effective over a pH range at which the protein is stable. The addition of protease inhibitors and reducing agents can help to maintain the integrity of a protein in solution and procedures of protein preparation and purification are routinely carried out on ice, as many proteins are thermally labile.

The unique 3-D arrangement of the constituent amino acids of a protein dictates its physicochemical properties. Thus, the optimal conditions in which a protein is both soluble and stable, following release into the extracellular milieu, must be determined empirically. With respect to heterologous expression systems, the host strain and growth conditions can equally affect the ability to obtain soluble protein following overexpression. For example, the choice of host strain can affect the folding, and thus structure, of the protein expressed. Manipulation of induction conditions can affect the rate and level of protein synthesis that in turn can affect solubility. Consequently, the preparation of soluble recombinant protein can often be one of the lengthiest stages in a purification strategy.

The aim of this chapter, therefore, was to assay the biological activity of the recombinant DSC1 components when produced *in vitro*. Following successful overexpression (Chapter 3), the solubility of each protein is discussed followed by a series of experiments designed to test their *in vitro* biological activity using EMSA.

5.2 Solubility of the recombinant DSC1 proteins

The successful cloning and overexpression of the full-length recombinant His-tagged DSC1 component polypeptides (His-Res1p - 75 kDa, His-Res2p - 75 kDa, His-Rep1p - 55 kDa, His-Rep2p - 30 kDa and His-Cdc10p - 90 kDa), in *E. coli* has been described in Chapter 3. For experimental analysis, these recombinant proteins must be both soluble and stable following release from bacteria.

Lysis of *E. coli* cells is routinely achieved by sonication, enzymatic lysis or mechanical disruption. The mechanical disruption method was favoured for this work, using a French Pressure apparatus. The French Press was chosen as it was extremely reproducible and the most convenient method available, particularly when managing large volumes of bacterial culture. The solubility of all proteins described in this chapter was routinely assessed (Methods 2.4.12). Following overexpression, cells were harvested by centrifugation and the pellet re-suspended in an appropriate volume of lysis buffer. Cells were then disrupted by passage through an automatic French pressure cell. Subsequently, an aliquot of this whole cell extract, together with samples of supernatant and pellet (obtained following centrifugation), were analysed by SDS-polyacrylamide gel electrophoresis. The solubility of each protein was determined by comparing the ratio present in the supernatant/pellet, following Coomassie blue staining or immunoblotting.

As discussed previously, the ability to produce soluble protein is dependent on multiple factors. Several parameters were altered in an attempt to achieve optimal conditions for the production of soluble protein, for each of the recombinant DSC1 components. Despite manipulation of lysis buffer composition (i.e. pH, salt concentration and buffer salt) no significant change was observed, with respect to the solubility of His-Res1p, His-Rep1p, His-Rep2p or His-Cdc10p. Consequently, these proteins were routinely dissolved in 50 mM KH_2PO_4, pH 7.5 (the optimal buffer conditions achieved). Figures 5.1-5.4, therefore allow a comparison of soluble protein production under 4 different induction conditions. Figures 5.1, 5.2 and 5.5 show the solubility of His-Rep1p, His-Cdc10p and His-Res2p following overexpression at different temperatures as viewed by Coomassie blue staining. Figures 5.3 and 5.4 show the solubility of His-Res1p and His-Rep2p at different temperatures as viewed by immunoblotting. Immunoblot analysis of both His-Res1p and His-Rep2p was necessary due to the low expression level of these proteins (Chapter 3).

His-Res1p was detected in the soluble fraction at 37°C (Figure 5.3 B) whereas His-Rep2p was completely insoluble following induction at this temperature (Figure 5.4 B). Both His-Res1p and His-Rep2p were partially soluble at all other temperatures tested (Figures 5.3 and 5.4). The fact that expression of these latter two proteins was detectable only by immunoblotting (under all conditions tested) reflects their low level of expression. Such low-

level expression may account for detection of both proteins in the soluble fraction, under most conditions tested. However, the low yield of these proteins imposes restrictions on their use for detailed structural work although the observation that at least 50% of these proteins were soluble was encouraging. In contrast, the high levels of protein obtained following overexpression of His-Rcp1p and His-Cdc10p may contribute to their insolubility. High-level overexpression of proteins can result in the physiological solubility limit of the cell being exceeded. Both of these proteins were insoluble and appeared in the pellet under all of the conditions tested (Figures 5.1 and 5.2).

Figure 5.5 shows that at 15°C, approximately 100 % of His-Res2p was recovered in the supernatant. Importantly, the ability to produce soluble His-Res2p was not solely reliant on manipulation of the growth conditions. In addition, the nature of the lysis buffer was crucial. Indeed, the production of soluble His-Res2p highlights the delicate balance of conditions required to obtain soluble protein (Section 5.3). Even so, assuming that this protein is correctly folded and active, following a scaling up in the process, structural studies should now be possible.

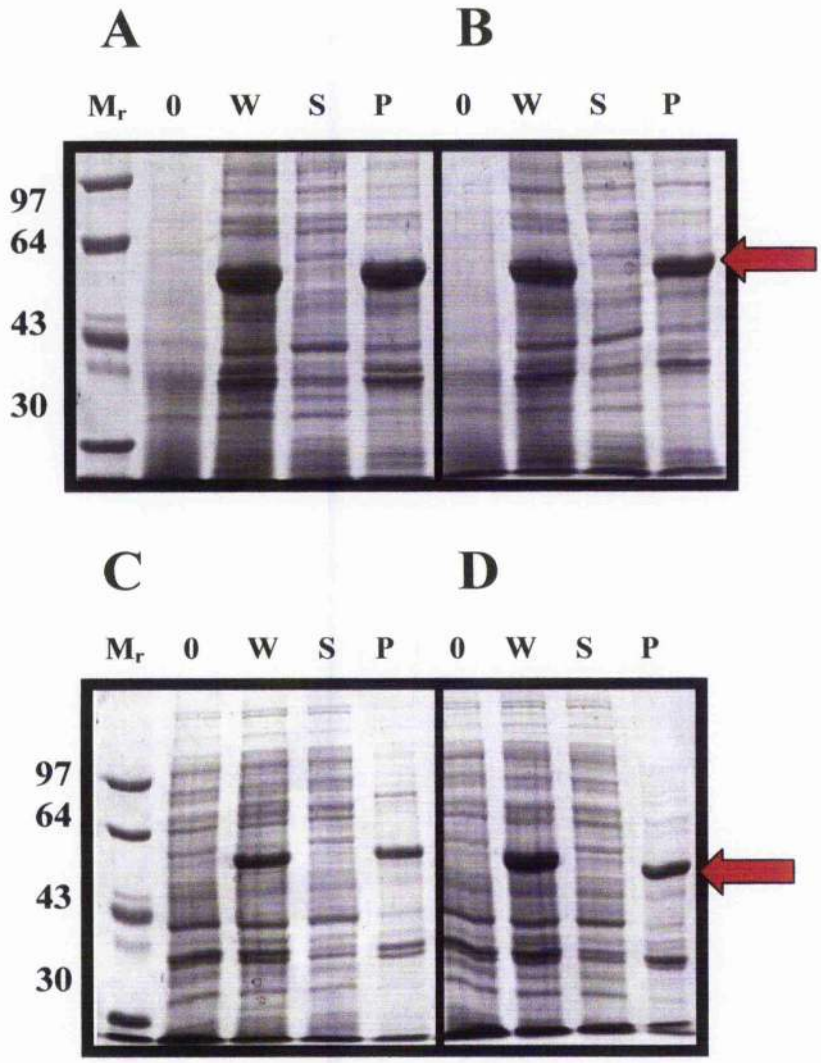


Figure 5.1 SDS-PAGE analysis of His-Rep1p solubility

His-Rep1p was overexpressed in BL21 CodonPlus (DE3)-RIL cells for 3 h at either 30˚C (**A**) or 37˚C (**B**) by induction with 1 mM IPTG. Alternatively overexpression was carried out at 15˚C (**C**) or 22˚C (**D**) by induction with 0.2 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h). The bacterial culture was then centrifuged, the pellet resuspended in 5 ml lysis buffer, and cells lysed by French pressure disruption. This whole cell extract (W) was then separated into soluble (S) and insoluble (P) fractions by centrifugation. The pellet was resuspended in 5 ml lysis buffer. The 0 h samples were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl/0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). Samples (1 ml) of the whole cell extract, the soluble fraction and the resuspended pellet were diluted with an equal volume of Laemmli sample buffer and similarly denatured. Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. Molecular weight markers are shown (M_r) with sizes indicated in kDa.

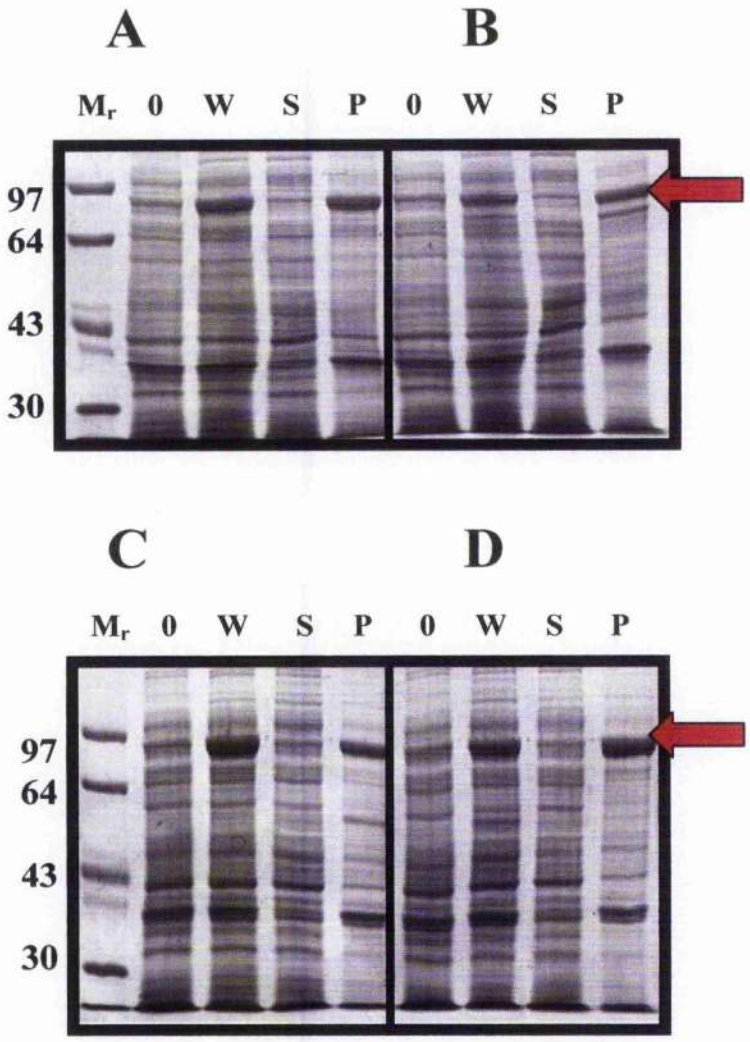


Figure 5.2 SDS-PAGE analysis of His-Cdc10p solubility

His-Cdc10p was overexpressed in BL21 CodonPlus (DE3)-RIL cells for 3 h at either 30°C (**A**) or 37°C (**B**) by induction with 1 mM IPTG. Alternatively overexpression was carried out at 15°C (**C**) or 22°C (**D**) by induction with 0.2 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h). The bacterial culture was then centrifuged, the pellet resuspended in 5 ml lysis buffer, and cells lysed by French pressure disruption. This whole cell extract (W) was then separated into soluble (S) and insoluble (P) fractions by centrifugation. The pellet was resuspended in 5 ml lysis buffer. The 0 h samples were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl/0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). Samples (1 ml) of the whole cell extract, the soluble fraction and the resuspended pellet were diluted with an equal volume of Laemmli sample buffer and similarly denatured. Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. Molecular weight markers are shown (M_r) with sizes indicated in kDa.

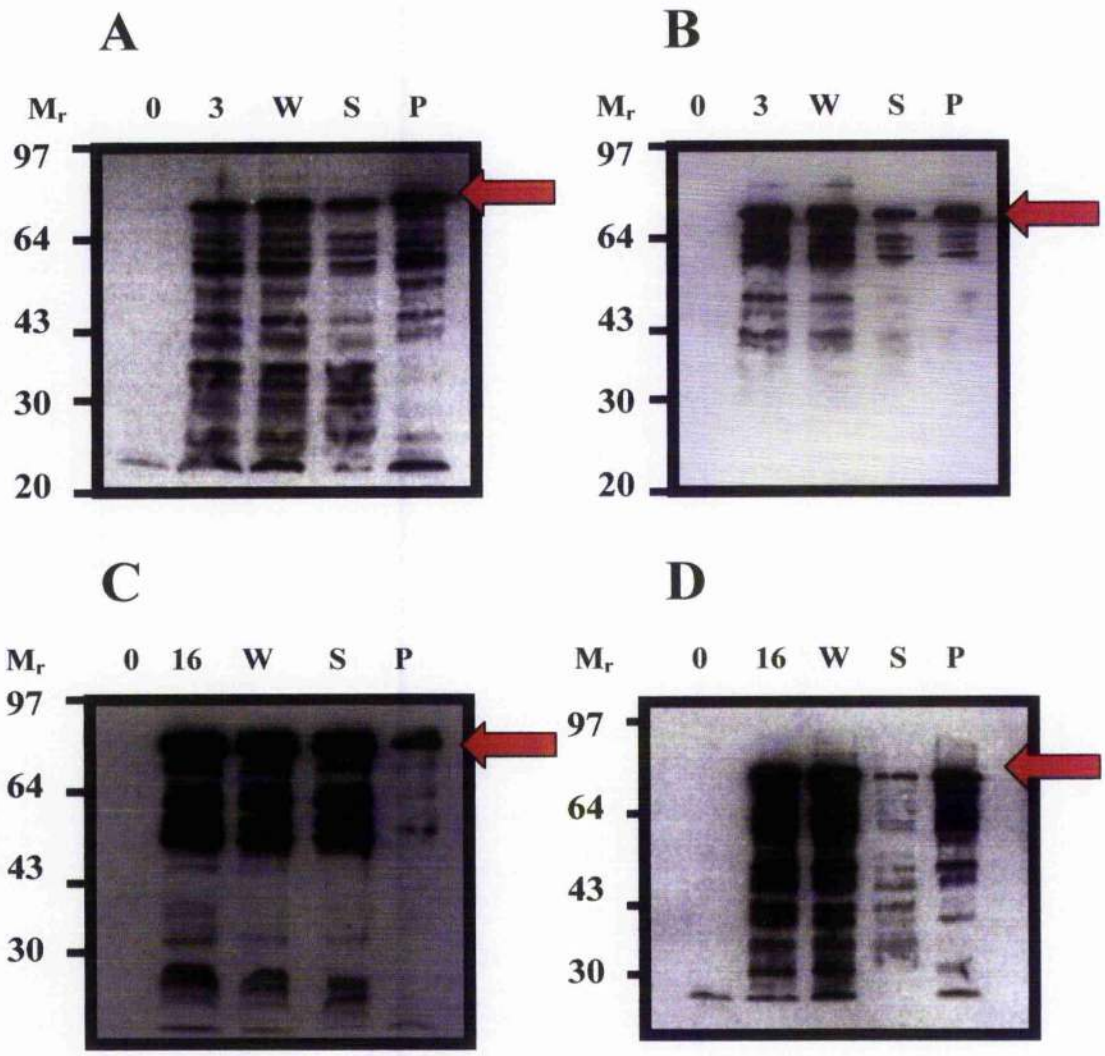


Figure 5.3 Immunoblot analysis of His-Res1p solubility

His-Res1p was overexpressed in BL21 CodonPlus (DE3)-RIL cells for 3 h at either 30°C (**A**) or 37°C (**B**) by induction with 1 mM IPTG. Alternatively overexpression was carried out at 15°C (**C**) or 22°C (**D**) by induction with 0.2 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at 3 h or 16 h. The bacterial culture was then centrifuged, the pellet resuspended in 5 ml lysis buffer, and cells lysed by French pressure disruption. This whole cell extract (W) was then separated into soluble (S) and insoluble (P) fractions by centrifugation and samples (1 ml) taken. Samples from 0 h and 3 or 16 h were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl/0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). All other samples were diluted with an equal volume of Laemmli sample buffer and similarly denatured. Samples were then analysed by immunoblotting as described (Methods 2.4.7-2.4.11). Blots were probed with Anti-PentaHis-HRP conjugate antibody (Qiagen) at 1:2500. The positions of molecular weight markers are indicated (M_r) in kDa.

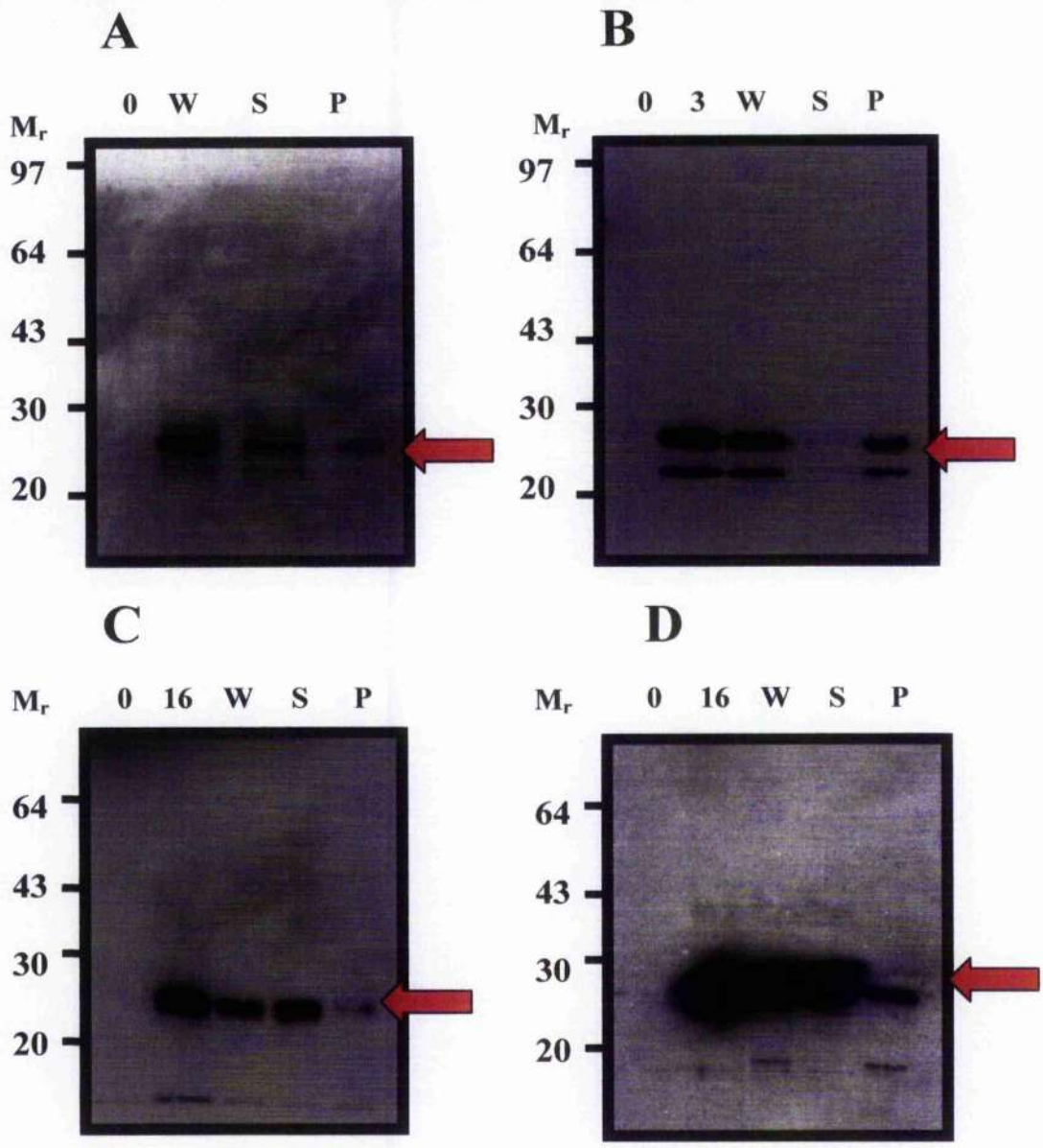


Figure 5.4 Immunoblot analysis of His-Rep2p solubility

His-Rep2p was overexpressed in BL21 CodonPlus (DE3)-RIL cells for 3 h at either 30˚C (**A**) or 37˚C (**B**) by induction with 1 mM IPTG. Alternatively overexpression was carried out at 15˚C (**C**) or 22˚C (**D**) by induction with 0.2 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at 3 h or 16 h. The bacterial culture was then centrifuged, the pellet resuspended in 5 ml lysis buffer, and cells lysed by French pressure disruption. This whole cell extract (W) was then separated into soluble (S) and insoluble (P) fractions by centrifugation and samples (1 ml) taken. Samples from 0 h and 3 or 16 h were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl/0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). All other samples were diluted with an equal volume of Laemmli sample buffer and similarly denatured. Samples were then analysed by immunoblotting as described (Methods 2.4.7-2.4.11). Blots were probed with Anti-PentaHis-HRP conjugate antibody (Qiagen) at 1:2500. The positions of molecular weight markers are indicated (M_r) in kDa

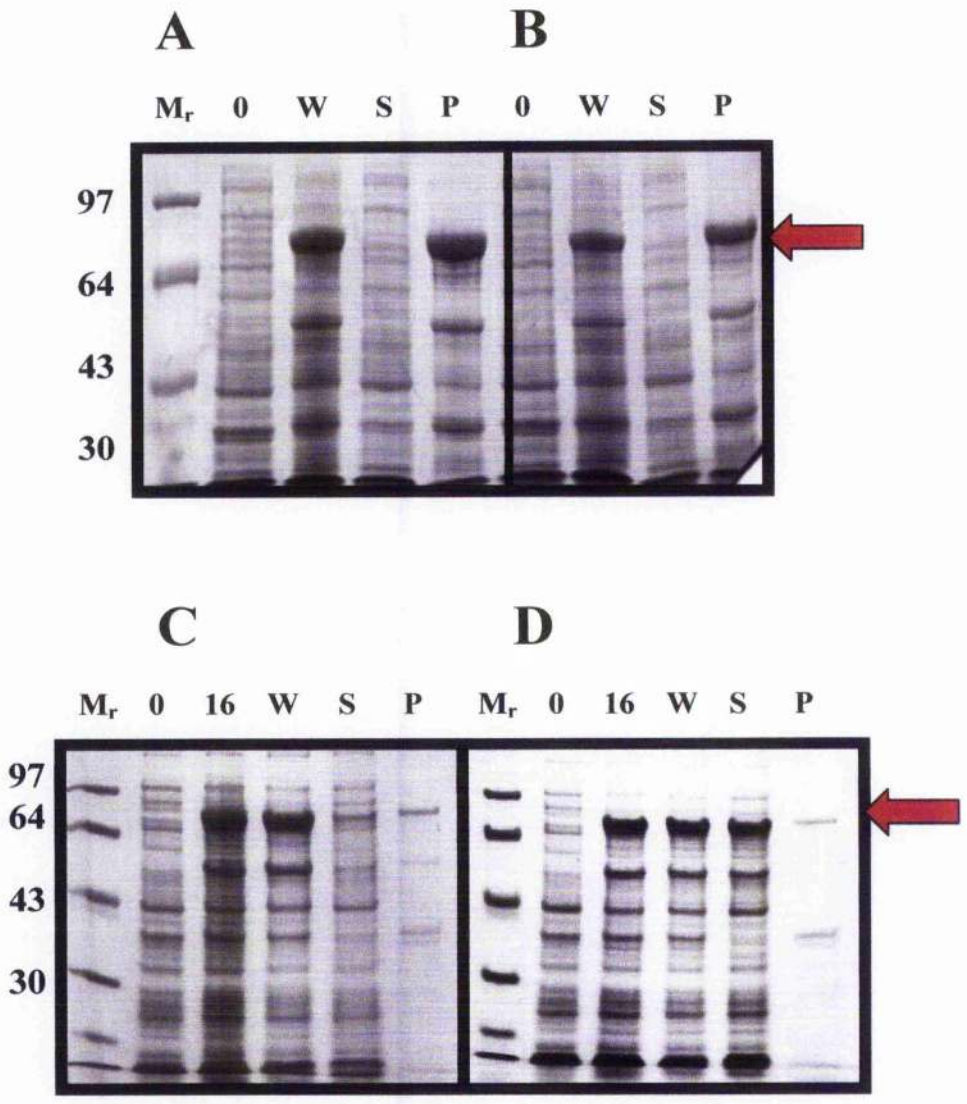


Figure 5.5 SDS-PAGE analysis of His-Res2p Solubility

His-Res2p was overexpressed in BL21 CodonPlus (DE3)-RIL cells for 3 h at either 30°C (**A**) or 37°C (**B**) by induction with 1 mM IPTG. Alternatively overexpression was carried out at 22°C (**C**) or 15°C (**D**) by induction with 0.2 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at 3 h or 16 h. The bacterial culture was then centrifuged, the pellet resuspended in 5 ml lysis buffer, and cells lysed by French pressure disruption. This whole cell extract (W) was then separated into soluble (S) and insoluble (P) fractions by centrifugation and samples (1 ml) taken. Samples from 0 h and 16 h were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl / 0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). All other samples were diluted with an equal volume of Laemmli sample buffer and similarly denatured. Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. Molecular weight markers are shown (M_r) with sizes indicated in kDa.

5.3 The effect of salt concentration on His-Res2p solubility

Lysis buffer composition plays a crucial role in determining both the solubility and stability of a protein following release from the cell. Attempts to enhance the solubility of His-Rep1p and His-Cdc10p were refractory to manipulation of both growth and lysis buffer conditions. In contrast, His-Res2p was significantly affected. Initial attempts to obtain soluble His-Res2p, following overexpression at 37°C or 30°C, coupled with a screening of different lysis buffer conditions, were unsuccessful. A small amount of His-Res2p was retrieved in the supernatant following overexpression at 22°C and overexpression at 15°C yielded significantly more soluble protein and manipulation of lysis buffer composition enhanced this further.

Following overexpression at 15°C, the concentration of salt in the lysis buffer proved pivotal in the ability to obtain soluble His-Res2p, following release from the cell. In lysis buffers containing ≥ 0.3 M NaCl (such concentrations of salt are routinely used in buffers due to the requirements of purification procedures), His-Res2p was almost exclusively present in the insoluble pellet (Figure 5.6 B and C). In contrast, reduction or complete removal of the salt concentration had a profound effect (Figure 5.6 A and D). In buffer solutions with either no or minimal salt (≤ 0.1 M NaCl), His-Res2p was completely soluble (although a minute amount is present in the pellet at 100 mM NaCl). In contrast, even a slight increase to 300 mM NaCl resulted in approximately 100% of expressed protein becoming insoluble; a similar effect was observed following lysis in buffers containing 1M NaCl.

The solubility of His-Res2p was extremely sensitive to salt concentration. This serves as an excellent example of the sensitivity of a protein to its immediate environment. The sensitivity of His-Res2p to high NaCl concentration has implications for the purification of this protein and is discussed in Chapter 6.

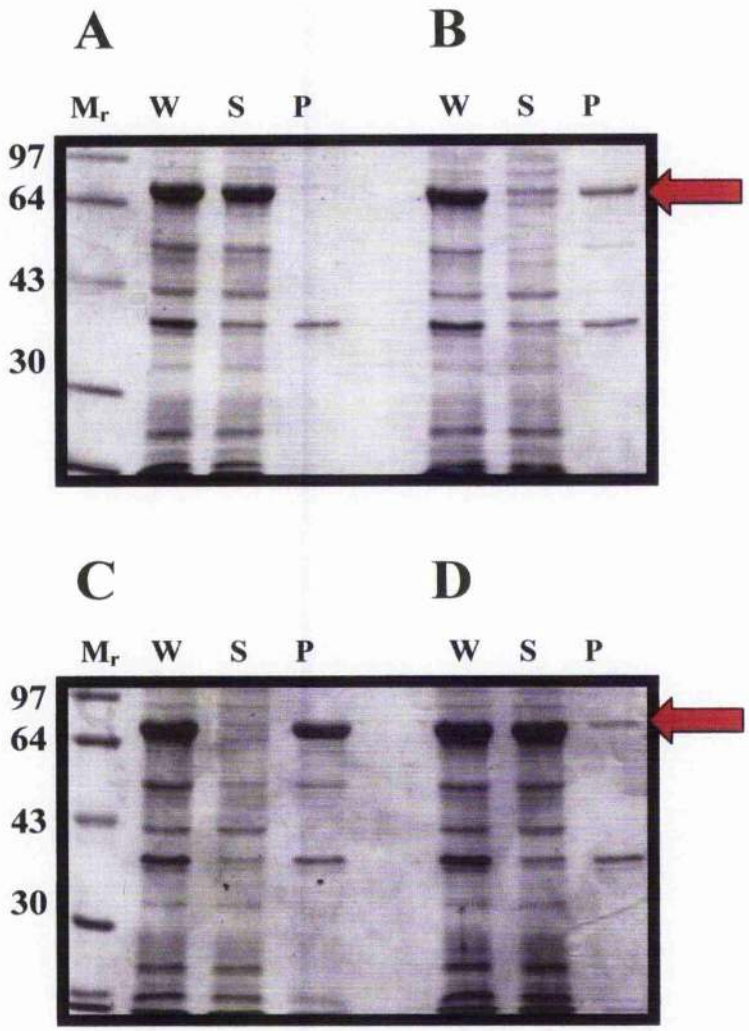


Figure 5.6 Effect of salt concentration on the solubility of His-Res2p

His-Res2p was overexpressed at 15°C in BL21 CodonPlus (DE3)-RIL cells by induction with 0.2 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at 16 h. The bacterial culture was then centrifuged, and the pellet resuspended in 5 ml lysis buffer: (**A**) 50mM KH_2PO_4, pH 7.5; (**B**) **A** + 1M NaCl; (**C**) **A** + 0.3M NaCl and (**D**) **A** + 0.1M NaCl. Cells were then lysed by French pressure disruption. The whole cell extract (W) was separated into soluble (S) and insoluble (P) fractions by centrifugation and samples (1 ml) taken. All samples were centrifuged, the pellets resuspended in an equal volume of Laemmli sample buffer and denatured by boiling for 5 min in the presence of DTT (150 mM). Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. Molecular weight markers are shown (M_r) with sizes indicated in kDa.

5.4 Biological activity assays – His-Res1p and His-Res2p

Both SpRes1p and SpRes2p, when produced *in vitro*, can bind specifically to MCB elements *in vitro* as detected by EMSA. Detection of both SpRes1p and SpRes2p MCB-specific DNA-binding complexes was dependent on co-expression with SpCdc10p, although SpRes2p alone could bind very weakly (Zhu et al., 1994; Zhu et al., 1997).

The ability of the bacterially produced His-Res1p and His-Res2p proteins to bind MCB elements *in vitro* was tested. Soluble extracts from *E. coli* cells, expressing either His-Res1p or His-Res2p (Methods 2.4.1 and 2.4.12), were incubated with a radio-labelled DNA probe and used in EMSA experiments. The DNA probe was identical to that detailed in Chapter 4 and contained the MCB1 cluster derived from the *cdc22*$^{+}$ promoter (Maqbool et al., 2003). The mutated MCB1m probe was again used to test specificity. Despite repeated attempts, binding of either His-Res1p or His-Res2p to MCB1 DNA using the EMSA was not detected under these conditions (Figure 5.7).

However, as mentioned above, binding of either SpRes1p or SpRes2p to MCB DNA *in vitro* was reportedly dependent on and/or enhanced by, the presence of SpCdc10p and so co-expression of His-Cdc10p with His-Res1p or His-Res2p was also attempted. It is possible that these proteins may associate in a co-translational manner *in vivo*. Therefore, co-expression with His-Cdc10p in *E. coli* might replicate this scenario since association of proteins as folding intermediates can increase stability and thus, successful co-expression has the potential to enhance protein solubility.

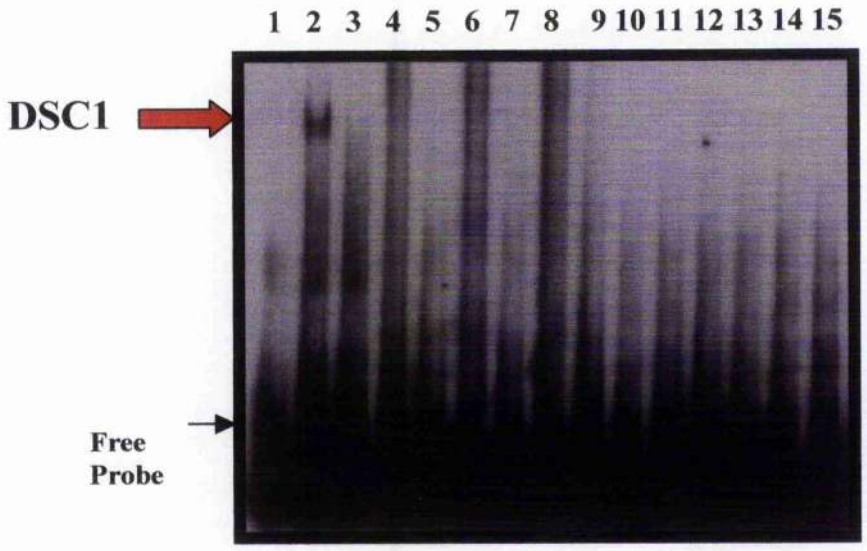


Figure 5.7 His-Res1p and His-Res2p binding to MCB motifs *in vitro*

EMSA analysis of protein extracts prepared from *E. coli* BL21 (DE3) CodonPlus-RIL cells transformed with pET-28c (GB 7), *His-res1*$^{+}$-pET-28c (GB 201) or *His-res2*$^{+}$-pET-28c (GB 178). Protein extracts were prepared and incubated with a radio-labelled probe derived from the *cdc22*$^{+}$ promoter (MCB1). To test the specificity of protein binding to the MCB sequence a probe containing a mutated MCB1 element (MCB1m) was included (lanes 3, 6, 7, 10, 11, 14 and 15). A protein extract from wild-type fission yeast cells (GG 217) was incubated with MCB1 and forms the characteristic DSC1 band-shift (lane 2). The upper and lower arrows on the left indicate the positions of the fission yeast DSC1 complex and unbound radio-labelled probe, respectively.

1 = free MCB1 probe

2 = wild-type protein extract + MCB1 probe

3 = wild-type protein extract + MCB1m probe

4 and 5 = *His-res1*$^{+}$-pET-28c protein extract + MCB1 probe

6 and 7 = *His-res1*$^{+}$-pET-28c protein extract + MCB1m probe

8 and 9 = *His-res2*$^{+}$-pET-28c protein extract + MCB1 probe

10 and 11 = *His-res2*$^{+}$-pET-28c protein extract + MCB1m probe

12 and 13 = pET-28c protein extract + MCB1 probe

14 and 15 = pET-28c protein extract + MCB1m probe

5.5 Cloning of *res1*$^+$, *rep1*$^+$ and *rep2*$^+$

To ensure that two proteins are co-expressed from separate plasmids in the same *E. coli* cell, there must be selection for both plasmids. Both *res1*$^+$ and *res2*$^+$ were cloned into the kanamycin resistant pET-28c vector, whereas *cdc10*$^+$ was cloned into the ampicillin resistant pET-14b (Chapter 3). The pET-14b plasmid is essentially identical to pET-28c, containing an N-terminal His-tag upstream of *Nde*I-*Bam*HI multiple cloning sites. However, in contrast, pET-14b contains the β-lactamase gene, conferring ampicillin resistance. The cloning of *res1*$^+$/*res2*$^+$ and *cdc10*$^+$ into plasmids with different antibiotic resistance markers allows phenotypic selection of co-transformants. Further to such studies using His-Cdc10p, co-expression between other DSC1 components was investigated. To this end, *res1*$^+$ was cloned into pET-14b to allow co-expression studies with Res2p. In addition, *rep1*$^+$ and *rep2*$^+$ were cloned into pET-14b to facilitate co-expression studies with these proteins.

Cloning of these cDNAs into pET-14b was performed as described previously (Chapter 3, section 3.2). Clones containing insert were analysed by restriction enzyme mapping and DNA sequencing (MWG-Biotech – using oligos listed in Appendix II) and a positive clone for each gene was stored at -70°C in the lab bacterial collection: GB 299 (*res1*$^+$), GB 289 (*rep1*$^+$) and GB 310 (*rep2*$^+$) – Appendix I.

Following optimisation of conditions, each clone was successfully over-expressed in the BL21 (DE3) CodonPlus-RIL strain. The expression levels and solubility were comparable to that achieved when these same cDNAs were expressed from the pET-28c vector: both His-Res1p and His-Rep2p were detectable only by immunoblot and were soluble, whereas His-Rep1p was detected by Coomassie blue staining and was insoluble (data not shown).

5.6 Co-expression of recombinant DSC1 components in *E. coli*

Following co-transformation (Methods 2.2.13) into BL21 (DE3) CodonPlus-RIL cells and small-scale protein inductions, the ability of two proteins to co-express was determined by Coomassie blue staining or immunoblotting, following SDS-PAGE.

Upon testing different combinations of DSC1 components, only co-expression of His-Res2p with His-Rep2p was detected: only His-Res2p was visualised by Coomassie staining whereas both proteins were detected by immunoblotting (Figure 5.8).

Unfortunately, despite successful co-expression, incubation of a soluble crude protein extract from His-Res2p/His-Rep2p co-expressing cells with MCB1 did not produce a detectable band-shift as detected by EMSA (data not shown). In addition, mixing of crude soluble extracts containing His-Res1p and His-Res2p was tested. Furthermore, reconstitution of the DSC1 band-shift complex following mixing of crude His-Res1p and His-Res2p bacterial supernatants with fission yeast protein extracts prepared from their respective deletion mutants

(His-Rcs2p to Δ*res2* and His-Res1p to Δ*res1* cells) was attempted. However, using either of these approaches, no MCB-specific DNA-binding activity was detected (data not shown).

Neither bacterially expressed His-Res1p nor His-Res2p appeared able to bind MCB-DNA *in vitro*. Despite successful expression of soluble His-Rep2p, the lack of a definitive assay has precluded the assignment of activity to this protein. Due to the insoluble nature of both His-Cdc10p and His-Rep1p, these proteins have been refractory to similar analysis.

In contrast, the bacterial expression of a biologically active Res1p protein has been reported (Ayte et al., 1995). A GST-Res1p fusion protein was shown to bind specifically to MCB elements *in vitro*. The DNA binding ability of this recombinant protein was attributed to SpRes1p homodimerisation, artificially promoted by the GST moiety. Accordingly, it was reasoned that the inability of either His-Res1p or His-Res2p to bind MCB-DNA might reflect the self-dimerisation potential of these proteins *in vitro*. It is possible that the soluble His-Res1p and His-Res2p proteins are correctly folded but cannot homodimerise, thus precluding DNA-binding. An investigation was undertaken to assess the ability of Res1p and Res2p to bind specifically to MCB DNA, when bacterially expressed as GST fusion proteins. The production of both SpRes1p and SpRes2p as GST-fusion proteins will also allow assay of His-Rep2p via binding assays.

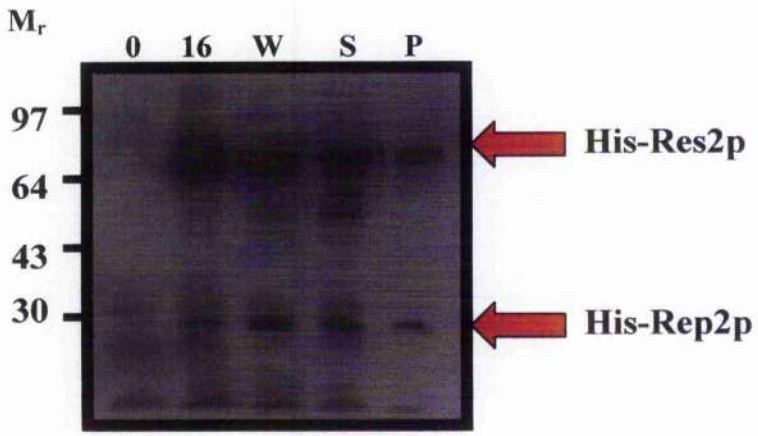


Figure 5.8 Co-expression of His-Res2p and His-Rep2p in *E. coli*

His-Res2p and HisRep2p were co-expressed at 15°C in BL21 CodonPlus (DE3)-RIL cells by induction with 0.2 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at 16 h. The bacterial culture was then centrifuged, the pellet resuspended in 5 ml 50mM KH_2PO_4, pH 7.5 and cells lysed by French pressure disruption. The whole cell extract (W) was separated into soluble (S) and insoluble (P) fractions by centrifugation and samples (1 ml) taken. Samples from 0 h and 16 h were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl/0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). All other samples were diluted with an equal volume of Laemmli sample buffer and similarly denatured. Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. Immunoblotting was carried out as described (Methods 2.4.7-2.4.11) and blots were probed with Anti-PentaHis-HRP conjugate antibody (Qiagen) at 1:2500. The positions of molecular weight markers are indicated (M_r) in kDa.

5.7 Cloning *res1*$^+$ and *res2*$^+$ cDNAs into pGEX-KG

The *res1*$^+$ and *res2*$^+$ cDNAs (Appendix I; GB 201 and GB 178, respectively) were amplified by PCR using primer pairs GO 464/GO 466 (*res1*$^+$) and GO 465/GO 466 (*res2*$^+$). Primers GO 464 and GO 465 were designed to add a 5' *Sma*I restriction site immediately upstream of the ATG of *res1*$^+$ and *res2*$^+$, respectively. Primer GO 466 adds a 3' *Xho*I restriction site downstream of the stop codon of each cDNA. Addition of these restriction ends was necessary to facilitate directional cloning into the pGEX-KG vector, in-frame with an N-terminal GST-tag.

Figure 5.9 provides a schematic representation of the PCR reaction and the primers used. All PCR reactions were carried out using the $Vent_R$ DNA polymerase (New England Biolabs).

5.7.1 Cloning *res1*$^+$ and *res2*$^+$ cDNAs into pCR2.1®

Following PCR amplification, each reaction mix was incubated with *Taq* polymerase at 72°C for 10 min to allow cloning into the linearised pCR2.1® vector. The cloning of PCR products using the TA Cloning® Kit (Invitrogen) has been described previously (Chapter 3, Section 3.3.2-3.3.3). Use of the TA Cloning® Kit, as an intermediate cloning step, was necessary due to the difficulty in digesting the *Sma*I and *Xho*I restriction ends present at the end of a PCR product (restriction enzymes often require substantial flanking sequence around their recognition site to digest efficiently). Following cloning into the TA vector, the cDNAs can then be easily excised for cloning into pGEX-KG.

Plasmid DNA was purified from overnight cultures (Methods 2.2.5) and analysed by restriction enzyme mapping to confirm the presence of an insert of the correct size (Figures 5.10A and 5.11A). Potential positive clones were sequenced (MWG-Biotech – using oligos listed in Appendix II) to confirm the presence of insert and the fidelity of the PCR. A positive clone for both *res1*$^+$ and *res2*$^+$ was stored at -70°C in the lab bacterial collection (Appendix I - GB 300 and GB 302).

5.7.2 Cloning *res1*$^+$ and *res2*$^+$ cDNAs into pGEX-KG

The plasmids GB 300 (*res1*$^+$) and GB 302 (*res2*$^+$) were digested with *Sma*I and *Xho*I as was the pGEX-KG vector to generate cohesive ends for ligation. The vector was also treated with calf intestinal alkaline phosphatase (Methods 2.2.7 and 2.2.8). The products of the restriction digests were analysed on a 1.5 % (w/v) agarose gel (Figures 5.10A and 5.11A) and bands of the appropriate size excised and purified using the QIAquick® Gel Extraction Kit (Methods 2.2.4 and 2.2.6). 5 µl samples were analysed on a 1.5 % (w/v) agarose gel to determine the vector:insert ratio to be used in the subsequent ligation reaction. Ligation reactions were carried out as described (Methods 2.2.9) and transformed into *E. coli* DH5α cells the following day (Methods 2.2.13). Plasmid DNA was purified from overnight cultures (Methods 2.2.5) and analysed by restriction enzyme mapping (Figures 5.10B and 5.11B) and DNA sequencing (MWG-Biotech – using oligos listed in Appendix II) to confirm cloning was in-frame with the N-terminal GST-tag. A positive clone for both *GST-res1*$^+$ and *GST-res2*$^+$ was stored at -70°C in the lab bacterial collection (Appendix I - GB 309 and GB 318, respectively).

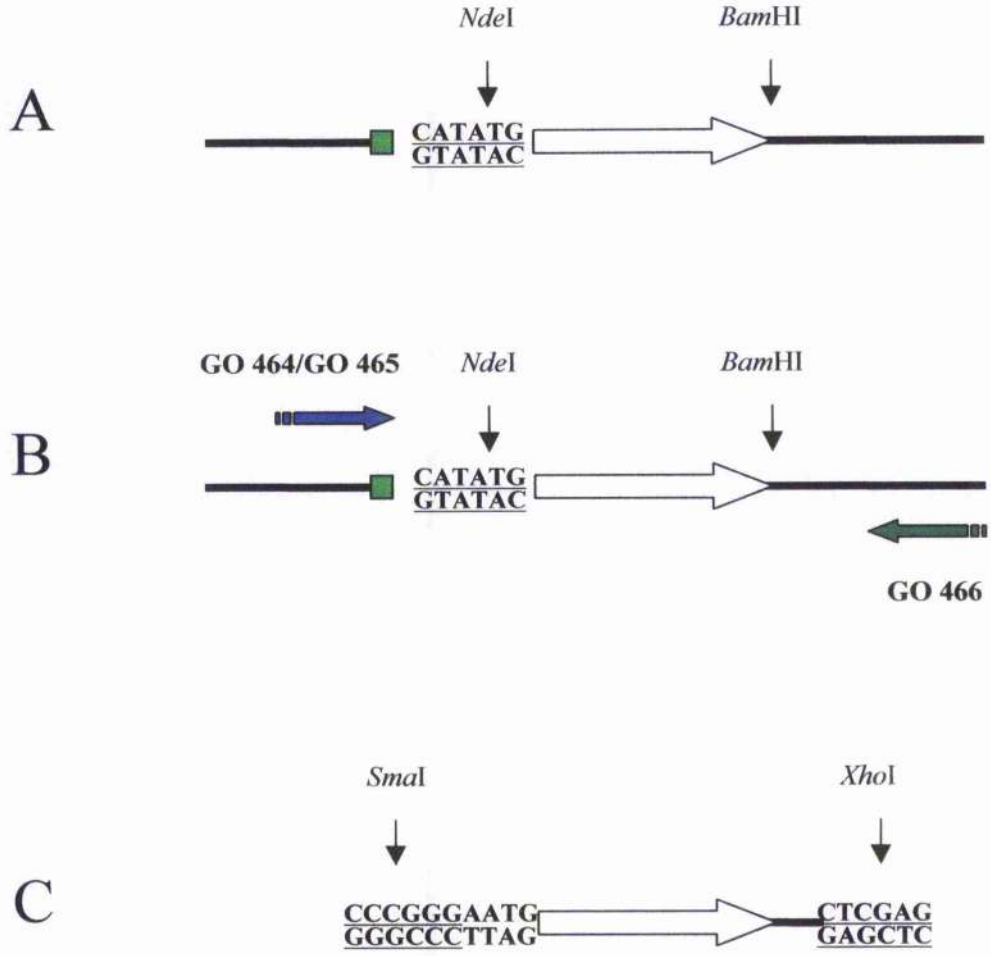


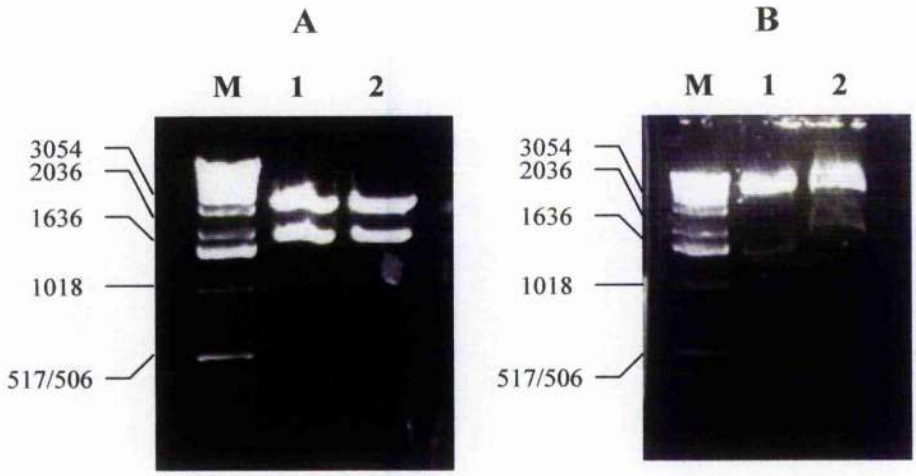
Figure 5.9 Schematic representation of the PCR amplifications for pGEX-KG cloning

The white arrow represents the ORFs of the *res1*$^{+}$ and *res2*$^{+}$ cDNAs cloned into pET 28c at *Nde*I-*Bam*HI. The His-tag is boxed in green, the black lines represent pET vector flanking sequence and the position of restriction sites are indicated by arrows.

A Restriction map representing the *Nde*I-*Bam*HI fragment of each cDNA cloned into the pET-28c vector (Chapter 3).

B The positions of the primers used in the PCR reaction are indicated in relation to the corresponding sequence of the pET-28c vector. GO 464 and GO 465 add a *Sma*I site (**CCCGGG**) immediately upstream of the **ATG** of *res1*$^{+}$ and *res2*$^{+}$, respectively. GO 466 includes the pET-28c *Xho*I (**CTCGAG**) site downstream of the stop codon.

C A schematic representation of the DNA fragments produced following PCR.



5.10 Restriction digestion analyses of the *res1*$^+$-pCR2.1® and *res1*$^+$-pGEX-KG plasmids

Clones containing insert were analysed by restriction mapping. Samples of each digest (5 µl) were analysed on a 1.5% (w/v) agarose gel stained with ethidium bromide and viewed under UV illumination. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs.

A Digestion of the *res1*$^+$-pCR2.1® plasmid: **1** *Eco*RI **2** *Sma*I-*Xho*I.

B Digestion of the *res1*$^+$-pGEX-KG plasmid: **1** *Hind*III **2** *Sac*I.

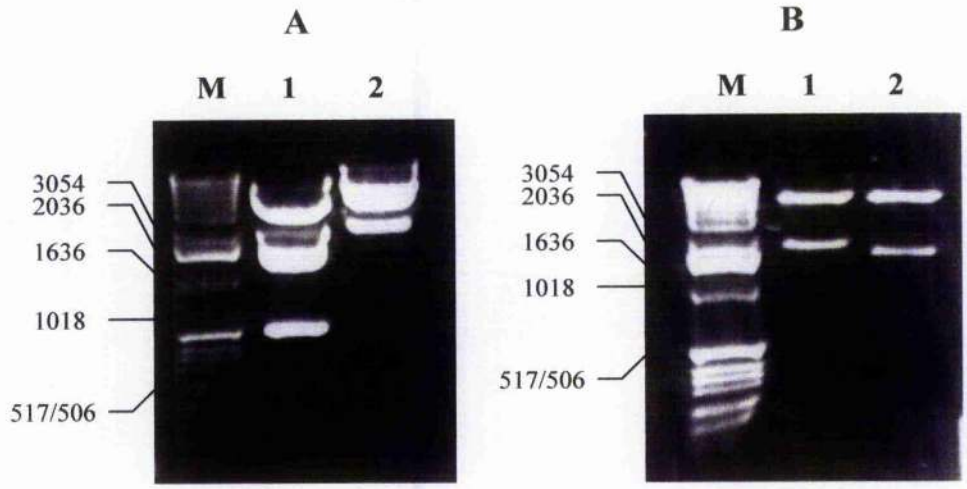
The restriction enzymes used in each digest and the predicted fragment sizes are summarised for each clone in Tables 5.1 and 5.2.

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Eco*RI**	**2**	**3900, 1900**
***Sma*I-*Xho*I**	**2**	**3900, 1900**

Table 5.1 *res1*$^+$-pCR2.1®

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Hind*III**	**2**	**5600, 1300**
***Sac*I**	**2**	**5200, 1700**

Table 5.2 *res1*$^+$-pGEX-KG



5.11 Restriction digestion analyses of the *res2*$^{+}$-pCR2.1® and *res2*$^{+}$-pGEX-KG plasmids

Clones containing insert were analysed by restriction mapping. Samples of each digest (5 µl) were analysed on a 1.5% (w/v) agarose gel stained with ethidium bromide and viewed under UV illumination. The 1 kb DNA ladder is shown (M) and sizes are indicated in base-pairs.

A Digestion of the *res2*$^{+}$-pCR2.1® plasmid: **1** *Eco*RI **2** *Sma*I-*Xho*I.

B Digestion of the *res2*$^{+}$-pGEX-KG plasmid: **1** *Sal*I **2** *Eco*RI.

The restriction enzymes used in each digest and the predicted fragment sizes are summarised for each clone in Tables 5.3 and 5.4. Fragments marked (*) appear as a result of partial digest.

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Eco*RI**	**4**	**3900,2000*, 1500,500**
***Sma*I-*Xho*I**	**2**	**3900, 2000**

Table 5.3 *res2*$^{+}$-pCR2.1®

Enzyme(s)	Number of Fragments	Fragment sizes (bp)
***Sal*I**	**2**	**5200, 1700**
***Eco*RI**	**2**	**5400, 1500**

Table 5.4 *res2*$^{+}$-pGEX-KG

5.8 Overexpression and solubility of GST-Res1p and GST-Res2p

As described for the His-tagged recombinant proteins (Chapter 3 and this Chapter), in order to obtain optimal conditions for the overexpression and solubility of both GST-Res1p and GST-Res2p, several *E. coli* expression strains were tested, together with a variety of growth and lysis buffer conditions. As before, the *E. coli* expression strain BL21 (DE3) CodonPlus-RIL gave optimal results for both recombinant proteins, following transformation with the clones containing the *GST-res1*$^+$ and *GST-res2*$^+$ cDNAs. Small-scale protein inductions were carried out to test for successful overexpression of each protein (Methods 2.4.1 and 2.4.12). Aliquots of the induction at 0 h and 3 h or 16 h together with samples of the whole cell extract (W), supernatant (S) and pellet (P) were then analysed by SDS-PAGE and the solubility of each protein determined by viewing after Coomassie blue staining or immunoblotting. Figures 5.12 and 5.13 show the overexpression and solubility of GST-Res1p and GST-Res2p at different temperatures as viewed by Coomassie blue staining. The GST-Res1p and GST-Res2p proteins resolve with a M_r approximately 98,000 (GST is a protein with a M_r value of 26,000). Figure 5.14 shows immunoblot analysis of the overexpression and solubility of GST-Res1p and GST-Res2p. For comparison, overexpression of the GST moiety alone is similarly depicted in Figure 5.15.

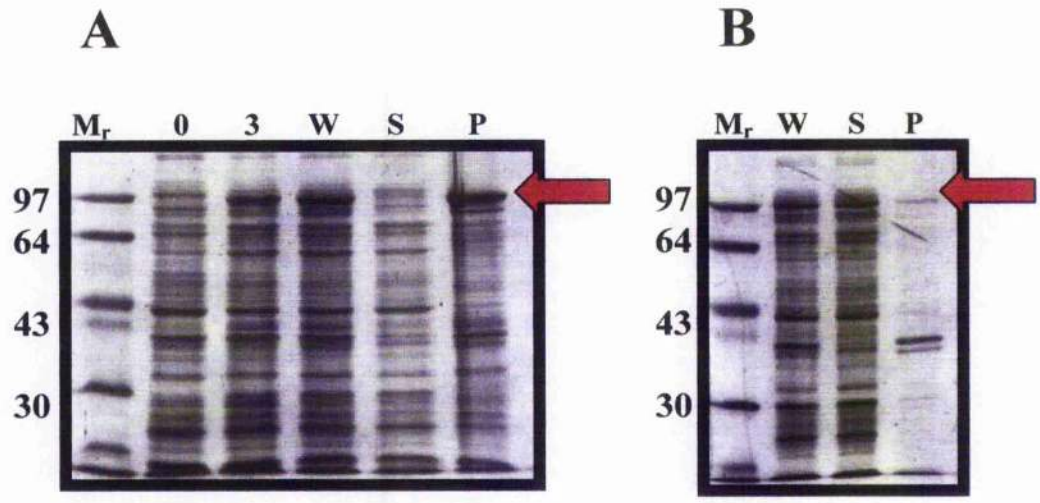


Figure 5.12 Overexpression of GST-Res1p

GST-Res1p was overexpressed at 30°C (**A**) or 15°C (**B**) in BL21 CodonPlus (DE3)-RIL cells by induction with 1 mM or 0.2 mM IPTG respectively. Samples (1 ml) were removed at the point of induction (0 h) and at 3 h or 16 h. The bacterial culture was then centrifuged and the pellet resuspended in 5 ml 1 x PBS. Cells were lysed by French pressure cell disruption. This whole cell extract (W) was then separated into soluble (S) and insoluble (P) fractions by centrifugation and samples (1 ml) taken. Samples from 0 h and 3 h were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl/0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). All other samples were diluted with an equal volume of Laemmli sample buffer and similarly denatured. Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. Molecular weight markers are shown (M_r) with sizes indicated in kDa.

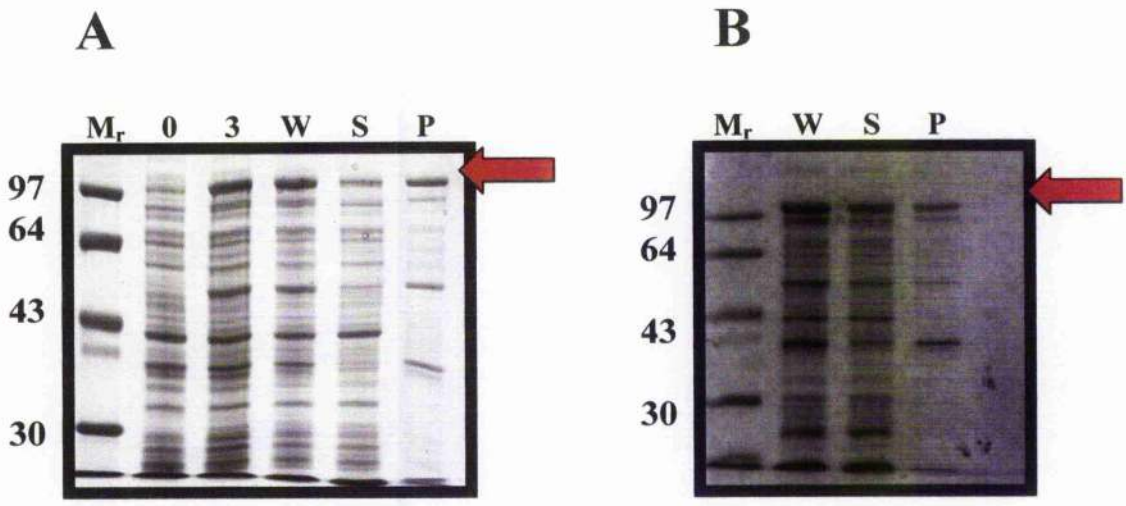


Figure 5.13 Overexpression of GST-Res2p

GST-Res2p was overexpressed at 30°C (**A**) or 15°C (**B**) in BL21 CodonPlus (DE3)-RIL cells by induction with 1 mM or 0.2 mM IPTG respectively. Samples (1 ml) were removed at the point of induction (0 h) and at 3 h or 16 h. The bacterial culture was then centrifuged and the pellet resuspended in 5 ml 1 x PBS. Cells were lysed by French pressure cell disruption. This whole cell extract (W) was then separated into soluble (S) and insoluble (P) fractions by centrifugation and samples (1 ml) taken. Samples from 0 h and 3 h were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl/0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). All other samples were diluted with an equal volume of Laemmli sample buffer and similarly denatured. Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. Molecular weight markers are shown (M_r) with sizes indicated in kDa.

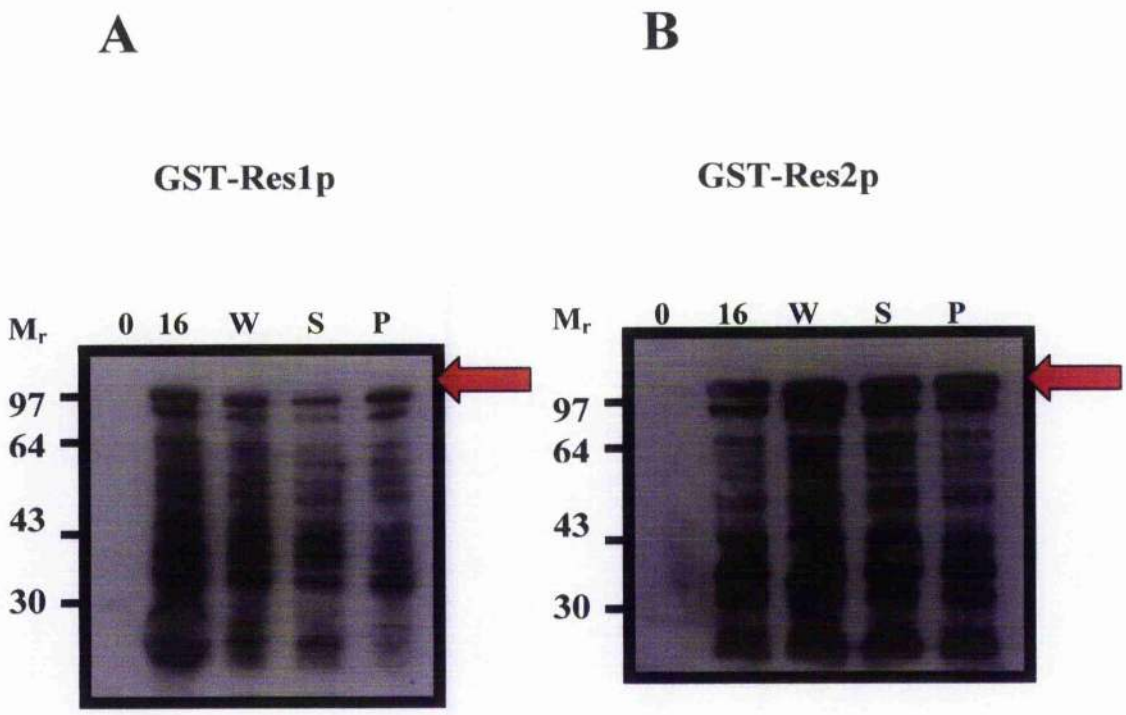


Figure 5.14 Immunoblot analysis of GST-Res1p and GST-Res2p

GST-Res1p (**A**) and GST-Res2p (**B**) were overexpressed at 15°C in BL21 CodonPlus (DE3)-RIL cells by induction with 0.2 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at 16 h. Each bacterial culture was then centrifuged and the pellets resuspended in 5 ml 1 x PBS. Cells were lysed by French pressure cell disruption. The whole cell extracts (W) were then separated into soluble (S) and insoluble (P) fractions by centrifugation and samples (1 ml) taken. Samples from 0 h and 16 h were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl/0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). All other samples were diluted with an equal volume of Laemmli sample buffer and similarly denatured. Samples were then analysed on a 10 % SDS-polyacrylamide gel stained with Coomassie brilliant blue. Immunoblotting was carried out as described (Methods 2.4.7-2.4.11) and blots were probed with Anti-GST-HRP conjugate antibody (Amersham) at 1:2500. The positions of molecular weight markers are indicated (M_r) in kDa.

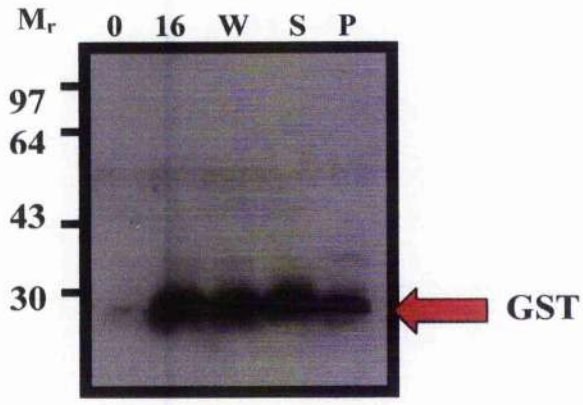


Figure 5.15 Immunoblot analysis of GST

GST was overexpressed at 15°C in BL21 CodonPlus (DE3)-RIL cells by induction with 0.2 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at 16 h. The bacterial culture was centrifuged and the pellet resuspended in 5 ml 1 x PBS. Cells were lysed by French pressure cell disruption. The whole cell extract (W) was separated into soluble (S) and insoluble (P) fractions by centrifugation and samples (1 ml) taken. Samples from 0 h and 16 h were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl/0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). All other samples were diluted with an equal volume of Laemmli sample buffer and similarly denatured. Samples were then analysed by immunoblotting as described (Methods 2.4.7-2.4.11). Blots were probed with Anti-GST-HRP conjugate antibody (Amersham) at 1:2500. The positions of molecular weight markers are indicated (M_r) in kDa.

5.9 Biological activity assays - GST-Res1p and GST-Res2p

Bacterially expressed GST-Res1p bound specifically to MCB elements *in vitro* (Ayte et al., 1995). Accordingly, the EMSA technique was employed to assay the biological activity of the recombinant GST-Res1p and GST-Res2p proteins.

Recombinant GST-Res1p and GST-Res2p were individually overexpressed at 15°C (Section 5.8) and recovered in the supernatant fraction. Soluble protein extracts (Methods 2.4.14) were incubated with a radio-labelled MCB1 DNA probe and used in EMSA experiments as described (Methods 2.2.14 - 2.2.16). The results of this assay are shown in Figure 5.16. Incubation of either GST-Res1p or GST-Res2p with MCB1 produced a co-migratory band-shift activity. In contrast, this band-shift activity was absent from extracts of GST alone incubated with MCB1. Both the GST-Res1p and GST-Res2p band-shift activities were specific to the MCB1 element, as judged by competition analysis and an inability to bind the mutated MCB1m sequence. Notably, both of these band-shifts were of a greater mobility than the DSC1 band-shift formed in wild-type fission yeast extracts, consistent with the absence of SpCdc10p, SpRep2p and SpRes2p or SpRes1p, respectively. In addition, the presence of the GST-tag, in both GST-Res1p and GST-Res2p containing complexes, was confirmed by a super-shift following addition of an anti-GST antibody (Figure 5.17).

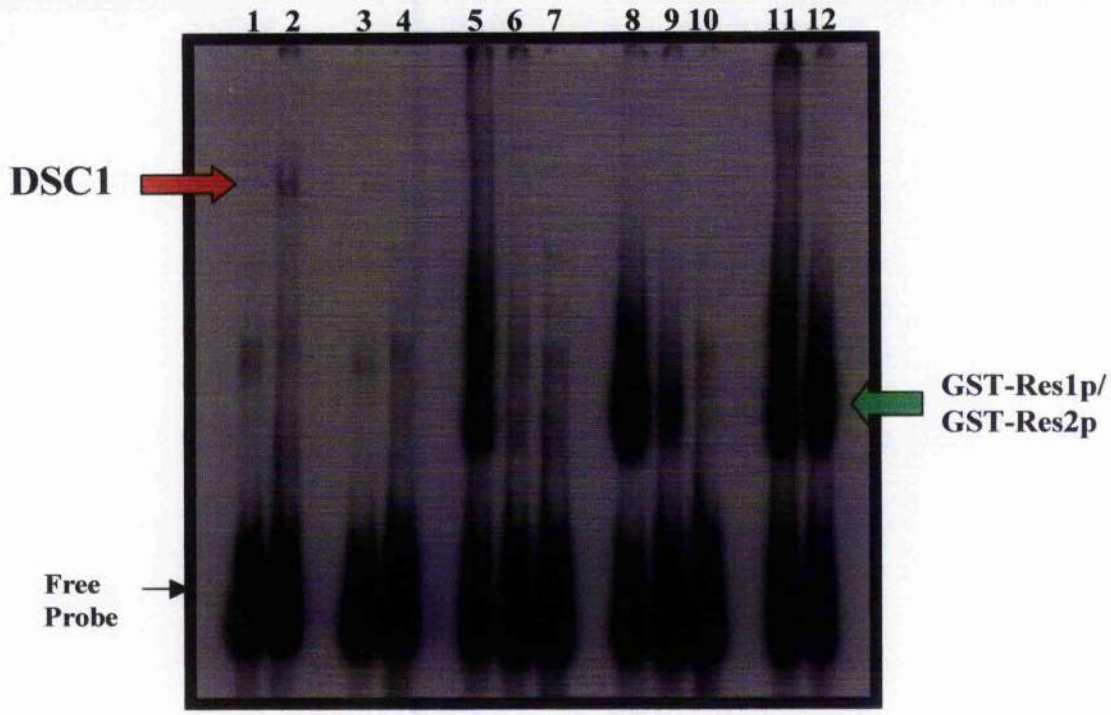


Figure 5.16 GST-Res1p and GST-Res2p bind specifically to MCB1 DNA *in vitro*

EMSA analysis of protein extracts prepared from *E. coli* BL21 (DE3) CodonPlus-RIL cells transformed with pGEX-KG (GB 159), *GST-res1*$^+$-pGEX-KG (GB 309) or *GST-res2*$^+$-pGEX-KG (GB 318). Protein extracts were prepared and incubated with a radio-labelled probe derived from the *cdc22*$^+$ promoter (MCB1). To check the specificity of the retarded complex an excess of unlabelled MCB1 probe was added (lanes 6 and 9) or a probe containing a mutated MCB1 (MCB1m) element (lanes 4, 7 and 10) or an excess of unlabelled MCB1m (lanes 11 and 12). A protein extract from wild-type fission yeast cells (GG 217) was incubated with MCB1 and forms a characteristic DSC1 band-shift for comparison (lane 2). The upper and lower arrows on the left indicate the positions of the DSC1 complex and the unbound radio-labelled probe, respectively. The arrow on the right indicates the GST-Res1p and GST-Res2p band-shifts.

1 = free MCB1 probe

2 = wild-type protein extract + MCB1 probe

3 = pGEX-KG protein extract + MCB1 probe

4 = pGEX-KG protein extract + MCB1m probe

5 = *GST-res1*$^+$-pGEX-KG protein extract + MCB1 probe

6 = *GST-res1*$^+$-pGEX-KG protein extract + excess unlabelled MCB1 probe

7 = *GST-res1*$^+$-pGEX-KG protein extract + MCB1m probe

8 = *GST-res2*$^+$-pGEX-KG protein extract + MCB1 probe

9 = *GST-res2*$^+$-pGEX-KG protein extract + excess unlabelled MCB1 probe

10 = *GST-res2*$^+$-pGEX-KG protein extract + MCB1m probe

11 = *GST-res1*$^+$-pGEX-KG protein extract + excess unlabelled MCB1m probe

12 = *GST-res2*$^+$-pGEX-KG protein extract + excess unlabelled MCB1m probe

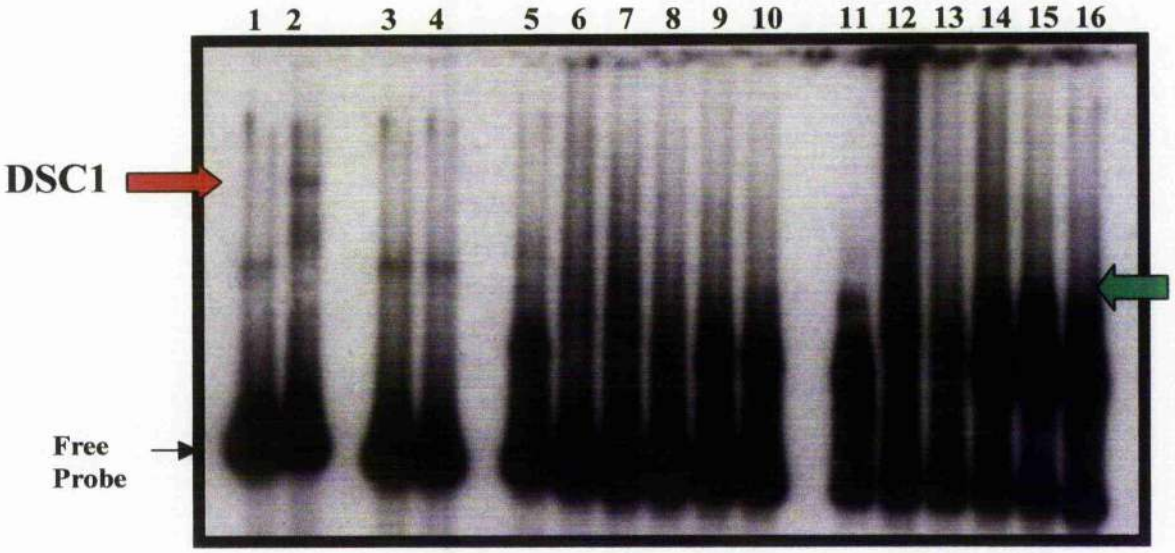


Figure 5.17 Super-shift EMSA assay of GST-Res1p and GST-Res2p with Anti-GST antibody

EMSA analysis of protein extracts prepared from *E. coli* BL21 (DE3) CodonPlus-RIL cells transformed with pGEX-KG (GB 159), p*GST-res1*$^+$ (GB 309) or p*GST-res2*$^+$ (GB 318). Protein extracts were prepared and incubated with a radio-labelled probe derived from the *cdc22*$^+$ promoter (MCB1). Protein extracts mixed with probe were incubated with various dilutions of an anti-GST antibody prior to loading on a 2% SeaPlaque: 1% agarose gel. A protein extract from wild-type fission yeast cells (GG 217) was incubated with MCB1 and forms a characteristic DSC1 band-shift for comparison (lane 2). The upper and lower arrows on the left indicate the positions of the DSC1 complex and the unbound radio-labelled probe, respectively. The green arrow indicates the super-shifted GST-Res2p complex.

1 = free MCB1 probe
2 = wild-type protein extract + MCB1 probe
3 = pGEX-KG protein extract + MCB1 probe
4 = pGEX-KG protein extract + MCB1 probe + Anti-GST (1:1)
5 = *GST-res1*$^+$-pGEX-KG protein extract + MCB1 probe
6 = *GST-res1*$^+$-pGEX-KG protein extract + MCB1 probe + Anti-GST (1:1)
7 = *GST-res1*$^+$-pGEX-KG protein extract + MCB1 probe + Anti-GST (1:2)
8 = *GST-res1*$^+$-pGEX-KG protein extract + MCB1 probe + Anti-GST (1:5)
9 = *GST-res1*$^+$-pGEX-KG protein extract + MCB1 probe + Anti-GST (1:10)
10 = *GST-res1*$^+$-pGEX-KG protein extract + MCB1 probe + Anti-GST (1:50)
11 = *GST-res2*$^+$-pGEX-KG protein extract + MCB1 probe
12 = *GST-res2*$^+$-pGEX-KG protein extract + MCB1 probe + Anti-GST (1:1)
13 = *GST-res2*$^+$-pGEX-KG protein extract + MCB1 probe + Anti-GST (1:2)
14 = *GST-res2*$^+$-pGEX-KG protein extract + MCB1 probe + Anti-GST (1:5)
15 = *GST-res2*$^+$-pGEX-KG protein extract + MCB1 probe + Anti-GST (1:10)
16 = *GST-res2*$^+$-pGEX-KG protein extract + MCB1 probe + Anti-GST (1:50)

5.10 Discussion

In this chapter, attempts to obtain soluble recombinant protein, following overexpression of His-Res1p, His-Res2p, His-Rep1p, His-Rep2p or His-Cdc10p in *E. coli*, have been described. Following this, a series of assays have been described to determine the biological activity of these soluble recombinant proteins when produced *in vitro*.

The ability to obtain a soluble protein, following release from the cell, can be dependent on several factors. With respect to heterologous protein production, not only must the conditions within the cell be suitable, facilitating efficient translation and protein folding, but also the protein must then be recovered and maintained in a solvent in which it is stable. Accordingly, a series of growth and lysis buffer conditions were screened, in attempts to recover each of the recombinant proteins in a soluble form.

The major advantage of heterologous expression systems is for the production of substantial quantities of recombinant protein. This is often counterbalanced by the affect that such high levels can impose on solubility. Indeed, recombinant protein may comprise ≥50% of total cell protein, and this can result in the formation of insoluble aggregates. Expression at lower temperatures, in combination with a reduced concentration of inducing agent, can often aid in the production of soluble protein (Schien and Noteborn, 1988). It has been suggested that lowering the induction temperature can slow down the rate of translation. In turn, the amount of protein that accumulates in a given time is reduced and as a result, the protein has more time to fold correctly, helping to prevent formation of insoluble aggregates.

Initial attempts to obtain soluble protein following overexpression of the His-tagged proteins at 37°C were unsuccessful, with the exception of His-Res1p. Given that the expression of His-Res1p was at a relatively low level (detectable only by immunoblotting), the partial solubility of this recombinant protein at this, and indeed all temperatures tested may reflect this observation. Similarly, His-Rep2p (also detectable only by immunoblot analysis) although completely insoluble at 37°C, appeared in the supernatant following expression at 30°C, 22°C or 15°C. Conversely, both His-Rep1p and His-Cdc10p were present in the insoluble fraction under all conditions tested, perhaps reflecting higher expression levels. These proteins may therefore be present as inclusion bodies, which can form for several reasons. The recombinant protein may require post-translational modification and/or molecular chaperones in order to fold correctly. In addition, the high levels of protein expressed within the cell may account for dense packing and aggregation (Lilie et al., 1998).

Even so, inclusion bodies can often be a useful source of pure protein. The solubilisation and purification of both His-Res1p and His-Cdc10p from inclusion bodies is discussed in detail in Chapter 6.

In contrast the solubility of His-Res2p increased significantly following induction at 15°C for 16 h. The ability to produce soluble His-Res2p was not solely dependent on finding the appropriate induction conditions, but was equally reliant on the composition of the lysis buffer. Using the variant solubility of His-Res2p in buffers of different salt concentration as an example, one of the many parameters that can dramatically affect the solubility of a protein has been highlighted. Furthermore, the apparent insolubility of His-Res2p in NaCl concentrations $\geq$ 300 mM has important consequences for the purification of this protein (discussed in Chapter 6).

Following production of soluble His-Res1p (albeit at low levels) and His-Res2p, these recombinant proteins were assayed for biological activity. Crude soluble extracts from *E. coli*, expressing either His-Res1p or His-Res2p, were incubated with a radio-labelled MCB probe. However, an *in vitro* MCB DNA-binding activity was not detected with either of these proteins using this technique.

Co-expression of His-Cdc10p with either His-Res1p or His-Res2p was attempted. Unfortunately, co-expression of either protein with His-Cdc10p proved unsuccessful, in terms of detectable protein (either His-Res1p or His-Res2p were detected but not His-Cdc10p by Coomassie blue staining or immunoblotting). In addition, co-expression of either His-Res1p or His-Res2p with each other and with His-Rep1p or His-Rep2p was tested. Association of these proteins might facilitate MCB-specific DNA binding of the His-Resp subunits in the absence of His-Cdc10p. However, although co-expression of His-Res2p/His-Rep2p was detected, and both of these proteins were present in the soluble fraction (as determined by immunoblotting – Figure 5.8), use of this protein extract did not yield detectable MCB binding *in vitro*.

These observations suggest that neither bacterially produced soluble His-Res1p or His-Res2p has DNA-binding activity. There are a number of possible explanations why the His-tagged Resp proteins cannot bind to DNA *in vitro*. The His-tag, present at the N-terminus of these proteins, may prevent proper folding and therefore occlude DNA-binding. To counter this possibility the presence of a His-tag at either the N- or C-terminus had no adverse effect on the specific DNA-binding properties of bacterially expressed ScSwi4p or ScMbp1p (Taylor et al., 2000).

With respect to the results observed with His-Res1p, it has been reported that *in vitro* translated SpRes1p cannot bind to MCB DNA, in the absence of SpCdc10p (Ayte et al., 1995). Similarly, the MCB-specific DNA binding activity of *in vitro* translated SpRes2p was dependent on co-expression with SpCdc10p, although a weak band-shift activity was detected with SpRes2p alone (Zhu et al., 1997).

The inability to co-express either His-Res1p or His-Res2p with His-Cdc10p has precluded a definitive analysis of the dependency of these proteins upon His-Cdc10p, for their DNA-binding activity *in vitro*. Using His-Res2p alone, no MCB specific band-shift activity

was detected. Pertinent to this result, Zhu and co-workers previously reported an inability to detect a MCB-specific DNA-binding complex by EMSA with *in vitro* translated SpRes2p, using three different MCB containing probes (Zhu et al., 1994). Therefore, the ability of SpRes2p alone to bind to MCB DNA *in vitro* may not be straightforward.

Intriguingly, it has been shown that C-terminal truncation of *in vitro* produced SpRes2p enhances MCB-specific DNA-binding. The C-termini of SpRes1p and SpRes2p are responsible for binding to SpCdc10p (Ayte et al., 1995; Zhu et al., 1997). The fact that only a weak band-shift was detected with SpRes2p alone, and that this was markedly increased by the presence of SpCdc10p, suggested that in the absence of SpCdc10p, the ability of SpRes2p to bind DNA is compromised (Zhu et al., 1997).

In agreement with this observation, analysis of the ScSwi4p protein *in vitro* suggests that, in the absence of ScSwi6p, the C-terminus of ScSwi4p is involved in contact with the N-terminal DNA-binding domain. This has led to the proposal that control of DNA-binding in ScSwi4p may be mediated, at least in part, through intramolecular auto-inhibition, in the absence of ScSwi6p (Baetz and Andrews, 1999).

Given the extensive sequence homology between the SpRes1p/SpRes2p and ScSwi4p proteins at the N and C-termini, it is tempting to suggest that such a mode of DNA-binding regulation may be conserved. It is perhaps relevant that the co-expression experiments of Zhu and co-workers utilised *in vitro* translated protein made in rabbit reticulocyte lysates and so it is possible that the DNA-binding activity exhibited by these proteins reflects a requirement for post-translational modification that only occurs in eukaryotes.

Arguing against this, bacterially expressed SpRes1p has been shown to bind specifically to MCB DNA by EMSA. This band-shift activity was attributed to the artificial homo-dimerisation of SpRes1p imparted by the GST moiety (Ayte et al., 1995). Further evidence for this was demonstrated by the MCB-specific binding of SpRes1p when fused at its N-terminus to the Epstein Barr Virus ZEBRA domain, a protein with intrinsic dimerisation potential (Ayte et al., 1997).

Such results have led these authors to attribute the apparent inability of SpRes1p to bind DNA *in vitro*, to homodimerisation potential. SpRes2p can homodimerise *in vitro*, in a SpCdc10p-dependent manner (Zhu et al. 1997). Both SpRes1p and SpRes2p are believed to form heterodimers in the mitotic DSC1 complex *in vivo* (Ayte et al., 1997; Whitehall et al., 1999).

This suggestion prompted an investigation to determine if GST-Res1p and GST-Res2p could bind specifically to MCB DNA, in contrast to the His-tagged isoforms. Following cloning into a GST expression vector and successful overexpression of soluble protein, the ability of GST-Res1p and GST-Res2p to bind specifically *in vitro* to MCB DNA was assayed. In accordance with the results of Ayte and co-workers, GST-Res1p produced a MCB-specific

band-shift complex. Notably, expression of SpRes2p as a GST-fusion protein also conferred MCB-specific DNA-binding. Both DNA-binding activities were co-migratory (a slightly stronger complex was detected with GST-Res2p – most likely reflecting the higher level of expression of this protein). The presence of the GST-tag in each band-shift complex was confirmed following super-shift analysis, using an anti-GST antibody (Figure 5.17).

The fact that *E. coli* is capable of producing biologically active GST-Res1p and GST-Res2p (in terms of MCB-specific DNA binding), indicates that post-translational modification is not essential for the DNA-binding activity of these proteins *in vitro*. Furthermore, although the band-shift activities are MCB-specific, the artificial dimerisation imparted by the GST moiety may confer this DNA binding capability. Indeed, if this were found to be true, then the inability of His-Res1p or His-Res2p to bind DNA might reflect the monomeric nature of these proteins *in vitro*. If either homo- or heterodimerisation *in vitro* is dependent on SpCdc10p (outwith artificial means), then the His-Resp proteins may not be inactive *per se*. The soluble nature of these proteins suggests that they may be correctly folded and so analysis of these proteins may still be worthwhile.

It should be noted that in all assays using GST-Res1p and GST-Res2p, crude soluble protein extracts were used. As is apparent from both Coomassie blue staining and immunoblotting, these extracts contain a heterogeneous population of both GST-Res1p and GST-Res2p proteins. In both cases the band-shift was reliant on the presence of either GST-Res1p or GST-Res2p (no specific complex was detected using extracts from *E. coli* expressing the GST-tag alone). Nevertheless, the possibility that MCB-specific DNA-binding is conferred by C-terminally truncated, rather than full-length, GST-Res1p or GST-Res2p molecules cannot be excluded. Such an explanation would correlate with relief of C-terminal mediated auto-inhibiton. The smeared appearance of the band-shift complex in both cases indicates such heterogeneity and has been noted by Ayte and co-workers (Ayte et al., 1995).

Despite the apparent MCB-specific DNA binding capacity of GST-Res1p and GST-Res2p, a more definitive answer regarding the DNA-binding properties of these proteins requires further analysis. A more direct analysis toward answering these questions necessitates purification of these proteins. Firstly, with respect to GST-Res1p and GST-Res2p, purification of full-length protein will allow determination of the nature of the DNA-binding activity. Is the MCB-specific DNA-binding activity generated by full-length GST-Res1p/GST-Res2p, or instead C-terminally truncated isoforms? In addition, thrombin cleavage of the GST-tag and purification of the resultant Res1p and Res2p proteins will allow separation of the effects of the GST-tag on DNA-binding. The presence of the GST-tag may stabilise these proteins and allow them to fold correctly. In contrast, the artificial homo-dimerisation imparted by the GST moiety may be required for DNA-binding.

Similarly, there are un-resolved issues over the results of the *in vitro* assays with both His-tagged Res1p and Res2p. Thrombin cleavage of the His-tag from purified His-Res1p and His-Res2p will allow separation of the effects of the His-tag, if any, on DNA-binding. Purification of these proteins and recovery in sufficient amounts will provide valuable material for use in biophysical studies allowing an analysis of the structure of these proteins. Of particular interest will be a comparison of the His-tagged and GST-tagged isoforms. In addition, purification of the GST-Resp proteins described in this chapter will allow an analysis of the activity of His-Rep2p using *in vitro* binding assays.

In the following chapter the purification of these recombinant proteins is described, to permit detailed functional, and ultimately structural, analyses.

Chapter 6

Purification of recombinant DSC1 proteins

6.1 Introduction

Chapter 5 focused on analysis of the biological activity of recombinant DSC1 components, when produced *in vitro*. These experiments showed that, when bacterially expressed as GST-fusion proteins, both Res1p and Res2p bound specifically to MCB containing DNA in EMSA assays. In contrast, their respective His-tagged isoforms displayed no such detectable DNA-binding activity. Similar assays using His-Rep2p were inconclusive and both His-Rep1p and His-Cdc10p were insoluble following overexpression. This precluded *in vitro* assays using either of these latter two proteins.

As discussed in the previous chapter, analysis of these results has raised several issues regarding the behaviour of the recombinant GST-Resp and His-Resp proteins *in vitro*. Specifically, with respect to both GST-Res1p and GST-Res2p, the nature of the protein responsible for MCB-specific DNA-binding has still to be fully resolved. *In vitro* DNA-binding assays employed crude soluble bacterial extracts, containing a heterogeneous mix of both full-length and truncated GST-Resp proteins. The effect(s) of the GST moiety upon DNA-binding requires further investigation, particularly with regard to *in vitro* homodimerisation and so the His-tagged isoforms of Res1p and Res2p merit further analysis. The inability of the His-Resp proteins to bind MCB-DNA *in vitro* may reflect an inability to homodimerise. Removal of the His and GST moieties will allow separation of the effect of these tags, if any, upon the behaviour of these proteins *in vitro*.

Detailed analysis of a protein requires it is free of any contaminants that may interfere with or influence its activity. In order to obtain a greater understanding of the behaviour of these recombinant proteins *in vitro*, purification must be undertaken. Purification of either GST-Res1p or GST-Res2p will also allow *in vitro* binding assays with His-Rep2p. Importantly, recovery of protein in sufficient yield and purity will also provide valuable material for detailed structural studies.

The purification of a protein from the many macromolecular compounds within a cell can often be a formidable task, particularly if it is present in low abundance. In addition, the protein may exist in several different isoforms within the cell. Such micro-heterogeneity can be problematic when attempting to obtain a homogeneous preparation, required for detailed structural studies. As discussed in Chapter 3, biophysical analyses require large amounts of pure and active protein that are often extremely difficult to obtain from the native source. Overexpression of proteins of interest, as recombinants with fusion tags in bacteria, often simplifies purification and increases yield.

Expression of a protein with a fusion tag is primarily useful for purification purposes; in some cases expression, solubility and stability may also be enhanced. The highly specific binding properties of the fusion tag can be exploited in affinity chromatography, providing a

relatively straightforward means of purification (generally applicable to any protein containing the tag). Inclusion of a genetically engineered protease recognition sequence at the C-terminus of the tag allows cleavage either during or after purification. The bacterial overexpression and solubility of each of the fission yeast DSC1 components, with an N-terminal His-tag, has been described previously (Chapter 3, Section 3.4). Both Res1p and Res2p have also been bacterially expressed as N-terminal GST-fusions (Chapter 5, section 5.8). Both the pET-28c and pGEX-KG vectors contain thrombin cleavage sites immediately C-terminal to their His and GST-tags, respectively.

The glutathione *S*-transferase enzyme has a subunit M_r of 26 kDa, and is frequently used as an N-terminal fusion partner in recombinant protein studies. GST-fusion proteins can then be purified based on their affinity for glutathione (Smith and Johnson, 1988). Similarly, fusion of a 6-Histidine-tag to the N or C-terminus of a protein affords a simple means of purification by exploiting the affinity of histidine for transition metal ions (Porath et al., 1975; Smith et al., 1988). In both cases, the fusion proteins can then be purified, following binding to their respective ligands (immobilised on a chromatographic support), by mass-action competitive binding. Elution is achieved by addition of reduced glutathione for GST-fusions, whereas imidazole competes with His-tagged proteins.

Ideally, for purification using these methods, the protein should be in soluble form. Even so, recovery of pure and active protein from insoluble aggregates is not unprecedented. Both His-Rep1p and His-Cdc10p were insoluble under all conditions tested, which prevented assay of their activities *in vitro*. The presence of insoluble proteins in inclusion bodies, however, often provides a convenient and straightforward means of obtaining pure protein in high yield (Carrio and Villaverde, 2002). The material present within inclusion bodies frequently represents mis-folded aggregates of protein. Accordingly, recovery of active protein from inclusion bodies routinely involves re-folding, which can often be an extremely complex process. Even so, several proteins purified from inclusion bodies have been found to obtain significant secondary and tertiary structure, with some displaying significant activity (Carrio and Villaverde, 2002). Purification of His-Rep1p and His-Cdc10p from inclusion bodies has the potential to provide a reproducible source of active protein in high yield.

In this chapter, purification of the recombinant DSC1 components is described with view to more definitive functional analyses *in vitro*. Following purification, attempts to obtain preliminary structural data for these proteins are described.

6.2 Purification of Recombinant His-tag fusion proteins using the BioCAD® system

Cloning, overexpression and solubility of the recombinant His-tagged DSC1 components has been described previously (Chapters 3 and 5).

The His-Res1p, His-Res2p and His-Rep2p recombinant proteins were purified from large-scale bacterial cultures by the method of Immobilised Metal ion Affinity Chromatography - IMAC (Porath et al., 1975; Smith et al., 1988) using the BioCAD® SPRINT™ Perfusion Chromatography® system (Methods 2.4.16-2.4.18). This is a fully automated purification system that employs a pre-programmed purification protocol. The ultraviolet (UV) elution profile is monitored at 280 nm, and can be plotted against a choice of parameters (e.g. % elution buffer, pH and conductivity).

In IMAC, a metal chelating group (typically imidodiacetate; $CH_2N(CH_2CO_2^-)_2$) is immobilised on a hydrophilic chromatographic support. The imidodiacetate groups chelate transition metal ions, allowing the remaining co-ordination sites to interact with proteins. The type of metal ion used can affect the binding strength, with copper providing tightest binding ($Cu^{2+} > Ni^{2+} > Zn^{2+} > Co^{2+}$). Proteins bind to these metal co-ordination sites through their surface amino acids, in particular histidine. Thus, proteins with a high availability of histidine residues (i.e. high affinity for transition metal ions) bind to the column, whilst other proteins elute during washing steps. Therefore, fusion of a 6-histidine tag to a recombinant protein provides an efficient handle for purification by this method. The active group in histidine, involved in binding to the metal co-ordination sites, is imidazole. Elution of bound protein is therefore achieved through mass action competitive binding, by a gradient of increasing imidazole concentration. Following initial trial purifications (e.g. using cobalt or nickel) the optimal conditions for binding and elution were achieved using zinc.

6.2.1 Purification of His-Res1p and His-Rep2p

Large-scale protein induction, preparation of clarified supernatant and preparation of the metal chelate column for purification, were carried out as described (Methods 2.4.17 and 2.4.18). Clarified supernatant was manually injected onto the POROS MC column in 5 ml aliquots, in the presence of Starting buffer. Elution of bound protein was achieved by an increasing linear gradient of imidazole (0.5 mM-500 mM), and was monitored by observing the absorbance profile at 280 nm plotted against % elution buffer; 2 ml fractions were collected automatically.

The results depicted in Figure 6.1 represent a typical purification run for His-Res1p, as recorded by the BioCAD® system. A single sharp peak in absorbance was initially observed, encompassing fractions ~6 to 9, followed by a gradual decline that levelled off around fraction 20. Samples of peak fractions were taken and analysed by SDS-PAGE and Coomassie blue staining, or immunoblotting using an antibody directed against the His-tag (Figures 6.2 and 6.3, respectively). The majority of protein, corresponding to the approximate M_r of His-Res1p (75 kDa), eluted in fractions 8 and 9, with faint bands of similar size visible in fractions 10 through 14. The sharp peak seen in Figure 6.1 correlated with the major bands detected in fractions 8 and 9 following Coomassie blue staining. However, immunoblot analysis of these same fractions indicated that a single protein corresponding to the expected size of His-Res1p was present in fractions 10-14 only (Figure 6.3 - a faint band of the expected size was also seen in fraction 9; fraction 14 is not shown on the blot). These results, indicated that the major peak observed was not consistent with elution of His-Res1p and was most likely due to non-specific binding of *E. coli* protein(s). Instead, His-Res1p appeared to elute in the shoulder to the right of this major peak, with the amount of His-Res1p retrieved following purification by this method being very low (Figure 6.2).

Similar results were obtained for His-Rep2p (data not shown). Manipulation of chromatography conditions had no significant effect on protein yield. A scale-up in the process was carried out in an attempt to maximise the amount of protein loaded, but was similarly ineffectual. The inability to purify appreciable amounts of His-Res1p or His-Rep2p was, therefore, attributed to the low levels at which these proteins were overexpressed. Attempts to increase the expression levels of these proteins have been described previously (Chapter 3, Section 3.4). The results presented represent the optimal conditions achieved to date. The inability to purify His-Res1p or His-Rep2p in significant amounts has implications for the future functional and structural study of these proteins.

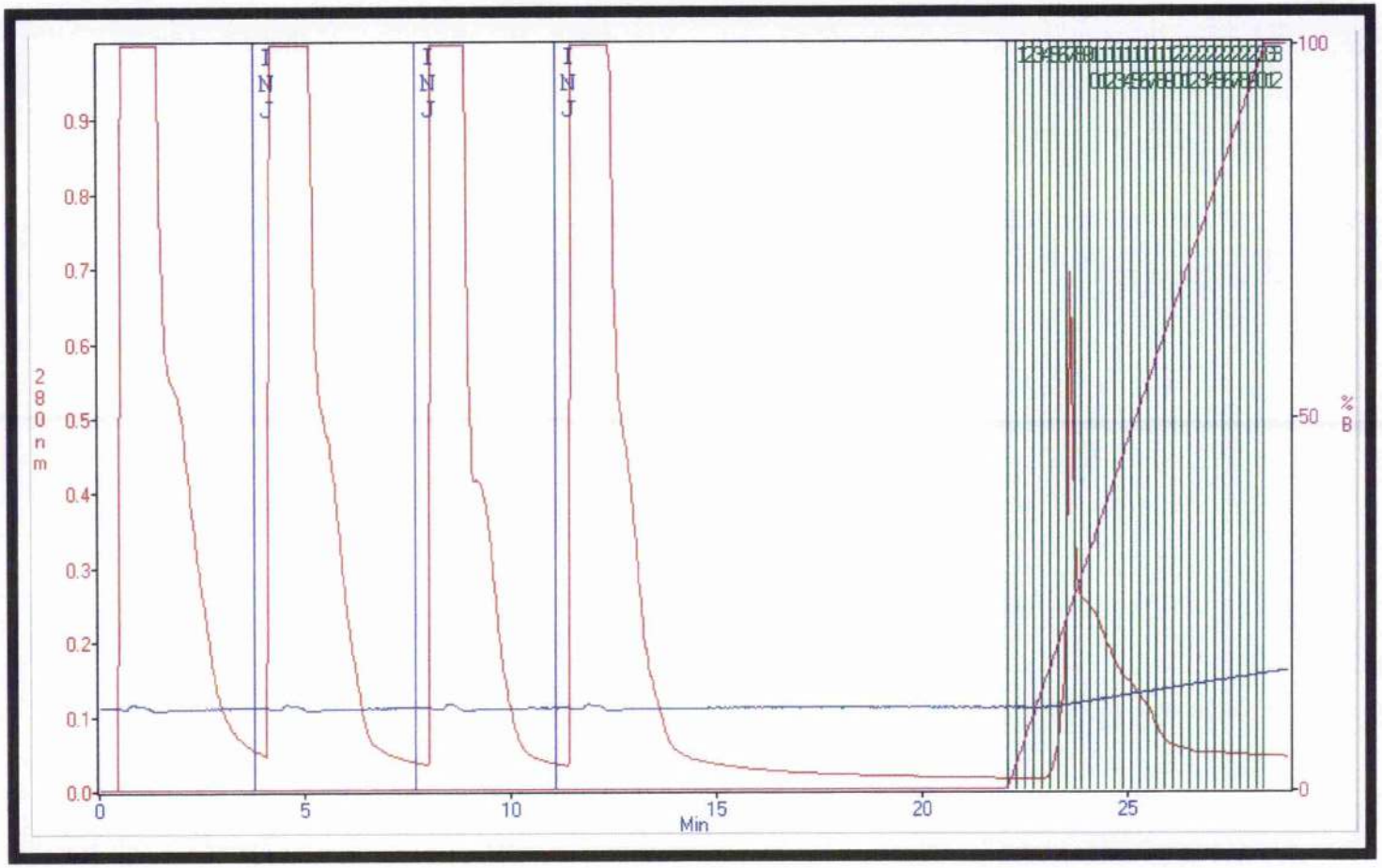


Figure 6.1 Purification profile for His-Res1p
The absorbance at 280 nm is indicated in red, whilst the conductivity and imidazole gradient are indicated in blue and pink, respectively. Fractions (numbered in green) were collected as indicated and aliquots analysed by 10% SDS-PAGE.

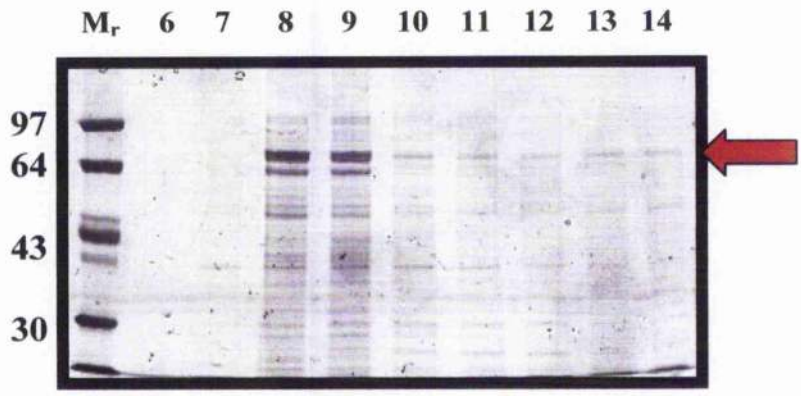


Figure 6.2 SDS-PAGE analysis of His-Res1p from IMAC purification

His-Res1p was overexpressed in BL21 CodonPlus (DE3)-RIL cells for 16 h at 15°C by induction with 0.2 mM IPTG. The bacterial culture (routinely 500 ml) was then centrifuged, the pellet resuspended in 20 ml lysis buffer, and cells lysed by French pressure disruption. The soluble fraction was obtained following centrifugation and loaded onto the column in 5 ml aliquots. Peak fractions were collected automatically and samples (100 µl) taken. All samples were diluted in an equal volume of Laemmli sample buffer and denatured by boiling for 5 min in the presence of DTT (150 mM). Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. Molecular weight markers are shown (M_r) with sizes indicated in kDa. The arrow indicates the expected position of full-length His-Res1p and numbers above each lane correspond to the fraction numbers indicated on the BioCAD® trace (Figure 6.1).

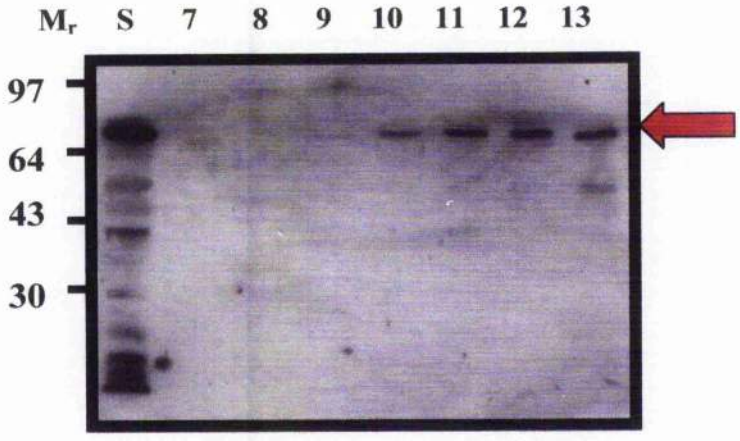


Figure 6.3 Immunoblot analysis of His-Res1p from IMAC purification

The peak fractions shown in Figure 6.2 were also analysed by immunoblot analysis. Following SDS-PAGE and transfer to nitrocellulose, the blot was probed with Anti-PentaHis-HRP conjugate antibody (Qiagen) at 1:2500. The positions of molecular weight markers are indicated (M_r) in kDa. The arrow indicates the expected position of full-length His-Res1p. A sample of the His-Res1p supernatant loaded onto the column is shown (lane S) and numbers above each lane correspond to the fraction numbers indicated on the BioCAD® trace (Figure 6.1).

6.2.2 Purification of His-Res2p

Large-scale protein induction, preparation of clarified supernatant and preparation of the metal chelate column for purification, were carried out as described (Methods 2.4.17 and 2.4.18). Loading of supernatant and elution were carried out as described previously for purification of His-Res1p (Section 6.2.1) and 2 ml fractions collected automatically. The results depicted in Figure 6.4 represent a typical purification run for His-Res2p, as recorded by the BioCAD® system. A single peak in absorbance was observed, encompassing fractions ~6 to 13, that peaked at fractions 9-10 and reached its nadir around fraction 17. Samples of peak fractions were taken and analysed by SDS-PAGE and Coomassie blue staining, or immunoblotting using an antibody directed against the His-tag (Figures 6.5 and 6.6, respectively). The majority of protein, corresponding to the approximate M_r of His-Res2p (75 kDa), eluted in fractions 7 through 13, as viewed by Coomassie blue staining, coincident with the major peak in absorbance (compare Figures 6.4 and 6.5). Immunoblot analysis of these same fractions detected a single protein corresponding to the expected size of His-Res2p; a faint band of the expected size was also seen in fraction 9 (Figure 6.6). These results indicated that the major peak in absorbance was consistent with elution of His-Res2p.

Purification of recombinant His-Res2p using this method yielded significant amounts of protein (typically 5-10 mg l^{-1} culture). Despite such a successful yield, the purity was compromised. A single contaminating protein was consistently observed to co-purify with His-Res2p, in all peak fractions analysed. This contaminant protein was ~ 60 kDa in size as viewed by Coomassie blue staining (compare with Figure 5.1 Chapter 5, for approximate size). In contrast to His-Res2p, this protein was not detected by immunoblotting using anti-His antibody.

In an attempt to remove this impurity, ion-exchange chromatography was attempted. This technique separates proteins by exploiting differences in ionic binding strength. The requirement for high salt concentrations in buffers used in this technique was complicated by the insolubility of His-Res2p in NaCl concentrations >100mM (Chapter 5 Section 5.3). Following IMAC, dialysis of eluted protein into ion-exchange buffers frequently resulted in protein precipitation, precluding use of this technique.

The inability to detect the contaminant protein following immunoblot analysis, with anti-His antibody, demonstrated that this protein did not contain the N-terminal His-tag that was present in full-length His-Res2p. It was therefore assumed that the co-purifying protein was either from *E. coli*, bound non-specifically to the column (or to His-Res2p), or alternatively an N-terminally degraded product of His-Res2p (missing the His-tag).

Close examination of the immunoblot analysis of His-Res2p purification revealed a small protein of approximately 20 kDa (Figure 6.6). It was therefore possible that this represented the N-terminal His-tag 'fragment' missing from the 'truncated' contaminant

protein. In order to identify the contaminant and confirm that the 75 kDa protein was His-Res2p, mass spectrometric analysis was undertaken. In the following section, the 75 kDa protein assumed to be His-Res2p is referred to as protein X and the co-purifying 60 kDa contaminant as protein Y.

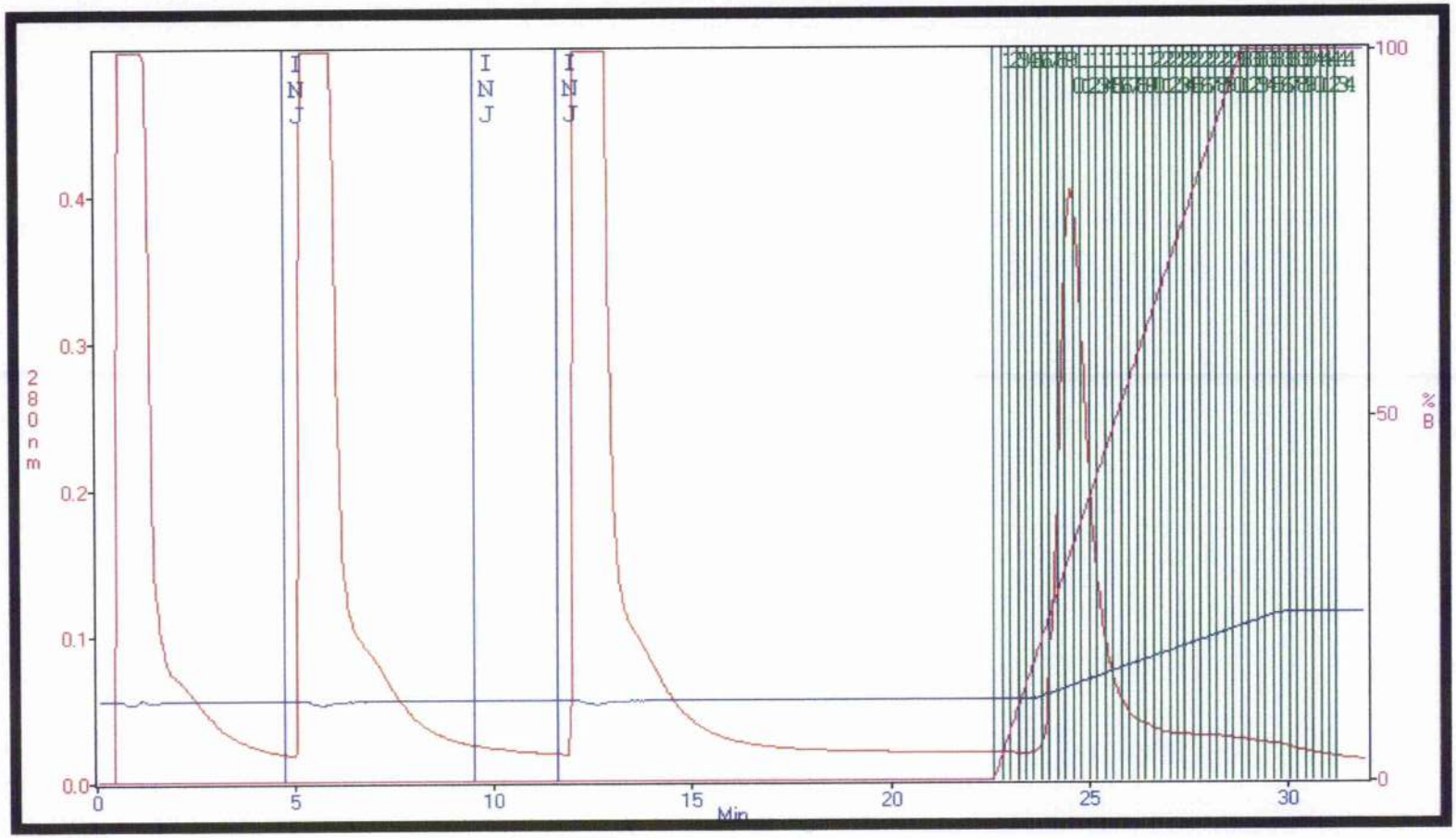


Figure 6.4 Purification profile for His-Res2p
The absorbance at 280 nm is indicated in red, whilst the conductivity and imidazole gradient are indicated in blue and pink, respectively. Fractions (numbered in green) were collected as indicated and aliquots analysed by 10% SDS-PAGE.

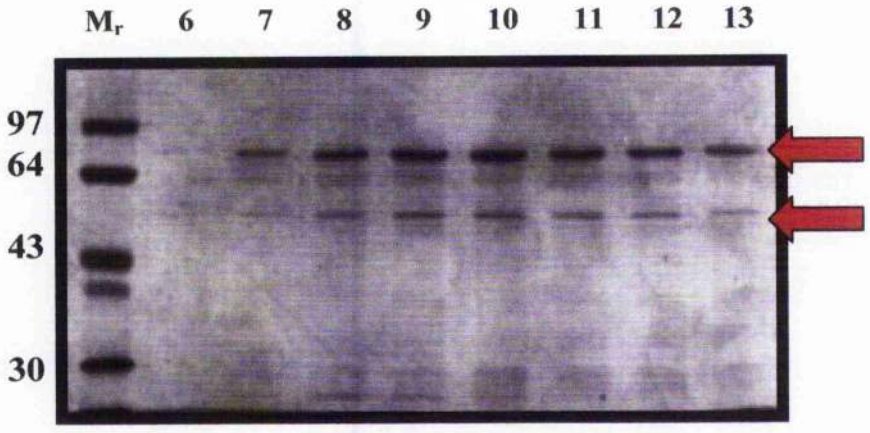


Figure 6.5 SDS-PAGE analysis of His-Res2p from IMAC purification

His-Res2p was overexpressed in BL21 CodonPlus (DE3)-RIL cells for 16 h at 15°C by induction with 0.2 mM IPTG. The bacterial culture (routinely 500 ml) was then centrifuged, the pellet resuspended in 20 ml lysis buffer, and cells lysed by French pressure disruption. The soluble fraction was obtained following centrifugation and loaded onto the column in 5 ml aliquots. Peak fractions were collected automatically and samples (100 µl) taken. All samples were diluted in an equal volume of Laemmli sample buffer and denatured by boiling for 5 min in the presence of DTT (150 mM). Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. The molecular weight markers are shown (M_r) with sizes indicated in kDa. The upper arrow indicates the expected position of full-length His-Res2p and the lower arrow indicates the position of the co-purifying protein. Numbers above each lane correspond to the fraction numbers indicated on the BioCAD® trace (Figure 6.4).

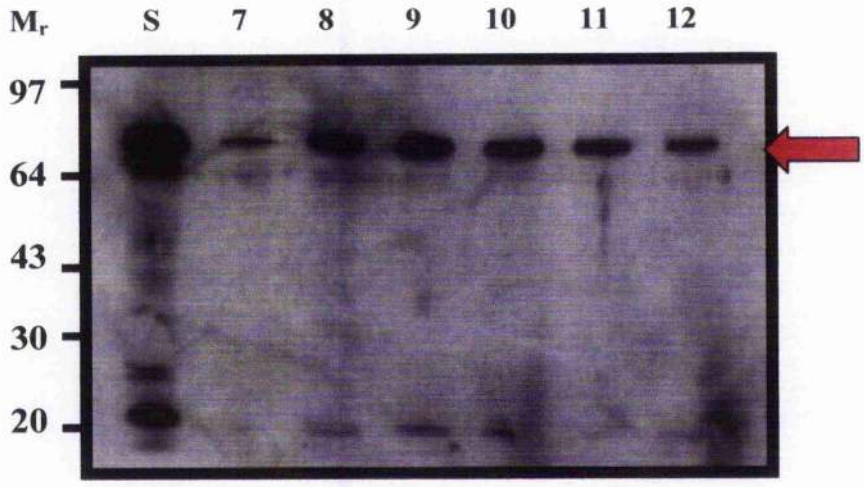


Figure 6.6 Immunoblot analysis of His-Res2p from IMAC purification

Peak fractions (7-12) from the purification were collected and samples taken. Samples were then analysed on a 10% SDS-polyacrylamide transferred to nitrocellulose and blotted with Anti-PentaHis-HRP conjugate antibody (Qiagen) at 1:2500. The positions of molecular weight markers are indicated (M_r) in kDa. The arrow indicates the expected position of full-length His-Res2p. A sample of the His-Res2p supernatant loaded onto the column is shown (lane S) and numbers above each lane correspond to the fraction numbers indicated on the BioCAD® trace (Figure 6.4).

6.2.3 Identification of protein X and protein Y

The identity of proteins X and Y, purified by IMAC in section 6.2.2, were confirmed by mass spectrometric analysis. Protein identification by mass spectrometry was carried out in the Sir Henry Wellcome Functional Genomics Facility (University of Glasgow) using standard methodology (Methods 2.4.20). A brief overview of this technique and its application to protein identification is described.

6.2.4 Protein identification by mass spectrometry

Mass spectrometry is a highly accurate and sensitive analytical technique used for measuring molecular weight. The molecular masses of proteins, and other biomolecules, can be accurately determined to within 0.01% of their total molecular weight (Pitt, 1996). Mass spectrometric analysis is achieved through the initial generation of gas phase ions of the molecule of interest, followed by analysis of the mass to charge (m/z) ratio of these ions. A mass spectrometer is essentially composed of three main parts: a source, in which ions are generated from the substance to be analysed; an analyser, in which ions are separated according to their mass; and a detector, in which a signal produced from the separated ions is transmitted to a data system where it is recorded in the form of a mass spectrum. There are several variations of this technique, primarily differing in the types of ionisation source and analyser used.

The application of mass spectrometry to biochemical analysis is particularly advantageous for protein identification by peptide sequencing, using tandem mass spectrometry (MS/MS). An MS/MS instrument typically contains two analysers separated by a collision cell. The first analyser is used to select sample ions according to their mass to charge (m/z) ratio. These selected (or 'parent') ions are then passed into the collision cell where they are fragmented by collision with gas molecules. The fragment (or 'daughter') ions are then analysed by the second analyser, again according to their m/z ratio. This technique is of particular use for generating peptide sequence information since the daughter ions (i.e. fragmented peptide) derive from a specifically selected parent ion (i.e. intact peptide), allowing information dependent on the amino acid sequence of the peptide to be obtained (Sheehan, 2000).

A schematic diagram, representing the process of protein identification using MS/MS is shown in Figure 6.7. The protein to be analysed must firstly be digested by a protease (e.g. trypsin). This generates peptides that are characteristic to the digest pattern of that protein. Clean up and separation of fragments in a digest mixture (e.g. by chromatography) prior to delivery into the source, allows the sequential analysis of individual peptides. Each of these peptides is then introduced into the mass spectrometer and, following ionisation, analysed ('weighed') to produce a mass spectrum. Analysis of the mass spectra produced from the intact peptides (known as a peptide mass fingerprint), can be used to search an on-line database in an

attempt to identify the protein. In addition, subsequent selection of specific parent ions and passage into the collision cell (where they are fragmented) produces a daughter ion mass spectrum. As fragmentation generally occurs by breaking of one of the peptide bonds in the molecule, a set of fragments is generated where the difference in mass between the fragments can be used to deduce the amino acid sequence. Consequently, the database search using MS/MS ion data is more specific, as it contains primary structure information. Protein identification is ultimately accomplished using complex computer algorithms that compare the experimental data (i.e. peptide molecular weights from the digested protein and/or MS/MS ion data from one or more of these peptides), with values calculated from a comprehensive primary structure database. The Mascot computer program searches primary structure databases for proteins matching the experimental data (www.matrixscience.com). A probabilistic scoring algorithm is then applied allowing proteins to be ranked in order of closest match. The absolute score reflects the fit of the theoretical and experimental data, and is typically reported as $-10*\log_{10}(P)$, where P is the probability. Thus, a low probability (that the experimental and calculated data match is random) is reflected in a high score (e.g. a P value of 10^{-20} becomes a score of 200). Furthermore, the significance of a match is measured against a threshold score, calculated for a random hit against a given database using the search parameter set chosen for the individual search. The threshold has a default setting of 5%, such that, scores below the threshold have a > 5% probability that the match is random. Conversely, scores above the threshold have a < 5% probability that the match is random.

Mass spectral data from proteins X and Y was used in a Mascot search. In both cases the top scoring protein was Res2p from *S. pombe*, with protein scores of 239 and 160, for X and Y respectively (the threshold value was 74). These scores represent the probability of a random match of approximately 1 in 10^{17} and 1 in 10^{9}, respectively. Therefore, it was concluded that protein X (approx. 75 kDa) was full-length His-Res2p whereas the co-purifying protein Y (approx. 60 kDa) was an N-terminally truncated product of His-Res2p. This was backed up by the absence from the dataset for the truncated form of some N-terminal peptides identified in the full-length transcript. The number of peptides matched and % sequence coverage of Res2p, from the database search using experimental data from proteins X and Y, are shown in Figures 6.10 and 6.11, respectively. Figures 6.8 and 6.9 depict chromatographic and mass spectral data obtained from analysis of protein X and are shown as examples. Similar data was retrieved from analysis of protein Y (data not shown).

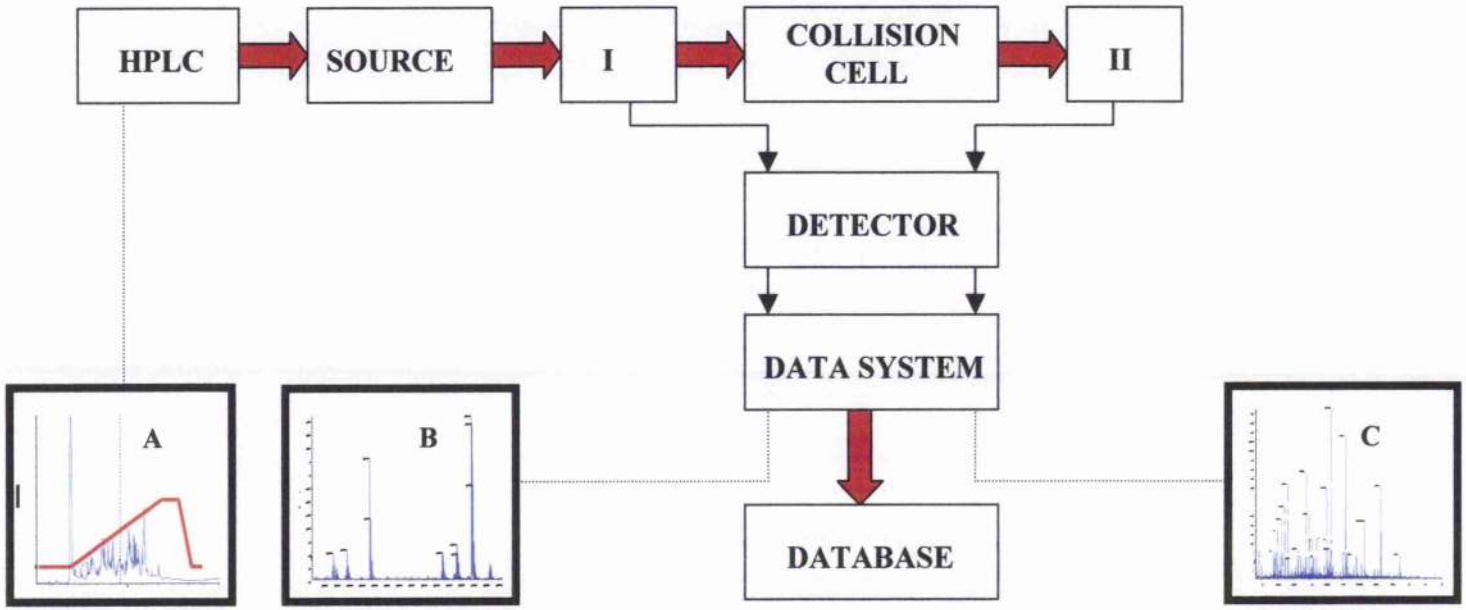


Figure 6.7 Schematic representation of protein identification by tandem mass spectrometry
1. Following trypsin digestion, peptides are separated by reversed-phase High Pressure Liquid Chromatography (HPLC) and sequentially introduced into the ionisation source. 2. Gas phase ions of each peptide ('parent') are then generated by electrospray ionisation (ESI) and analysed (**I**) to generate a parent ion mass spectrum (peptide mass fingerprint). 3. Selected peptide ions are then delivered into the collision cell where they are fragmented. The fragment ('daughter') ions are subsequently analysed (**II**) and a daughter ion mass spectrum produced. 4. The parent and daughter ion mass spectral information produced in the data system is then used in a database search to identify the protein. The chromatogram (Box **A**), parent ion (Box **B**) and daughter ion (Box **C**) mass spectra, produced from analysis of protein X, are shown in greater detail in Figures 6.8 and 6.9. Information retrieved from a database search, using the experimental data from proteins X and Y, is shown in Figures 6.10 and 6.11, respectively.

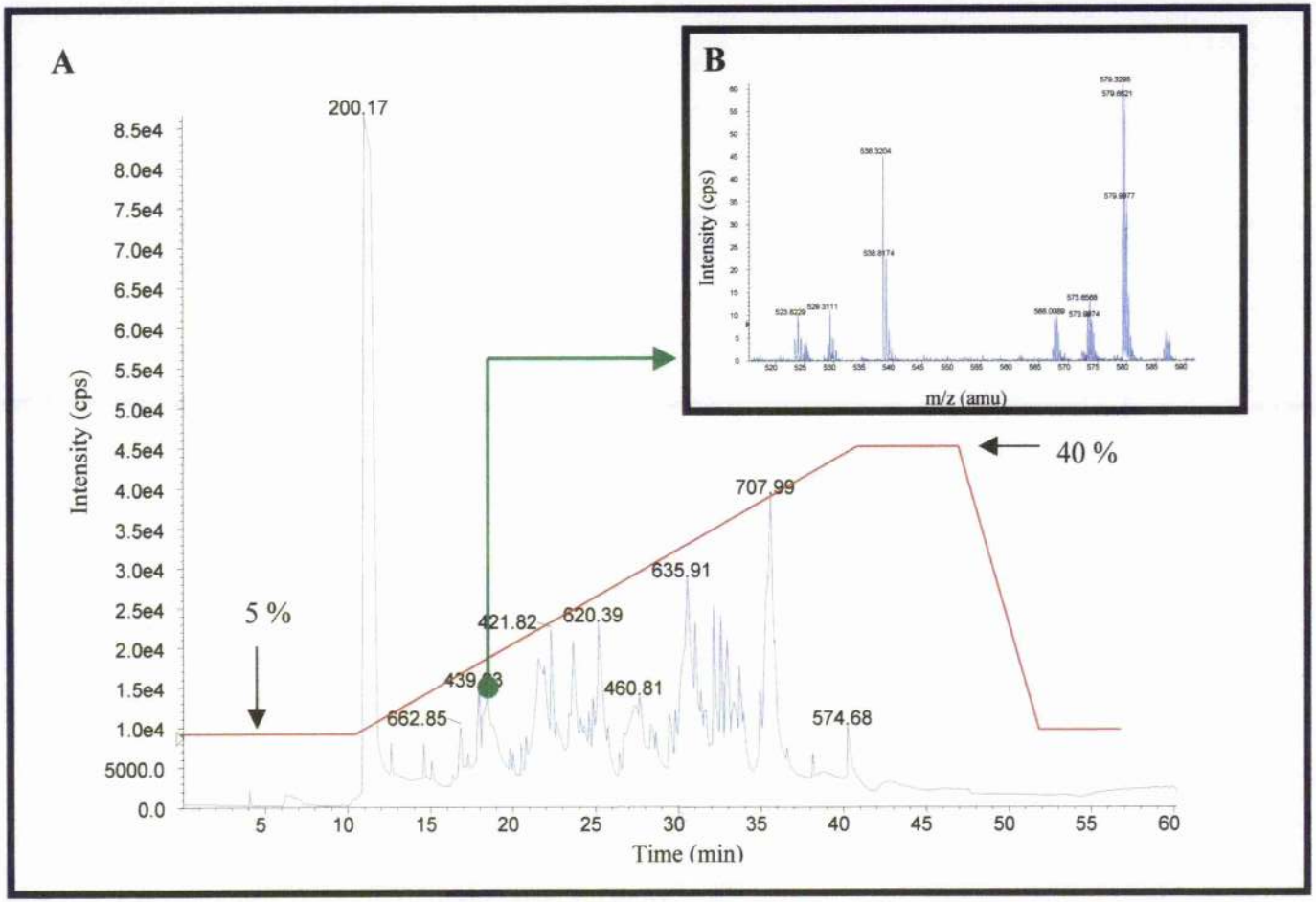


Figure 6.8 Chromatogram

A. Chromatogram produced from reversed-phase HPLC separation of peptides from a trypsin digest of protein X. The red line represents the acetonitrile gradient. **B**. Mass spectral analysis of a peptide purified by HPLC. The region on the chromatogram corresponding to elution of this peptide is highlighted (green arrow).

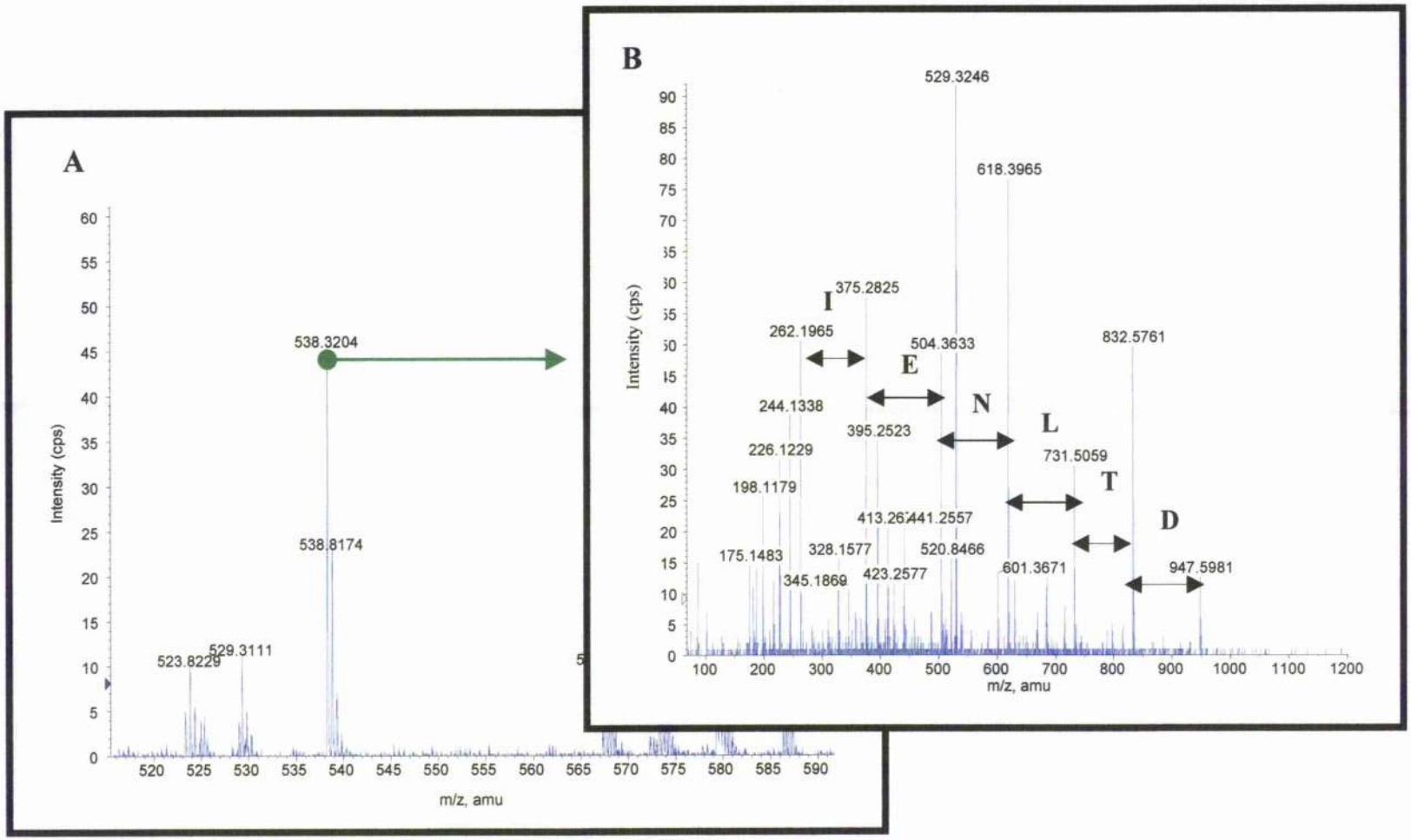


Figure 6.9 Parent and daughter ion mass spectra

A The parent ion mass spectrum depicted in Figure 6.8 B is shown here in greater detail. **B** (inset). The daughter ion mass spectrum produced upon fragmentation of the peptide highlighted in **A** (green arrow). The amino acid sequence of the parent peptide is indicated by the single letter codes and reads from right to left. The actual mass of this peptide (underlined in Figure 6.10) is approximately 1070 Da and thus the parent ion carries a charge value of 2.

MAPRSSAVHV AVYSGVEVYE CFIKGVSVMR RRRDSWLNAT QILKVADFDK PQRTRVLERQ
VQIGAHEKVQ GGYGKYQGTW VPFQRGVDLA TKYKVDGIMS PILSLDIDEG KAIAPKKKQT
KQKKPSVRGR RGRKPSSLSS STLHSVNEKQ PNSSISPTIE SSMNKVNLPG AEEQVSATPL
PASPNALLSP NDNTIKPVEE LGMLEAPLDK YEESLLDFFL HPEEGRIPSF LYSPPPDFQV
NSVIDDDGHT SLHWACSMGH IEMIKLLLRA NADIGVCNRL SQTPLMRSVI FTNNYDCQTF
GQVLELLQST IYAVDTNGQS IFHHIVQSTS TPSKVAAAKY YLDCILEKLI SIQPFENVVR
LVNLQDSNGD TSLLIAARNG AMDCVNSLLS YNANPSIPNR QRRTASEYLL EADKKPHSLL
QSNSNASHSA FSFSGISPAI ISPSCSSHAF VKAIPSISSK FSQLAEEYES QLREKEEDLI
RANRLKQDTL NEISRTYQEL TFLQKNNPTY SQSMENLIRE AQETYQQLSK RLLIWLEARQ
IFDLERSLKP HTSLSISFPS DFLKKEDGLS LNNDFKKPAC NNVTNSDEYE QLINKLTSLQ
ASRKKDTLYI RKLYEELGID DTVNSYRRLI AMSCGINPED LSLEILDAVE EALTREK

Figure 6.10 Matched peptides from protein X
The amino acid sequence of SpRes2p is shown above with matched peptides from the MS/MS ion data search of protein X, shown in **red**. Sequence coverage was 43% and 38 peptides were matched. The daughter ion mass spectrum of the peptide underlined (QDTLNEISR) is shown in Figure 6.9B.

MAPRSSAVHV AVYSGVEVYE CFIKGVSVMR RRRDSWLNAT QILKVADFDK PQRTRVLERQ
VQIGAHEKVQ GGYGKYQGTW VPFQRGVDLA TKYKVDGIMS PILSLDIDEG KAIAPKKKQT
KQKKPSVRGR RGRKPSSLSS STLHSVNEKQ PNSSISPTIE SSMNKVNLPG AEEQVSATPL
PASPNALLSP NDNTIKPVEE LGMLEAPLDK YEESLLDFFL HPEEGRIPSF LYSPPPDFQV
NSVIDDDGHT SLHWACSMGH IEMIKLLLRA NADIGVCNRL SQTPLMRSVI FTNNYDCQTF
GQVLELLQST IYAVDTNGQS IFHHIVQSTS TPSKVAAAKY YLDCILEKLI SIQPFENVVR
LVNLQDSNGD TSLLIAARNG AMDCVNSLLS YNANPSIPNR QRRTASEYLL EADKKPHSLL
QSNSNASHSA FSFSGISPAI ISPSCSSHAF VKAIPSISSK FSQLAEEYES QLREKEEDLI
RANRLKQDTL NEISRTYQEL TFLQKNNPTY SQSMENLIRE AQETYQQLSK RLLIWLEARQ
IFDLERSLKP HTSLSISFPS DFLKKEDGLS LNNDFKKPAC NNVTNSDEYE QLINKLTSLQ
ASRKKDTLYI RKLYEELGID DTVNSYRRLI AMSCGINPED LSLEILDAVE EALTREK

Figure 6.11 Matched peptides from protein Y
The amino acid sequence of SpRes2p is shown above with matched peptides from the MS/MS ion data search of protein Y, shown in **red**. Sequence coverage was 33% and 23 peptides were matched.

6.3 Purification of GST-Res1p and GST-Res2p

Cloning, overexpression and solubility of the recombinant GST-Res1p and GST-Res2p proteins has been described previously (Chapter 5 Sections 5.7-5.8). GST-Res1p and GST-Res2p were purified from large-scale bacterial cultures by affinity chromatography using the BioCAD® SPRINT™ Perfusion Chromatography® system, attached to a GSTrap FF 5 ml column (Methods 2.4.13-2.4.15).

Large-scale protein induction and preparation of clarified supernatants for purification were carried out as described (Methods 2.4.2 and 2.4.14). Clarified supernatant was manually injected onto the GSTrap FF 5 ml column in 5 ml aliquots, in the presence of 1 x PBS. Elution of bound protein was achieved by addition of elution buffer (containing 10 mM reduced glutathione), and was monitored by observing the absorbance profile at 280 nm; 2 ml fractions were collected automatically.

The results depicted in Figure 6.12 represent a typical purification run for GST-Res2p, as recorded by the BioCAD® system. A single peak in absorbance was observed, encompassing fractions ~ 6 to 8. Samples of peak fractions were taken and analysed by SDS-PAGE and Coomassie blue staining (Figure 6.13). The majority of protein, corresponding to the approximate M_r of GST-Res2p, eluted in fractions 7 and 8 with a faint band of similar size visible in fraction 9. Indeed, the peak seen in Figure 6.10 correlated with the major bands detected in fractions 7 and 8 following Coomassie blue staining. Therefore, these results indicated that the major peak in absorbance was consistent with elution of GST-Res2p.

Despite the majority of GST-Res2p being present in the supernatant fraction, the yield following purification was poor. Similar results were obtained upon purification of GST-Res1p with even lower yield (attributed to the lower expression levels observed with this protein, in comparison to GST-Res2p - data not shown).

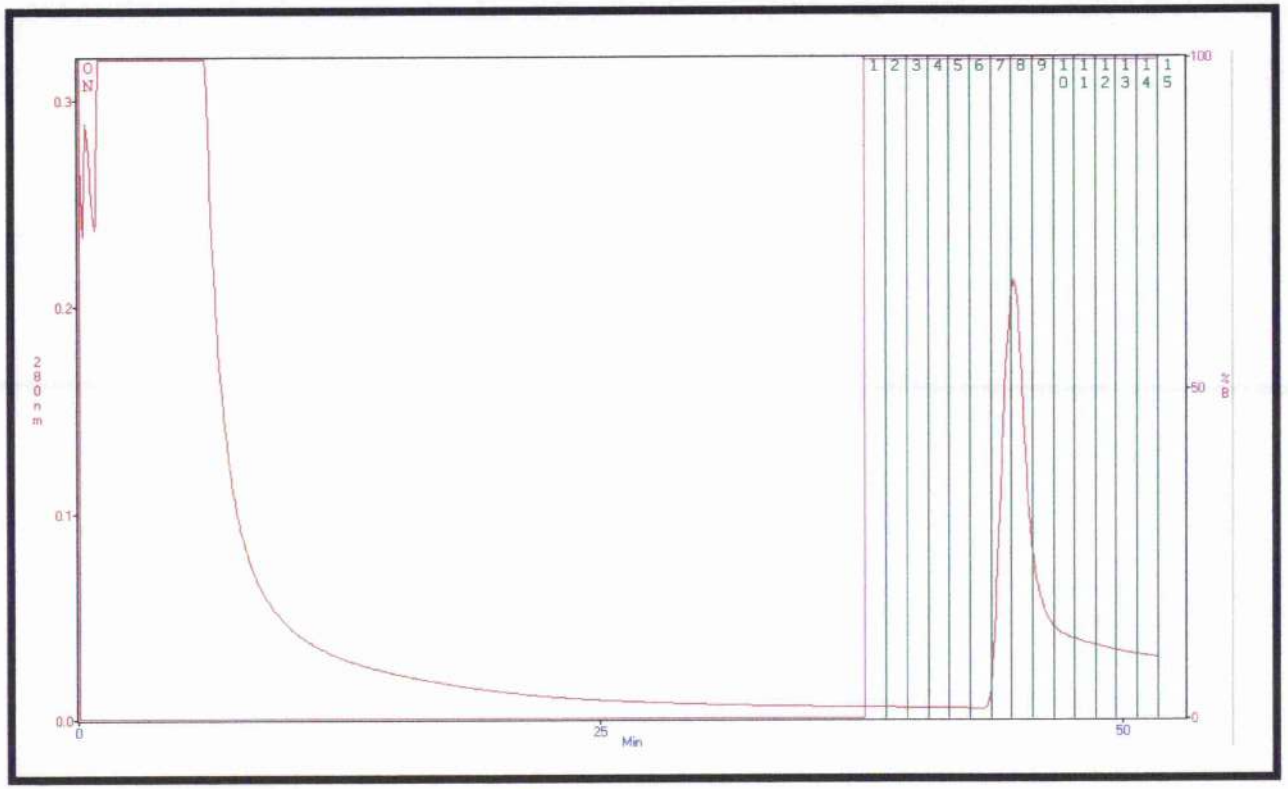


Figure 6.12 Purification profile for GST-Res2p
The absorbance at 280 nm is indicated in red, whilst the % of elution buffer is indicated in pink. Fractions (numbered in green) were collected as indicated and aliquots analysed on 10% SDS-PAGE.

Figure 6.13 SDS-PAGE analysis of GST-Res2p from the GSTrap column

GST-Res2p was overexpressed in BL21 CodonPlus (DE3)-RIL cells for 16 h at 15°C by induction with 0.2 mM IPTG. The bacterial culture (routinely 500 ml) was then centrifuged, the pellet resuspended in 20 ml lysis buffer, and cells lysed by French pressure disruption. The soluble fraction was obtained following centrifugation and loaded onto the column in 5 ml aliquots. Peak fractions were collected automatically and samples (100 µl) taken. All samples were diluted in an equal volume of Laemmli sample buffer and denatured by boiling for 5 min in the presence of DTT (150 mM). Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. Molecular weight markers are shown (M_r) with sizes indicated in kDa. The arrow indicates the expected position of full-length GST-Res2p. The whole cell extract (W), soluble (S) and insoluble (P) samples are shown and numbers above each lane correspond to the fraction numbers indicated on the BioCAD® trace (Figure 6.12).

6.4 Purification of His-Rep1p and His-Cdc10p from inclusion bodies

As described previously, both His-Rep1p and His-Cdc10p were insoluble following overexpression in *E. coli* under all conditions tested (Chapter 5, Sections 5.2 and 5.10). Consequently, this prevented assay of their biological activity *in vitro*. In addition, with regard to His-Cdc10p, this also had implications for the *in vitro* assays of His-Res1p and His-Res2p (Chapter 5, Sections 5.6 and 5.10).

High-level overexpression of recombinant proteins in *E. coli* can often result in the formation of inclusion bodies. Inclusion bodies are thought to contain aggregates of mis-folded protein. Although the protein may contain the correct primary structure, proper folding (and hence full biological activity) may require post-translational modification and/or molecular chaperones (Lilie et al, 1998). The accumulation of a recombinant protein in inclusion bodies can nevertheless have a number of advantages: primarily, accumulation of high levels of protein in a very pure form. Native protein can be recovered following solubilisation of inclusion bodies in denaturing solvents or detergents, followed by re-folding/re-naturation and so the solubilisation and purification of His-Rep1p and His-Cdc10p from inclusion bodies was attempted.

6.4.1 Solubilisation of His-Rep1p and His-Cdc10p using N-lauroylsarcosine

Solubilisation of both His-Rep1p and His-Cdc10p was based on the adaptation of a technique described in the Novagen protein folding kit. Attempts to solubilise both proteins were made in a range of detergents, with greatest success found with buffers containing N-lauroylsarcosine. This is an anionic detergent that has been reportedly successful in solubilising proteins from inclusion bodies to their native or near native state, without affecting biological activity (Frangioni & Neel, 1993).

His-Rep1p and His-Cdc10p were overexpressed in *E. coli* as described (Methods 2.4.1). Cells were harvested by centrifugation and the pellet resuspended in an appropriate volume of inclusion body wash buffer. Inclusion bodies were isolated following cell disruption and centrifugation as described (Methods 2.4.19). Subsequently, aliquots of the whole cell extract, supernatant and pellet (obtained following centrifugation), were analysed by SDS-polyacrylamide gel electrophoresis. The solubility of each protein was estimated by visual comparison of the amounts present in supernatant and pellet fractions, following Coomassie blue staining. Figures 6.14 and 6.15 show the results obtained following solubilisation and purification of His-Rep1p and His-Cdc10p, respectively, from inclusion bodies using N-lauroylsarcosine. Both proteins were approximately 100% soluble in buffer containing 0.3% (w/v) N-lauroylsarcosine.

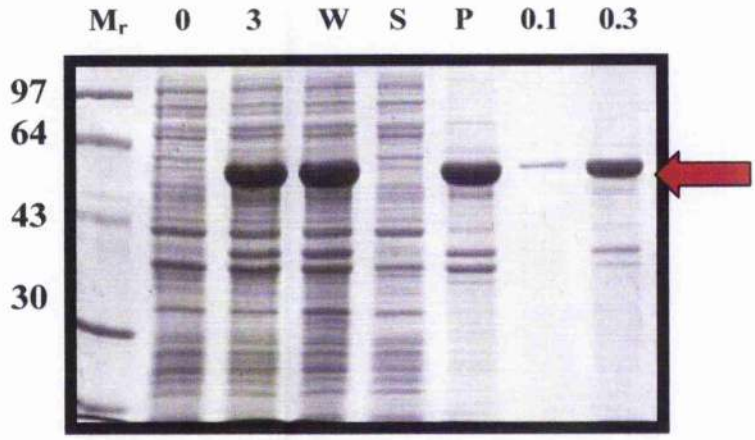


Figure 6.14 SDS-PAGE analysis of His-Rep1p solubilised in N-lauroylsarcosine

His-Rep1p was overexpressed in BL21 CodonPlus (DE3)-RIL cells for 3 h at 37°C by induction with 1 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at 3 h. The bacterial culture was then centrifuged, the pellet resuspended in 5 ml Inclusion body wash buffer, and cells lysed by French pressure disruption. This whole cell extract (W) was then separated into soluble (S) and insoluble (P) fractions by centrifugation and samples (1 ml) taken. The pellet was washed extensively in Inclusion body wash buffer prior to resuspension in 50 mM KH_2PO_4, pH 7.5, supplemented with either 0.1% (v/v) or 0.3% (v/v) N-lauroylsarcosine. Following agitation and centrifugation and an aliquot of supernatant taken for analysis. Samples from 0 h and 3 h were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl/0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). All other samples were diluted with an equal volume of Laemmli sample buffer and similarly denatured. Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. The arrow indicates the expected position of full-length His-Rep1p. Molecular weight markers are shown (M_r) with sizes indicated in kDa.

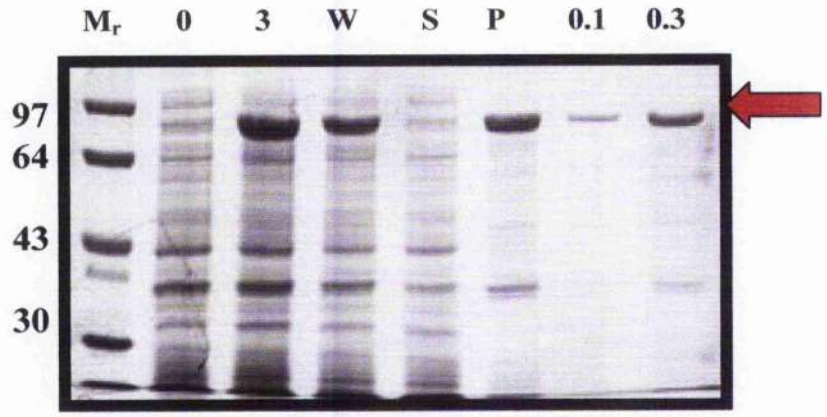


Figure 6.15 SDS-PAGE analysis of His-Cdc10p solubilised in N-lauroylsarcosine

His-Cdc10p was over-expressed in BL21 CodonPlus (DE3)-RIL cells for 3 h at 37°C by induction with 1 mM IPTG. Samples (1 ml) were removed at the point of induction (0 h) and at 3 h. The bacterial culture was then centrifuged, the pellet resuspended in 5 ml Inclusion body wash buffer, and cells lysed by French pressure disruption. This whole cell extract (W) was then separated into soluble (S) and insoluble (P) fractions by centrifugation and samples (1 ml) taken. The pellet was washed extensively in Inclusion body wash buffer prior to re-suspension in 50 mM KH_2PO_4, pH 7.5, supplemented with either 0.1% (v/v) or 0.3% (v/v) N-lauroylsarcosine. Following agitation and centrifugation an aliquot of supernatant was taken for analysis. Samples from 0 h and 3 h were centrifuged, the pellets resuspended in Laemmli sample buffer (10 µl/0.1 absorbance unit) and denatured by boiling for 5 min in the presence of DTT (150 mM). All other samples were diluted with an equal volume of Laemmli sample buffer and similarly treated. Samples were then analysed on a 10% SDS-polyacrylamide gel stained with Coomassie brilliant blue. The arrow indicates the expected position of full-length His-Cdc10p. Molecular weight markers are shown (M_r) with sizes indicated in kDa.

6.5 Investigating the structure and function of the purified recombinant DSC1 proteins

Attempts to purify His-Res1p, His-Rep2p, GST-Res1p and GST-Res2p were unsuccessful in that both yield and purity were unsatisfactory. In contrast, purification of His-Res2p, His-Rep1p and His-Cdc10p yielded amounts sufficient to allow detailed structural analyses to be undertaken. However, the issue of whether these latter three recombinant proteins were biologically active remained unresolved.

Following purification of these proteins, attempts were made to recover them in buffer solutions more appropriate for biophysical analyses. With respect to His-Res2p, the high concentration of imidazole present in the elution buffer had to be removed. Similarly, the level of N-lauroylsarcosine may have to be reduced or removed to permit structural studies of His-Rep1p and His-Cdc10p. Buffer exchange is routinely achieved by dialysis, involving multiple changes of buffer through a semi-permeable membrane with a molecular weight cut-off, or centrifugation using a centricon tube (Methods 2.4.3-2.4.4). Using either approach, all three proteins frequently precipitated despite attempting a variety of conditions (e.g. temperature) and buffers. In addition, thrombin cleavage of His-Res2p was unsuccessful despite manipulation of cleavage conditions. Unfortunately, due to time constraints, studies using purified His-Res2p and His-Cdc10p in band-shift assays were not initiated. Nevertheless, initial attempts to obtain structural information from His-Cdc10p and His-Rep1p were undertaken.

6.6 Circular dichroism

Detailed structural information from a protein is typically obtained from high-resolution techniques such as X-ray crystallography and NMR spectroscopy. Both of these techniques require significant amounts of material (several tens of mg) and results are rarely rapidly obtained. Indeed, for successful X-ray analysis crystals must be produced that diffract sufficiently to permit structural detail at the atomic level. Typically, this may take several weeks, months or even years to achieve. Furthermore, in NMR, full structural determinations are limited to small proteins of maximal molecular weight ~ 30 kDa.

In contrast, circular dichroism (CD) studies require significantly less material (1-2 mg ml^{-1} or less) and results are rapidly obtained (typically within a few hours). Despite these advantages, this is a lower resolution technique and so detailed structural information is not generally obtained from CD spectra. However, the secondary structure content of a protein can be empirically determined. Perhaps the most valuable use of CD is in the study of protein folding and protein-protein, protein-nucleic acid and protein-ligand interactions.

CD is a spectroscopic technique that involves measurement of the differential absorption of the left and right circularly polarised components of plane-polarised light. Upon passage through an optically active sample, one of these components (i.e. left or right) will be absorbed to a greater extent than the other (as a result of the asymmetric nature of chiral molecules). Consequently, subsequent recombination of these components generates elliptically polarised light. The occurrence of ellipticity is called circular dichroism. CD spectra are recorded in a spectropolarimeter where the differential absorption of left and right circularly polarised light is detected and converted into units of ellipticity, θ (degrees cm^2 $dmol^{-1}$).

CD measurements can be performed in two main spectral regions, the far-UV region (240-180 nm) and the near-UV region (320-260 nm). In the far-UV region the principal absorbing species is the peptide bond and so the CD spectrum is sensitive to the main chain conformation. The common secondary structure motifs, α-helix, β-sheet and β-turn exhibit distinct CD spectra in this region, therefore CD studies can be used to estimate the secondary structure content of a protein. In addition, random coil, representing regions of a protein that do not encompass the major secondary structural motifs, can be identified as it absorbs in a region similar to that of β-sheet (Kelly and Price, 2000).

Several computer programs are available to estimate the secondary structure content of a protein from the properties of its CD spectra (Provencher and Glöckner, 1981; Sreerama and Woody, 1993). Experimental CD values are compared to a database containing CD spectra of proteins whose structures have been solved by X-ray diffraction. Using this approach, a secondary structure content estimate of the protein is obtained. The validity of these estimates can then be compared/contrasted with previous known data. Therefore, CD spectra can provide

low-resolution (yet reliable) secondary structural information on a protein. In the near UV region the environment of aromatic amino acids such as tryptophan, tyrosine, phenylalanine and cystinyl groups can be detected. In a folded protein the side chains of these amino acids are likely to be placed in a chiral environment, thus giving rise to spectra that can provide a fingerprint of the tertiary structure of the protein.

In this section, CD was used to determine the secondary structures of His-Cdc10p and His-Rep1p. In addition, the stability and unfolding characteristics of both proteins were investigated, by monitoring changes in their CD spectra following denaturation in increasing concentrations of guanidinium chloride.

6.6.1 Secondary structure determination of His-Rep1p and His-Cdc10p

His-Cdc10p and His-Rep1p were solubilised and purified from inclusion bodies as described in Section 6.4.1. Specifically for CD measurements, inclusion bodies were solubilised in 50 mM KH_2PO_4, 150 mM NaF pH 7.5, supplemented with 0.1% (v/v) N-lauroylsarcosine. The lower amount of 0.1% (v/v) N-lauroylsarcosine was preferred (to 0.3%) as the sample was of greater purity in this preparation and a minimal amount of detergent was more suitable for CD studies. CD spectroscopy was carried out in the Scottish Circular Dichroism Facility, University of Glasgow, using standard methodology (Methods 2.4.21). In all cases, the CD spectrum was obtained following subtraction of the baseline (buffer only) spectrum from the sample spectrum.

6.6.2 CD spectra of His-Rep1p and His-Cdc10p

Purified His-Rep1p (0.24 mg ml^{-1}) and His-Cdc10p (0.22 mg ml^{-1}) were prepared for CD analysis as described above. An initial CD scan revealed that there was a significant level of nucleic acid contamination in both protein preparations. Nucleic acids display distinct CD spectra in the near UV region and therefore interfere with any signal produced from the protein. Consequently, analysis of His-Rep1p and His-Cdc10p in the near-UV region was not possible. An example of the nucleic acid contamination of His-Rep1p is shown in Figure 6.16. Furthermore, analysis of these proteins in the far-UV region was only permitted down to a lower wavelength limit of 180 nm due to the poor signal to noise ratio. Figures 6.17 and 6.18 show the CD spectra obtained for His-Rep1p and His-Cdc10p, respectively. Secondary structure estimates were calculated using the CONTIN procedure (Provencher and Glöckner, 1981) and are displayed in Tables 6.1 and 6.2 for His-Rep1p and His-Cdc10p, respectively. These results indicate that the majority of detergent solubilised full-length His-Rep1p consists of α-helix (~78%), with ~17% β-sheet/turn and the remaining ~6% as random coil. For detergent solubilised full-length His-Cdc10p the majority of the protein consists of β-sheet/turn (~40%), with ~30% α-helix and the remaining ~30% as random coil.

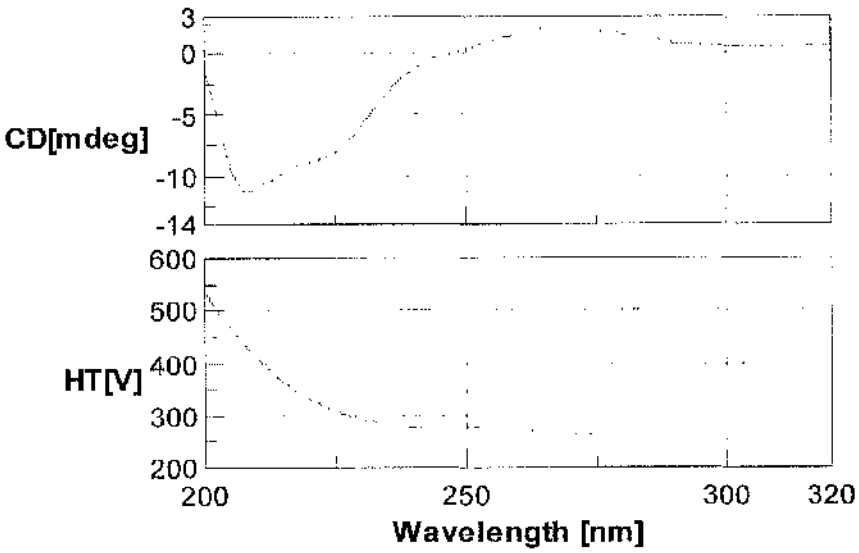


Figure 6.16 Nucleic acid contamination of His-Rep1p

CD spectrum incorporating both near and far-UV regions shows nucleotide contamination of His-Rep1p. Measurements were recorded in a quartz cell of pathlength 0.02 cm, containing purified His-Rep1p (0.24 mg ml^{-1}) in 50 mM KH_2PO_4, 150 mM NaF pH 7.5, supplemented with 0.1% (v/v) N-lauroylsarcosine.

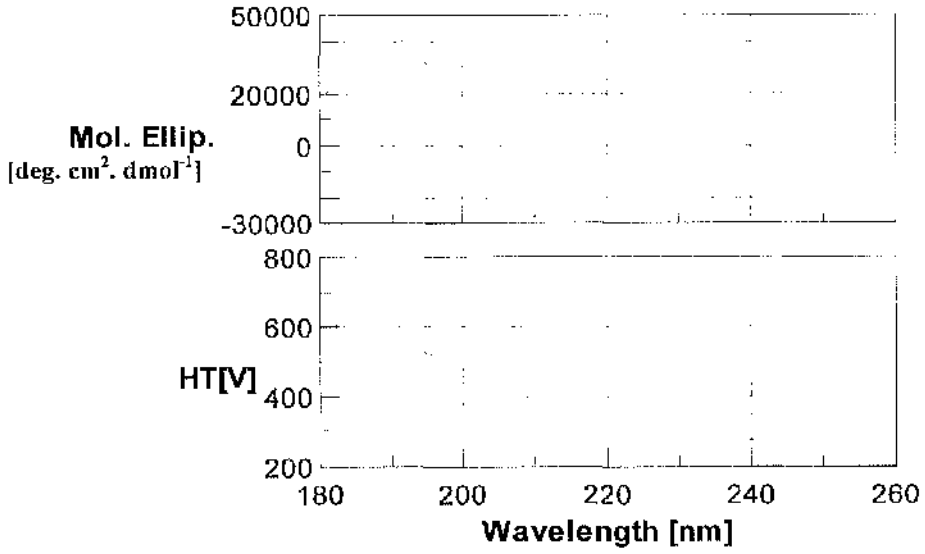


Figure 6.17 Secondary structure determination of His-Rep1p

The far-UV scan of His-Rep1p. Measurements were recorded in a quartz cell of pathlength 0.02 cm, containing purified His-Rep1p (0.24 mg ml^{-1}) in 50 mM KH_2PO_4, 150 mM NaF pH 7.5, supplemented with 0.1% (v/v) N-lauroylsarcosine.

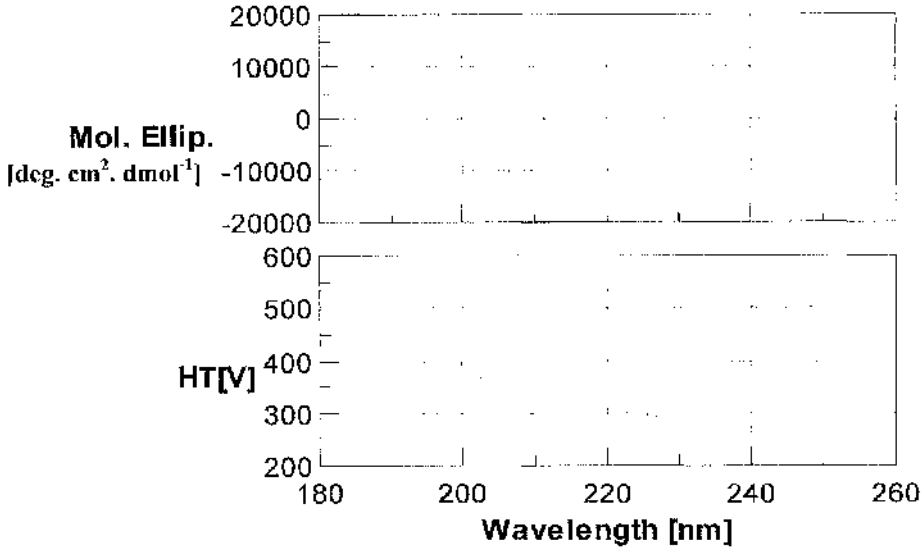


Figure 6.18 Secondary structure determination of His-Cdc10p

The far-UV scan of His-Cdc10p. Measurements were recorded in a quartz cell of pathlength 0.02 cm, containing purified His-Cdc10p (0.22 mg ml^{-1}) in 50 mM KH_2PO_4, 150 mM NaF pH 7.5, supplemented with 0.1%(v/v) N-lauroylsarcosine.

Secondary Structure Element	Total Protein %
α-helix	77.4
β-sheet	2.1
β-turn	14.8
Random coil	5.8

Table 6.1 Secondary structure content estimate of His-Rep1p

The secondary structure content of recombinant His-Rep1p was estimated from analysis of the far-UV CD spectra using the CONTIN procedure (Provencher and Glöckner, 1981).

Secondary Structure Element	Total Protein %
α-helix	29.6
β-sheet	20.7
β-turn	20.9
Random coil	28.8

Table 6.2 Secondary structure content estimate of His-Cdc10p

The secondary structure content of recombinant His-Cdc10p was estimated from analysis of the far-UV CD spectra using the CONTIN procedure (Provencher and Glöckner, 1981).

6.6.3 Assessment of His-Rep1p and His-Cdc10p stability

The stabilities of His-Rep1p and His-Cdc10p were also assessed, following chemical denaturation with guanidinium chloride (GdmCl). Purified His-Rep1p (0.24 mg ml^{-1}) and His-Cdc10p (0.22 mg ml^{-1}) samples were prepared in increasing guanidinium chloride concentrations as described (Methods 2.4.22).

The unfolding of His-Rep1p and His-Cdc10p in GdmCl was monitored by changes in far-UV CD and expressed in terms of % total change of this parameter occurring between 0 M and 6 M GdmCl. The unfolding of His-Rep1p and His-Cdc10p as monitored by CD is presented in graphical form in Figures 6.19 and 6.20, respectively.

From the appearance of the graph shown in Figure 6.19, it appears that there is a biphasic unfolding event occurring. An initial major unfolding event takes place between 0 and 0.5 M GdmCl, followed by an apparent stabilisation until the GdmCl concentration is increased to approximately 2 M. From this point onwards, a second unfolding event begins that ends at 6 M. Similarly, the unfolding of His-Cdc10p follows a biphasic pattern. The initial unfolding event takes place between 0-2 M GdmCl. The second event occurs sharply between 2 M and ~ 3 M and perhaps a third event may be occurring between 3 M and 6 M (Figure 6.20).

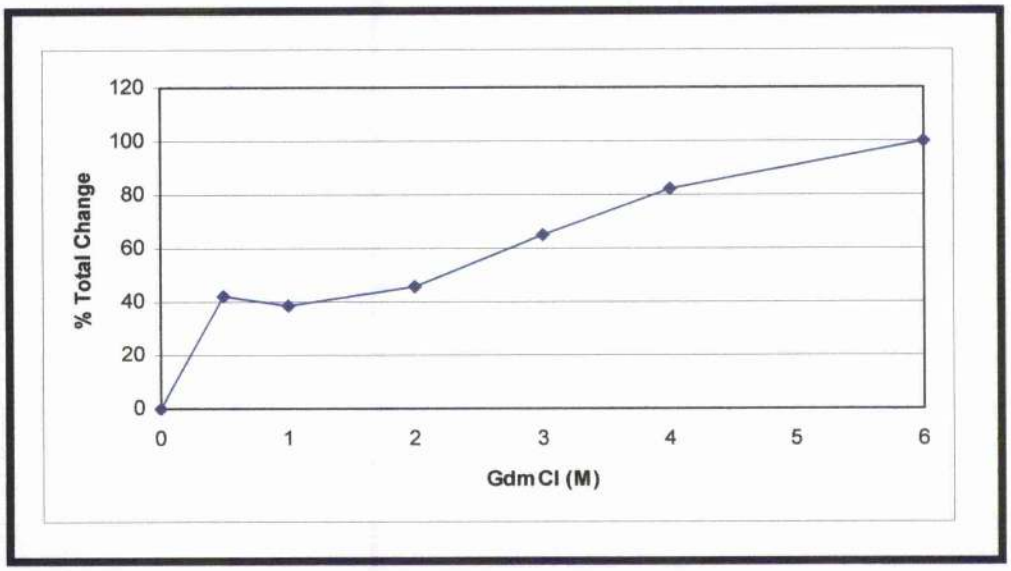


Figure 6.19 Unfolding profile of His-Rep1p during denaturation with guanidinium chloride

Purified His-Rep1p (0.24 mg ml^{-1}) was incubated at room temperature in increasing concentrations of GdmCl: 0 M, 0.5 M, 1 M, 2 M, 3 M, 4 M and 6 M prior to measurement. The CD changes were expressed relative to the total change observed between 0 M and 6 M GdmCl. CD was measured as change in ellipticity (θ) at 222 nm.

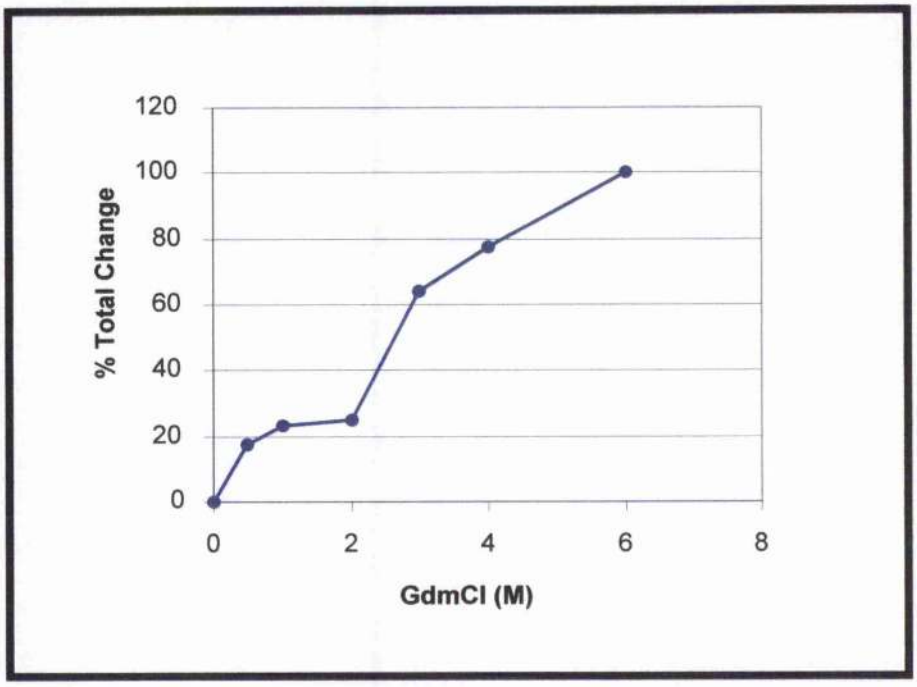


Figure 6.20 Unfolding profile of His-Cdc10p during denaturation with guanidinium chloride

Purified His-Cdc10p (0.22 mg ml^{-1}) was incubated at room temperature in increasing concentrations of GdmCl: 0M, 0.5M, 1M, 2M, 3M, 4 M and 6 M prior to measurement. The CD changes were expressed relative to the total change observed between 0 M and 6 M GdmCl. CD was measured as change in ellipticity (θ) at 222 nm.

6.7 Discussion

The aim of this chapter was to purify the bacterially overexpressed recombinant fission yeast DSC1 components. Following on from the results discussed in Chapter 5, several unresolved issues required further investigation. The initial aim of purification was to obtain protein with which to carry out further *in vitro* functional assays. In addition, purification of protein of sufficient yield and purity would allow the initiation of structural studies.

Unfortunately, both the yield and purity of His-Res1p and His-Rep2p were unsatisfactory, following purification by IMAC. This was attributed to the low level overexpression of these proteins in *E. coli*. Attempts to increase the overexpression level have been discussed previously (Chapter 3 Section 3.4). A scaling up of the purification process neither improved yield nor purity.

In contrast, the yields of the GST-Resp proteins obtained following overexpression in *E. coli* were significantly greater, and it was anticipated that purification would provide large amounts of protein. Disappointingly, neither GST-Res1p nor GST-Res2p was purified in sufficient amounts, following affinity purification on an immobilised glutathione column. Due to time constraints, the optimisation of purification of these proteins has not been fully investigated and as such, the results presented represent preliminary data. Reasons for poor recovery may include steric hindrance of the GST moiety, by intramolecular interactions with the attached Resp protein. It should be noted that in both cases, DNA sequencing has confirmed that the GST-tag was in-frame with the cDNA. These results were particularly frustrating, given that several interesting and potentially informative experiments, regarding the *in vitro* DNA-binding behaviour of these proteins, relied on their purification.

The most favourable case was in purification of His-Res2p. This protein was reproducibly purified in large amounts following IMAC (5-10 mg l^{-1} culture). Despite consistent co-purification with a ~60,000 Da protein, subsequent mass spectrometric analysis identified this protein as an N-terminally truncated product of His-Res2p (Section 6.2.3). Unfortunately, recovery of purified His-Res2p, in buffers suitable for structural studies, frequently resulted in precipitation. Indeed, initial attempts to obtain CD spectra from this protein were unsuccessful, primarily due to precipitation. Intriguingly, the spectral studies displayed a high level of nucleic acid contamination in purified samples of His-Res2p. This was despite the fact that both DNAase and RNAase were routinely used in purification procedures. Given that this protein is a DNA-binding protein its association with nucleic acids is not surprising. Precipitation of His-tagged proteins purified by IMAC has also been noted elsewhere and attributed to contamination with Zn^{2+} ions from the chromatography column (Lindner et al., 1992). However, despite addition of EDTA, precipitation continued.

His-Cdc10p and His-Rep1p were solubilised from inclusion bodies, using the anionic detergent N-lauroylsarcosine. The yield and purity of these proteins was sufficient to initiate structural studies. Similarly to His-Res2p, dialysis into more suitable buffers routinely resulted in precipitation and aggregation of protein. Nevertheless, CD studies were initiated on His-Cdc10p and His-Rep1p, as this technique was somewhat tolerant to the presence of low levels of detergent. Initial CD spectra of these proteins also revealed a significant level of nucleic acid contamination; this precluded near-UV analysis of these proteins and imposed restrictions on the far-UV studies. Far-UV CD spectra were obtained for both proteins permitting secondary structure content estimations and stability studies were initiated by monitoring unfolding following denaturation in increasing concentrations of guanidinium chloride. The results of the CD studies using these proteins should be treated with caution, as neither of these proteins has been shown to be active *in vitro*. It is highly likely that to obtain these proteins in their fully native state, re-folding will be necessary. The results obtained here serve as an example of the type of initial structural information that can be rapidly obtained following successful purification.

Structure-function studies of proteins need large amounts of material. The ability to obtain sufficient amounts requires the development of efficient and reproducible methods for overexpression and purification of the proteins of interest. As exemplified in this (and previous chapters) development of such protocols requires considerable investigation. The necessity to produce large amounts of protein for such studies can frequently result in aggregation and precipitation as has been the case with His-Res2p. Assuming that the problem of precipitation can be overcome, structural studies may then be initiated given the yield and purity of this protein obtained, following purification by IMAC.

With respect to His-Res1p, His-Rep2p and the GST-Resp proteins, optimisation of purification must be addressed in order to provide sufficient material for further study. Several experiments of significant interest, with purified GST-Resp proteins could then be addressed. In summary, due to time constraints, the results presented within this chapter provide a basis for further study.

Chapter 7

General discussion

7.1 Introduction

In both budding and fission yeasts, the highly related MBF/SBF and DSC1 transcription factor complexes are responsible for controlling the G1-S phase-specific gene expression programme essential for passage through START and entry into a new round of mitotic cell division. Similarly, in humans, a G1-S phase-specific gene expression programme exists that is also essential for passage through the Restriction point and entry into mitosis. In contrast to the two yeast systems, the E2F transcription factor family of proteins bear no significant resemblance to any of the yeast proteins at the amino acid sequence level.

In budding yeast, MBF and SBF are heterodimeric transcription factors composed of the single sequence-specific DNA-binding protein subunits, ScMbp1p and ScSwi4p, respectively, combined with the common regulatory subunit ScSwi6p that is present in both complexes. Fission yeast DSC1 is believed to exist as a heterotetramer, consisting of two distinct yet related sequence-specific DNA-binding proteins, SpRcs1p and SpRes2p, each of which binds to a single molecule of the regulatory SpCdc10p protein. The budding yeast and fission yeast proteins constitute a family of related transcription factors that share considerable homology, located to three main regions. All six proteins possess the centrally located ankyrin-repeat motifs, whose primary function is unclear, but may be involved in both protein-protein interactions and in affording flexibility. The ScMbp1p/ScSwi4p and SpRes1p/SpRcs2p proteins also share considerable homology within their N-terminal DNA-binding regions, whilst ScSwi6p and SpCdc10p do so in their C-termini. To date, structural studies on this family of transcription factors have been limited to the budding yeast proteins; specifically the DNA-binding domain of ScMbp1p and ankyrin repeat domain of ScSwi6p (Xu et al., 1997; Taylor et al., 1997; Foord et al., 1999). In contrast, no structural data beyond the amino acid sequence level is available for any of the fission yeast members of this family. Thus, successful cloning and expression of the fission yeast DSC1 components will provide an invaluable resource for detailed structural and functional studies.

7.2 Summary of results

7.2.1 Chapter 3: Cloning and overexpression

In Chapter 3 the cloning strategy was described, which allowed the successful and reproducible overexpression of the individual components of the fission yeast DSC1 complex in *E. coli*. Site-directed mutagenesis was employed to remove an artificial internal stop codon within the *cdc10*$^{+}$ open reading frame.

Each of the individual components of the DSC1 complex: *cdc10*$^{+}$, *res1*$^{+}$, *res2*$^{+}$, *rep1*$^{+}$ and *rep2*$^{+}$ was cloned into the pET-28c vector (pET-14b for *cdc10*$^{+}$) in-frame with an N-terminal 6-histidine tag (to facilitate purification by affinity chromatography) and successfully expressed following induction at 37°C. High-level expression of the Res2p, Rep1p and Cdc10p components of DSC1 as His-tagged fusion proteins was achieved as visualised by SDS-PAGE and Coomassie blue staining (Figures 3.7-3.8). Overexpression of His-Res1p and His-Rep2p was achieved at lower levels, such that they were detectable only following immunoblot analysis (Figures 3.7, 3.10 and 3.9, 3.11).

7.2.2 Chapter 4: Biological activity assays *in vivo*

Chapter 4 described the use of fission yeast (the organism from which these proteins originate) as host to assay the effects of the N-terminal His-tag upon the biological activity of each of the recombinant proteins *in vivo*.

To this end, the cloning strategy was described that facilitated expression of the *His-res1*$^{+}$, *His-res2*$^{+}$, *His-rep1*$^{+}$ and *His-rep2*$^{+}$ cDNAs in fission yeast from the pREP vector. Due to time constraints (primarily the lengthy site-directed mutagenesis protocol), *His-cdc10*$^{+}$ was not cloned into the pREP vector and so the ability of *His- cdc10*$^{+}$ to function *in vivo* was not determined.

The biological activity of each recombinant protein was assayed following expression in the Δ*res1* mutant background. The wild type *res1*$^{+}$, *res2*$^{+}$, *rep1*$^{+}$ and *rep2*$^{+}$ genes can each suppress the cold-sensitive phenotype of this strain (Tanaka et al., 1992; Miyamoto et al., 1994; Sugiyama et al., 1994; Nakashima et al., 1995). Similarly, the *His-res1*$^{+}$, *His-res2*$^{+}$, *His-rep1*$^{+}$ and *His-rep2*$^{+}$ cDNAs each suppressed the cold-sensitive lethality of Δ*res1* cells when ectopically expressed (Figures 4.8-4.11). The biological activities of His-Res1p and His-Res2p were also assayed following expression in Δ*res1* and Δ*res2* mutant backgrounds. Ectopic expression of *res1*$^{+}$ and *res2*$^{+}$ in Δ*res1* and Δ*res2* cells, re-constitutes the characteristic DSC1 band-shift activity that is otherwise lost in these mutant strains (Zhu et al., 1997; Ayte et al., 1997). A 'DSC1-like' band-shift activity was observed in Δ*res2* cells upon ectopic expression of *His-res2*$^{+}$ (Figures 4.12-4.13). Surprisingly, no 'DSC1-like' band-shift activity was detected

when *His-res1*$^+$ was expressed in Δ*res1* cells (Figure 4.14). That each recombinant protein was able to rescue the cold-sensitive lethality of the Δ*res1* mutant suggested that the His-tag imposed no impairment to function *in vivo*.

7.2.3 Chapter 5: Recombinant protein solubility

In Chapter 5, attempts to obtain soluble recombinant protein, following overexpression of His-Res1p, His-Res2p, His-Rep1p, His-Rep2p or His-Cdc10p in *E. coli*, were described. Following this, a series of assays were outlined to determine the biological activity of these bacterially expressed proteins.

A range of expression and lysis buffer conditions were tested to optimise the amount and solubility of each recombinant protein. Both His-Res1p and His-Rep2p were retrieved in the soluble fraction under most conditions tested, although yields were low (Figures 5.3-5.4). The ability to produce soluble His-Res2p was dependent on both induction conditions and the nature of the solvent. Approximately 100% of His-Res2p was retrieved in soluble form following expression at 15°C and importantly, release into lysis buffer containing ≤ 100 mM NaCl (Figures 5.5-5.6). Despite analysis under a range of conditions, both His-Rep1p and His-Cdc10p were retrieved as insoluble aggregates following release from the cell (Figures 5.1-5.2) and so biological activity assays using either of these two proteins were precluded.

7.2.4 Chapter 5: Biological activity assays *in vitro*

Following production of soluble His-Res1p and His-Res2p, these recombinant proteins were assayed for biological activity *in vitro* using the EMSA technique. In this study, neither His-Res1p nor His-Res2p (obtained in crude soluble *E. coli* extracts following overexpression) produced a detectable MCB DNA-binding activity when incubated with a radio-labelled MCB probe (Figure 5.7).

Failure to detect DNA-binding with either of these proteins may have been due, at least in part, to the absence of Cdc10p, since *in vitro* DNA-binding by either protein is dependent on Cdc10p, although weak Res2p DNA-binding has been reported alone (Ayte et al., 1995; Zhu et al., 1997). The insoluble nature of His-Cdc10p in this study prevented such analyses and so an alternative strategy employed co-expression studies in an attempt to circumvent this problem. Unfortunately, co-expression of either His-Res1p or His-Res2p with His-Cdc10p proved unsuccessful in terms of detectable protein. In addition, co-expression of either His-Res1p or His-Res2p with each other and with His-Rep1p or His-Rep2p was tested, reasoning that association of these proteins might facilitate MCB-specific DNA binding of the Resp subunits, in the absence of Cdc10p. Although co-expression of His-Res2p/His-Rep2p was detected, and both of these proteins were present in the soluble fraction (as determined by immunoblotting – Figure 5.8), use of this protein extract did not yield detectable MCB binding *in vitro*. Mixing of

either His-Res1p or His-Res2p with fission yeast protein extracts prepared from *Δres1* or *Δres2* cells, respectively, were similarly ineffectual. To summarise, neither bacterially produced soluble His-Res1p nor His-Res2p showed detectable DNA-binding activity *in vitro*, in the absence of Cdc10p.

Previous studies have indicated that in the absence of Cdc10p, at least *in vitro*, Res1p requires artificially mediated homodimerisation in order to exhibit MCB-specific DNA-binding (Ayte et al., 1995). It had been shown previously that bacterially expressed Res1p as a GST-fusion protein could bind specifically to an MCB DNA probe in *in vitro* band-shift assays (Ayte et al., 1995). In this study, both Res1p and Res2p were subsequently cloned and expressed as N-terminal GST-fusion proteins to investigate dimerisation-dependent MCB-specific DNA-binding *in vitro*, in comparison to the His-tagged isoforms. Hence, the cloning strategy was described, which allowed the successful and reproducible overexpression of the Res1p and Res2p components of the fission yeast DSC1 complex in *E. coli* as N-terminal GST-fusion proteins. Each protein was successfully overexpressed in soluble form following induction at 15°C (Figures 5.12-5.14). Furthermore, both GST-Res1p and GST-Res2p, in contrast to the His-tagged isoforms, bound specifically to MCB DNA in band-shift assays (Figures 5.16-5.17).

7.2.5 Chapter 6: Purification and analysis of the recombinant DSC1 components

Following the solubility studies and *in vitro* activity assays carried out in Chapter 5, purification of the recombinant DSC1 components and initiation of structural analyses were described in Chapter 6. The solubilisation and purification of His-Rep1p and His-Cdc10p from inclusion bodies was also described. The aims of this chapter were twofold: firstly, to obtain pure protein of sufficient yield and purity with which to carry out further *in vitro* functional analyses and secondly, to allow the initiation of structural studies.

Unfortunately, both yield and purity of His-Res1p and His-Rep2p were unsatisfactory, following purification by IMAC (Figure 6.2). This was attributed to the low-levels of these proteins produced following overexpression in *E. coli*. Despite attempts to increase expression levels and scale-up the purification, the amounts retrieved following purification were insufficient to pursue structural studies.

In contrast, the yields of GST-Res1p and GST-Res2p obtained following overexpression in *E. coli* were significantly greater, and it was anticipated that purification would provide sufficient amounts of material with which to instigate structure/function studies. Frustratingly, however, neither GST-Res1p nor GST-Res2p was retrieved in sufficient yield or purity, following affinity purification (Figure 6.13). Due to time constraints, optimisation of the purification procedure for these proteins has not been fully investigated and as such, the results presented represent preliminary data.

His-Res2p was reproducibly expressed and purified in high yield (5-10 mg l^{-1} culture). Initial observations noted the continued presence of a 'contaminant' protein species of M_r approximately 60 kDa, although this was subsequently identified as a N-terminally degraded product of His-Res2p, by mass spectrometric analysis (Figures 6.4-6.11). Unfortunately, recovery of purified His-Res2p, in buffer solutions suitable for structural studies, frequently resulted in precipitation. Initial attempts to obtain CD spectral data from this protein were unsuccessful, primarily due to precipitation. The spectral studies showed a high level of nucleic acid contamination in purified samples of His-Res2p. This was despite the fact that routine measures (e.g. use of both DNAase and RNAase) were employed to prevent this problem.

Both His-Rep1p and His-Cdc10p proteins were refractory to manipulation of induction and solvent conditions, although each protein was ultimately solubilised and purified from inclusion bodies, using the anionic detergent N-lauroylsarcosine in yields that were sufficient to initiate structural studies (Figures 6.14-6.15). Similarly to His-Res2p, however, dialysis into more suitable buffers routinely resulted in precipitation and aggregation of protein. Initial CD spectra of these proteins also revealed a significant level of nucleic acid contamination (Figure 6.16). Nevertheless, CD studies were initiated on His-Cdc10p and His-Rep1p, as this technique is tolerant to low levels of detergent. The results of these CD studies demonstrated that each protein had significant secondary structure and were at least partially folded (Figures 6.17-6.20 and Tables 6.1-6.2). Given that both proteins required detergent solubilisation, it is highly likely that to obtain these proteins in their fully native state, re-folding strategies are required.

7.3 Future experimental work

The aim of this work was to provide a stable reproducible and active source of recombinant components of the fission yeast DSC1 complex in sufficient yield and purity to allow detailed structural and functional analyses. Whilst significant steps have been taken toward this ultimate goal, several key issues require further investigation.

Importantly, this study has demonstrated that bacterially expressed Res1p and Res2p are biologically active, in terms of MCB-specific DNA-binding activity, when expressed as N-terminal GST-fusion proteins. This indicates that post-translational modification is not essential for the DNA-binding activity of these proteins *in vitro*. This had been demonstrated previously for GST-Res1p (Ayte et al., 1995). Initial attempts to purify these proteins in sufficient yield for further analyses, specifically structural studies, proved unsuccessful. Future work would therefore concentrate on optimising the purification of these two proteins to provide a convenient source of pure protein with which to undertake detailed biophysical analyses.

In contrast, no such DNA-binding activity was detected with either of the His-tagged Resp proteins. Previous studies reported that the MCB-specific DNA-binding activity of *in*

vitro translated Res1p and Res2p was dependent on co-expression with Cdc10p, although a weak band-shift activity could be detected with Res2p alone (Ayte et al., 1995; Zhu et al., 1997). In this thesis, the inability to obtain soluble His-Cdc10p protein (or to co-express either His-Res1p or His-Res2p with His-Cdc10p) has prevented a definitive analysis of the dependency of these proteins upon Cdc10p for *in vitro* DNA-binding activity. Even so, this study has demonstrated the solubilisation of His-Cdc10p from inclusion bodies, although it is likely that the production of active His-Cdc10p protein from this source will require re-folding. If this can be achieved, then the issue of His-Res1p and His-Res2p DNA-binding activity may be more fully investigated.

Of particular interest is why the GST-Resp proteins should exhibit MCB-specific DNA-binding activity, yet the His-Resp proteins do not. Clearly such differences may reflect the different N-terminal tags. It is possible that this may be due to a tag-mediated protein folding effect with the His-Resp proteins, presumably, unable to fold into their native tertiary structure, in contrast to their GST-tagged isoforms. However, the soluble nature of both His-Res1p and His-Res2p suggests this would be unlikely. It has been demonstrated in this study that His-tagged Res1p and Res2p are able to function *in vivo*. Perhaps a more obvious explanation for the difference in DNA-binding activity between these fusion proteins lies in the intrinsic dimerisation potential of the GST-tag. In contrast to the His-tagged proteins, GST-fusion proteins are expressed as homodimers.

As discussed in Chapter 5, Res1p and Res2p are believed to bind to DNA as heterodimers *in vivo* and Res1p and Res2p can homo-dimerise *in vitro* in a Cdc10p-dependent manner (Ayte et al., 1997; Zhu et al., 1997; Whitehall et al., 1999). Thus, it is highly likely that the difference in activity displayed by the GST-Resp and His-Resp proteins reflects intrinsic dimerisation potential. However, this conclusion cannot be reached solely from the results presented in this study, and require further investigation to be proved.

The GST-Resp fusion proteins used in band-shift assays herein were obtained from crude soluble bacterial extracts (Figure 5.16). Whilst the majority of protein in these extracts was full-length GST-Res1p or GST-Res2p, they were not homogeneous and so it could not be determined whether DNA-binding was mediated by full-length, or C-terminally truncated GST-Resp molecules. Importantly, C-terminally truncated Res2p molecules show enhanced DNA-binding *in vitro*, and the highly related ScSwi4p protein displays C-terminally mediated auto-inhibition of DNA-binding (Zhu et al., 1997; Baetz and Andrews, 1999). Thus, it is tempting to speculate that the inability of the His-Resp proteins reflects an inability to homo-dimerise and/or C-terminally mediated inhibition, in the absence of Cdc10p.

In order to further resolve these issues, pure preparations of full-length GST-Resp fusion proteins must be obtained for use in DNA-binding assays. Initial attempts to do this were described in Chapter 6. Due to time constraints the conditions for purification have not been

optimised. Thrombin cleavage of each tag could be employed in order to separate the effects of the respective tags upon DNA-binding activity. Of particular importance will be assay of the ability of each protein to bind DNA in the absence of GST-mediated dimerisation.

In this respect, the His-Resp fusion proteins merit further analysis. Full-length His-Res2p has been successfully over-expressed and purified in high yield, providing a valuable resource for structural studies. Importantly, to date, this represents the first member of this family of transcription factors to have been produced intact and in yields sufficient to permit detailed biophysical analyses. As discussed above (and in Chapter 5) the apparent inactivity of either His-Resp protein may not simply be due to mis-folded but rather due to an inability to homo-dimerise *in vitro*. So, whilst a reproducible source of pure His-Res2p in high yield has been produced, final optimisation of the stability of this protein requires further investigation. Specifically, protein precipitation has proven to be a frequent and frustrating problem following attempts to study this purified recombinant protein. Assuming that the problem of precipitation can be overcome, structural studies may then be initiated. Similarly, following optimisation of the GST-Resp purifications, structural analyses should be possible. Comparative studies with His-Res2p, for example using CD, will be of particular interest. In addition, the purified GST-Res1p and GST-Res2p may also be utilised for pull-down assays with the His-tagged DSC1 components.

In future studies it may be worthwhile to investigate the use of alternative heterologous hosts for the expression of the recombinant DSC1 component proteins. Eukaryotic hosts such as the methylotrophic yeast *Pichia pastoris* and insect cells are now more commonly in use. In particular, exploitation of these systems may be required to produce sufficient amounts of His-Res1p and His-Rep2p, which have been unattainable in *E. coli*. Ultimately following optimisation and extension of much of the work presented in this study, the long-term goal of reconstituting an *in vitro* DSC1 complex should be possible providing an invaluable resource for future investigation into the structure and function of these components.

References

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D. (1994). *Molecular Biology of the Cell* (3rd Edition). Garland publishing.

Amon, A., Tyers, M., Futcher, B., and Nasmyth, K. (1993). Mechanisms That Help the Yeast Cell Cycle Clock Tick: G2 Cyclins Transcriptionally Activate G2 Cyclins and Repress G1 Cyclins. *Cell* **74**, 993-1007.

Andrews, B.J., and Herskowitz, I. (1989a). Identification of a DNA Binding Factor Involved in Cell-Cycle Control of the Yeast *HO* Gene. *Cell* **57**, 21-29.

Andrews, B.J., and Herskowitz, I. (1989b). The yeast SWI4 protein contains a motif present in developmental regulators and is part of a complex involved in cell-cycle dependent transcription. *Nature* **342**, 830-833.

Andrews, B. J., and Moore, L. A. (1992). Interaction of the yeast Swi4 and Swi6 cell cycle regulatory proteins *in vitro*. *Proceedings of the National Academy of Sciences USA* **89**, 11852-11856.

Aves, S. J., Durkacz, B. W., Carr, A., and Nurse, P. (1985). Cloning, sequencing and transcriptional control of the *Schizosaccharomyces pombe cdc10* 'start' gene. *The EMBO Journal* **4**, 457-463.

Ayte, J., Leis, J. F., and DeCaprio, J. A. (1997). The Fission Yeast Protein $p73^{res2}$ Is an Essential Component of the Mitotic MBF Complex and a Master Regulator of Meiosis. *Molecular and Cellular Biology* **17**, 6246-6254.

Ayte, J., Leis, J. F., Herrera, A., Tang, E., Yang, H., and DeCaprio, J. A. (1995). The *Schizosaccharomyces pombe* MBF Complex Requires Heterodimerization for entry into S Phase. *Molecular and Cellular Biology* **15**, 2589-2599.

Ayte, J., Schweitzer, C., Zarzov, P., Nurse, P., and DeCaprio, J. A. (2001). Feedback regulation of the MBF transcription factor by cyclin Cig2. *Nature Cell Biology* **3**, 1043-1050.

Baber-Furnari, B. A., Rhind, N., Boddy, M. N., Shanahan, P., Lopez-Girona, A., and Russell, P. (2000). Regulation of Mitotic Inhibitor Mik1 Helps to Enforce the DNA Damage Checkpoint. *Molecular Biology of the Cell* **11**, 1-11.

Baetz, K., and Andrews, B. (1999). Regulation of Cell Cycle Transcription Factor Swi4 through Auto-Inhibition of DNA-Binding. *Molecular and Cellular Biology* **19**, 6729-6741.

Bardin, A. J., and Amon, A. (2001). MEN and SIN: What's the Difference? *Nature Reviews Molecular Cell Biology* **2**, 815-826.

Breeden, L. L. (2003). Periodic Transcription: A Cycle Within a Cycle. *Current Biology* **13**, R31-R38.

Bartlett. R., and Nurse, P. (1990). Yeast as a Model System for Understanding the Control of DNA Replication in Eukaryotes. *Bioessays* **12**, 457-463.

Baum, B., Nishitani, H., Yanow, S., and Nurse, P. (1998). Cdc18 transcription and proteolysis couple S phase to passage through mitosis. *The EMBO Journal* **17**, 5689-5698.

Baum, B., Wuarin, J., and Nurse, P. (1997). Control of S-phase periodic transcription in the fission yeast mitotic cycle. *The EMBO Journal* **16**, 4676-4688.

Bell, S. P., and Stillman, B. (1992). ATP-dependent recognition of eukaryotic origins of DNA replication by a multiprotein complex. *Nature* **357**, 128-134.

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

Birkenbihl, R. P., and Subramani, S. (1995). The *rad21* Gene Product of *Schizosaccharomyces pombe* Is a Nuclear, Cell Cycle-regulated Phosphoprotein. *The Journal of Biological Chemistry* **270**, 7703-7711.

Blagosklonny, M. V., and Pardee, A. B. (2002). The Restriction Point of the Cell Cycle. *Cell Cycle* **1**, 103-110.

Blanco, M. A., Sanchez-Diaz, A., De Prada, J. M., and Moreno, S. (2000). $APC^{ste9/srw1}$ promotes degradation of mitotic cyclins in G1 and is inhibited by cdc2 phosphorylation. *The EMBO Journal* **19**, 3945-2955.

Boddy, M. N., and Russell, P. (2001). DNA replication checkpoint. *Current Biology* **11**, R953-R956.

Booher, R. N., Alfa, C. E., Hyams, J. S., and Beach, D. H. (1989). The Fission Yeast cdc2/cdc13/suc1 Protein Kinase: Regulation of Catalytic Activity and nuclear Localization. *Cell* **58**, 485-497.

Bork, P. (1993). Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally? *Proteins* **17**, 363-374.

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry* **72**, 248-254.

Breeden, L., and Nasmyth, K. (1987 a). Cell Cycle Control of the Yeast *HO* Gene: *Cis*- and *Trans*-Acting Regulators. *Cell* **48**, 389-397.

Breeden, L., and Nasmyth, K. (1987 b). Similarity between cell-cycle genes of budding yeast and fission yeast and the *Notch* gene of *Drosophila*. *Nature* **329**, 651-654.

Brown, G. W., and Kelly, T. J. (1998). Purification of Hsk1, a Minichromosome Maintenance Protein Kinase from Fission yeast. *The Journal of Biological Chemistry* **273**, 22083-22090.

Brown, G. W., and Kelly, T. J. (1999). Cell cycle regulation of Dfp1, an activator of the Hsk1 protein kinase. *Proceedings of the National Academy of Sciences USA* **96**, 8443-8448.

Caligiuri, M., and Beach, D. (1993). Sct1 Functions in Partnership with Cdc10 in a Transcription Complex That Activates Cell Cycle START and Inhibits Differentiation. *Cell* **72**, 607-619.

Carr, A. M., and Caspari, T. (2003). Checkpoint Controls Halting the Cell Cycle. *In* : Egel, R. ed. The Molecular Biology of *Schizosaccharomyces pombe*: Genetics, Genomics and Beyond. Springer, 41-54.

Carrio, M., M., and Villaverde, A. (2002). Construction and deconstruction of bacterial inclusion bodies. *Journal of Biotechnology* **96**, 3-12.

Chaung, R. -Y., Chretien, L., Dai, J., and Kelly, T. J. (2002). Purification and Characterization of the Schizosaccharomyces pombe Origin Recognition Complex. *The Journal of Biological* **277**, 16920-16927.

Chaung, R. -Y., and Kelly, T. J. (1999). The fission yeast homologue of Orc4p binds to replication origin DNA via multiple AT-hooks. *Proceedings of the National Academy of Sciences USA* **96**, 2656-2661.

Cho, R. J., Huang, M., Campbell, M. J., Dong, H., Steinmetz, L., Sapinoso, L., Hampton, G., Elledge, S. J., Davis, R. W., and Lockhart, D. J. (2001). Transcriptional regulation and function during the human cell cycle. *Nature Genetics* **27**, 48-54.

Christensen, P. U., Bentley, N. J., Martinho, R. G., Nielsen, O., and Carr, A. (2000). Mik1 levels accumulate in S phase and may mediate an intrinsic link between S phase and mitosis. *Proceedings of the National Academy of Sciences USA* **97**, 2579-2584.

Clarke, K. L., Halay, E. D., Lai, E., and Burley, S. K. (1993). Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* **364**, 412-420.

Connolly, T., and Beach, D. (1994). Interaction between the Cig1 and Cig2 B-Type Cyclins in the Fission Yeast Cell Cycle. *Molecular and Cellular Biology* **14**, 768-776.

Correa-Bordes, J., Gulli, M. -P., and Nurse, P. (1997). $p25^{rum1}$ promotes proteolysis of the mitotic B-cyclin $p65^{cdc13}$ during G1 of the fission yeast cell cycle. *The EMBO Journal* **16**, 4657-4664.

Cosma, M. P., Tanaka, T., and Nasmyth, K. (1999). Ordered Recruitment of Transcription and Chromatin Remodelling Factors to a Cell Cycle and Developmentally Regulated Promoter. *Cell* **97**, 299-311.

Costanzo, M., Schub, O., and Andrews, B. (2003). G_1 Transcription Factors Are Differentially Regulated in Saccharomyces cerevisiae by the Swi6-Binding protein Stb1. *Molecular and Cellular Biology* **23**, 5064-5077.

Cunliffe, L., White, S., and McInerny, C. J. (2004). DSC1-MCB regulation of meiotic transcription in *Schizosaccharomyces pombe*. *Molecular Genetics and Genomics* **271**, 60-71.

Daga, R. R., and Jimenez, J. (1999). Translational control of the Cdc25 cell cycle phosphatase: a molecular mechanism coupling mitosis to cell growth. *Journal of Cell Science* **112**, 3137-3146.

DeBruin, A., Maiti, B., Jakoi, L., Timmers, C., Buerki, R., and Leone, G. (2003). Identification and Characterization of *E2F7*, a Novel Mammalian E2F Family Member Capable of Blocking Cellular Proliferation. *The Journal of Biological Chemistry* **278**, 42041-42049.

De Bruin, R. A. M., McDonald, W. H., Kalashnikova, T. I., Yates III, J., and Wittenberg, C. (2004). Cln3 Activates G1-Specific Transcription via Phosphorylation of the SBF Bound Repressor Whi5. *Cell* **117**, 887-898.

Diffley, J. F. X., and Labib, K. (2002). The chromosome replication cycle. *Journal of Cell Science* **115**, 869-872.

Ding, R., and Smith, G. R. (1998). Global control of meiotic recombination genes by *Schizosaccharomyces pombe rec16* (*rep1*). *Molecular Genetics and Genomics* **258**, 663-670.

Dirick, L., Moll, T., Auer, H. and Nasmyth, K. (1992). A central role for *SWI6* in modulating cell cycle Start-specific transcription in yeast. *Nature* **357**, 508-513.

Dyson, N. (1998). The regulation of E2F by pRB-family proteins. *Genes and Development* **12**, 2245-2262.

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

Espinoza, F. H., Ogas, J., Herskowitz, I., and Morgan, D. O. (1994). Cell Cycle Control by a Complex of the Cyclin HCS26 (PCL1) and the Kinase PHO85. *Science* **266**, 1388-1391.

Fernandez-Sarabia, M. J., McInerny, C. J., Harris, P., Gordon, C., and Fantes, P. A. (1993). The cell cycle genes *cdc22*$^{+}$ and *suc22*$^{+}$ of the fission yeast *Schizosaccharomyces pombe* encode the large and small subunits of ribonucleotide reductase. *Molecular Genetics and Genomics* **238**, 241-251.

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

Foord, R., Taylor, I. A., Sedgwick, S. G., and Smerdon, S. J. (1999). X-ray structural analysis of the yeast cell cycle regulator Swi6 reveals variations of the ankyrin fold and has implications for Swi6 function. *Nature Structural Biology* **6**, 157-165.

Ford, H. L., and Pardee, A. B. (1999). Cancer and the Cell Cycle. *Journal of Cellular Biochemistry Supplements* **32/33**, 166-172.

Forsburg, S. L. (2004). Eukaryotic MCM Proteins: Beyond Replication Initiation. *Microbiology and Molecular Biology Reviews* **68**, 109-131.

Forsburg, S. (1993). Comparison of *Schizosaccharomyces pombe* expression systems. *Nucleic Acids Research* **21**, 2955-2956.

Forsburg, S. L., and Nurse, P. (1991). Cell cycle regulation in the yeasts *Saccharomyces cerevisiae and Schizosaccharomyces pombe*. *Annual Reviews of Cell Biology* **7**, 227-56.

Frangioni, J., V., and Neel, B., G. (1993). Solubilization and Purification of Enzymatically Active Glutathione *S*-Transferase (pGEX) Fusion Proteins. *Analytical Biochemistry* **210**, 179-187.

Gaymonet, M., Spanos, A., Wells, G. P., Smerdon, S. J., and Sedgwick, S. G. (2004). Clb6/Cdc28 and Cdc14 Regulate Phosphorylation Status and Cellular Localization of Swi6. *Molecular and Cellular Biology* **24**, 2277-2285.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H., and Oliver, S. G. (1996). Life with 6000 genes. *Science* **274**, 546-67.

Gordon, C., and Fantes, P. A. (1986). The *cdc22*$^{+}$ gene of Schizosaccharomyces *pombe* encodes a cell cycle-regulated transcript. *The EMBO Journal* **5**, 2981-2985.

Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999). ESPript: multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305-308.

Gregan, J., Lindner, K., Brimage, L., Franklin, R., Namdar, M., Hart, E. H., Aves, S., and Kearsey, S. (2003). Fission Yeast Cdc23/Mcm10 Functions After Pre-replicative Complex Formation to Promote Cdc45 Chromatin Binding. *Molecular Biology of the Cell* **14**, 3876-3887.

Guan, K. –L., and Dixon, J. E. (1991). Eukaryotic Proteins Expressed in *Escherichia coli*: An Improved Thrombin Cleavage and Purification procedure of Fusion Proteins with Glutathione *S*-Transferase. *Analytical Biochemistry* **192**, 262-267.

Hagstrom, K. A., and Meyer, B. J. (2003). Condensin and Cohesin: More than Chromosome Compactor and Glue. *Nature Reviews Genetics* **4**, 520-534.

Hartwell, L. H. (1974). *Saccharomyces cerevisiae* Cell Cycle. *Bacteriological Reviews* **38**, 164-198.

Hartwell, L. H., Culotti, J., Pringle, J. R., and Reid, B. J. (1974). Genetic control of the cell division cycle in yeast. *Science* **183**, 46-51.

Hartwell, L. H. (1991). Twenty-Five Years of Cell Cycle Genetics. *Genetics* **129**, 975-980.

Harper, J. W., Burton, J. L., and Solomon, M. J. (2002). The anaphase-promoting complex: it's not just for mitosis any more. *Genes and Development* **16**, 2179-2206.

Harrington, L. A., and Andrews, B. J. (1996). Binding to the yeast Swi4,6-dependent cell cycle box, CACGAAA, is cell cycle regulated *in vivo*. *Nucleic Acids Research* **24**, 558-565.

Harris, P., Kersey, P. J., McInerny, C. J., and Fantes, P. A. (1996). Cell cycle, DNA damage and heat shock regulate suc22+ expression in fission yeast. *Molecular Genetics and Genomics* **252**, 284-291.

Hartwell, L. H., and Weinert, T. A. (1989). Checkpoints: Controls That Ensure the Order of Cell Cycle Events. *Science* **246**, 629-634.

Hofmann, J. F. X., and Beach, D. (1994). *cdt1* is an essential target of the Cdc10/Sct1 transcription factor: requirement for DNA replication and inhibition of mitosis. *The EMBO Journal* **13**, 425-434.

Humphrey, T. (2000). DNA damage and cell cycle control in *Schizosaccharomyces pombe*. *Mutation Research* **451**, 211-226.

Iyer, V. R., Horak, C. E., Scafe, C. S., Botstein, D., Snyder, M., and Brown, P. O. (2001). Genomic binding sites of the yeast cell-cycle transcription factors SBF and MBF. *Nature* **409**, 533-538.

Jallepalli, P. V., Brown, G. W., Muzi-Falconi, M., Tien, D., and Kelly, T. J. (1997). Regulation of the replication initiator protein $p65^{cdc18}$ by CDK phosphorylation. *Genes and Development* **11**, 2767-2779.

Johnston, L. H., and Lowndes, N. F. (1992). Cell cycle control of DNA synthesis in budding yeast. *Nucleic AcidsResearch* **20**, 2403-2410.

Kellogg, D. R. (2003). Wee1-dependent mechanisms required for coordination of cell growth and cell division. *Journal of Cell Science* **116**, 4883-4890.

Kelly, S. M., and Price, N. C. (2000). The Use of Circular Dichroism in the Investigation of Protein Structure and Function. *Current Protein and Peptide Science* **1**, 349-384.

Kelly, T. J., and Brown, G. W. (2000). Regulation of Chromosome Replication. *Annual Reviews of Biochemistry* **69**, 829-880.

Kelly, T. J., Martin, G. S., Forsburg, S. L., Stephen, R. J., Russo, A., and Nurse, P. (1993). The Fission Yeast *cdc18*$^{+}$ Gene Product Couples S Phase to START and Mitosis. *Cell* **74**, 1-20.

Koch, C., Moll, T., Neuberg, M., Ahorn, H. and Nasmyth, K. (1993). A Role for the Transcription Factors Mbp1 and Swi4 in Progression from G1 to S Phase. *Science* **261**, 1551-1557.

Koch, C., and Nasmyth, K. (1994). Cell cycle regulated transcription in yeast. Current Opinion in Cell Biology **6**, 451-459.

Koch, C., Schleiffer, A., Ammerer, G., and Nasmyth, K. (1996). Switching transcription on and off during the yeast cell cycle: Cln/Cdc28 kinases activate bound transcription factor SBF (Swi4/Swi6) at Start, whereas Clb/Cdc28 kinases displace it from the promoter in G_2. *Genes and Development* **10**, 129-141.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **277**, 680-685.

Lee, B., and Amon, A. (2001). Meiosis: how to create a specialized cell cycle. *Current Opinion in Cell Biology* **13**, 770-777.

Lee, J. -K., and Hurwitz, J. (2001). Processive DNA helicase activity of the minichromosome maintenance proteins 4, 6 and 7 complex requires forked DNA structures. *Proceedings of the National Academy of Sciences USA* **98**, 54-59.

Lee, J. -K., Seo, Y. -S., and Hurwitz, J. (2003). The Cdc23 (Mcm10) protein is required for the phosphorylation of minichromosome maintenance complex by the Dfp1-Hsk1 kinase. *Proceedings of the National Academy of Sciences USA* **100**, 2334-2339.

Lei, M., and Tye, B. K. (2001). Initiating DNA synthesis: from recruiting to activating the MCM complex. *Journal of Cell Science* **114**, 1447-1454.

Levine, M., and Tjian, R. (2003). Transcription regulation and animal diversity. *Nature* **424**, 147-151.

Lew, D. J., Weinert, T., and Pringle, J. R. (1997). Cell Cycle Control in *Saccharomyces cerevisiae*. In The Molecular and Cellular Biology of the Yeast *Saccharomyces*: Cell Cycle and Cell Biology, Pringle, J. R., Broach, J. R. and Jones, E. W., eds. : Cold Spring Harbor Laboratory Press, 607-695.

Lilie, H., Schwarz, E., and Rudolph, R. (1998). Advances in refolding of proteins produced in *E. coli*. *Current Opinion in Biotechnology* **9**, 497-501.

Lindner, P., Guth, B., Wülfing, C., Krebber, C., Steipe, B., Müller, F., and Plückthum, A. (1992). Purification of Native Proteins from the Cytoplasm and Periplasm of Escherichia coli Using IMAC and Histidine Tails: A Comparison of Proteins and Protocols. *METHODS: A Companion to Methods in Enzymology* **4**, 41-56.

Lodish, H., Baltimore, D., Berk, A., Zipursky, S. L., Matsudaira, P., and Darnell, J. (1995). *Molecular Cell Biology* (3rd edition). Scientific American Books.

Logan , N., Delavaine, L., Graham, A., Reilly, C., Wilson, J., Brummelkamp, T. R., Hijmans, E. M., Bernards, R., and LaThangue, N. B. (2004). E2F-7: a distinctive E2F family member with an unusual organization of DNA-binding domains. *Oncogene* **23**, 5138-5150.

Lopez-Girona, A., Mondesert, O., Leatherwood, J., and Russell, P. (1998). Negative Regulation of Cdc18 DNA Replication Protein by Cdc2. *Molecular Biology of the Cell* **9**, 63-73.

Lowndes, N. F., Johnston, A. L. and Johnston, L.H. (1991). Coordination of expression of DNA synthesis genes in budding yeast by a cell-cycle regulated *trans* factor. *Nature* **350**, 247-250.

Lowndes, N., McInerny, C. J., Johnston, A. L., Fantes, P. A., and Johnston, L.H. (1992a). Control of DNA synthesis genes in fission yeast by the cell-cycle gene *cdc10*$^{+}$. *Nature* **355**, 449-453.

Lowndes, N., Johnston, A. L., Breeden, L., and Johnston, L.H. (1992b). *SWI6* protein is required for transcription of the periodically expressed DNA synthesis genes in budding yeast. *Nature* **357**, 505-508.

Lu, Q., Bauer, J. C., and Greener, A. et al. (1997). Using *Schizosaccharomyces pombe* as a host for expression and purification of eukaryotic proteins. *Gene* **200**, 135-144.

Lygerou, Z., and Nurse, P. (1999). The fission yeast origin recognition complex is constitutively associated with chromatin and is differentially modified through the cell cycle. *Journal of Cell Science* **112**, 3703-3712.

Lygerou, Z., and Nurse, P. (2000). Controlling S-phase Onset in Fission Yeast. *Cold Spring Harbor Symposia on Quantitative Biology* **65**, 332-332.

MacKay, V. L., Mai, B., Waters, L., and Breeden, L. L. (2001). Early Cell Cycle Box-Mediated Transcription of *CLN3* and *SWI4* Contributes to the Proper Timing of the G_1-to-S Transition in Budding Yeast. *Molecular and Cellular Biology* **21**, 4140-4148.

MacNeill, S., and Nurse, P. (1997). Cell Cycle Control in Fission Yeast. *In* : Pringle, J. R., Broach, J. R. and Jones, E. W., eds. The Molecular and Cellular Biology of the Yeast *Saccharomyces*: Cell Cycle and Cell Biology. Cold Spring Harbor Laboratory Press, 697-763.

Maqbool, Z., Kersey, P. J., Fantes, P. A., and McInerny, C. J. (2003). MCB-mediated regulation of cell cycle-specific *cdc22*$^{+}$ transcription in fission yeast. *Molecular Genetics and Genomics* **269**, 765-775.

Marahens, Y., and Stillman, B. (1992). A Yeast Chromosomal Origin of DNA Replication Defined by Multiple Functional Elements. *Science* **255**, 817-823.

Marks, J., Fankhauser, C., Reymond, A., and Simanis, V. (1992). Cytoskeletal and DNA structure abnormalities result from bypass of requirement for the *cdc10* start gene in the fission yeast *Schizosaccharomyces pombe*. *Journal of Cell Science* **101**, 517-528.

Martin-Castellanos, C., Blanco, M. A., De Prada, J. M., and Moreno, S. (2000). The puc1 Cyclin Regulates the G1 Phase of the Fission Yeast Cell Cycle in Response to Cell Size. *Molecular Biology of the Cell* **11**, 543-544.

Martin-Castellanos, C., Labib, K., and Moreno, S. (1996). B-type cyclins regulate G1 progression in fission yeast in opposition to the p25^{rum1} cdk inhibitor. *The EMBO Journal* **15**, 839-849.

Masukata, H., Huberman, J. A., Frattini, M. G., and Kelly, T. J. (2003). DNA Replication in *S. pombe*. *In* : Egel, R. ed. The Molecular Biology of *Schizosaccharomyces pombe*: Genetics, Genomics and Beyond. Springer, 73-94.

Mata, J., Lyne, R., Burns, G., and Bähler, J. (2002). The transcriptional program of meiosis and sporulation in fission yeast. *Nature Genetics* **32**, 143-147.

Maundrell, K. (1993). Thiamine-repressible expression vectors pREP and pRIP for fission yeast. *Gene* **123**, 127-130.

McCollum, D., and Gould, K. L. (2001). Timing is everything: regulation of mitotic exit and cytokinesis by the MEN and SIN. *Trends in Cell Biology* **11**, 89-95.

McInerny, C. J. (2004). Cell cycle-regulated transcription in fission yeast. *Biochemical Society Transactions*. In Press.

McInerny, C. J., Kersey, P. J., Creanor, J., and Fantes, P. A. (1995). Positive and negative roles for cdc10 in cell cycle gene expression. *Nucleic Acids Research* **23**, 4761-4768.

McInerny, C. J., Partridge, J. F., Mikesell, G. E., Creemer, D. P., and Breeden, L. L. (1997). A novel Mcm1-dependent element in the *SWI4*, *CLN3*, *CDC6*, and *CDC47* promoters activates M/G_1-specific transcription. *Genes and Development* **11**, 1277-1288.

McIntosh, E. M., Atkinson, T., Storms, R. K. and Smith, M. (1991). Characterisation of a short *cis*-acting DNA sequence which conveys cell-cycle stage dependent transcription in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **11**, 329-337.

McIntosh, P. B., Taylor, I. A., Frenkiel, T. A., Smerdon, S. J., and Lane, A. (2000). The influence of DNA binding on the backbone dynamics of the yeast cell-cycle protein Mbp1. *Journal of Biomolecular NMR* **16**, 183-196.

Measday, V., Moore, L., Ogas, J., Tyers, M., and Andrews, B. (1994). The PCL2 (ORFD)-PHO85 cyclin-dependent kinase complex: a cell cycle regulator in yeast. *Science* **266**, 1391-1391.

Merrill, G. F., Morgan, B. A., Lowndes, N. F., and Johnson, L. H. (1992). DNA Synthesis Control in Yeast: An Evolutionarily Conserved Mechanism for Regulating DNA Synthesis Genes? *BioEssays* **14**, 823-830.

Miyamoto, M., Tanaka, K., and Okayama, H. (1994). $res2^+$, a new member of the *cdc10*$^+$/*SWI4* family, controls the 'start' of mitotic and meiotic cycles in fission yeast. *The EMBO Journal* **13**, 1873-1880.

Moll, T., Dirick, L., Auer, H., Bonkovsky, J. and Nasmyth, K. (1992). SWI6 is a regulatory subunit of two different cell cycle START-dependent transcription factors in *Saccharomyces cerevisiae*. *Journal of Cell Science* **16**, 87-96.

Mondesert, O., McGowan, C. H., and Russell, P. (1996). Cig2, a B-Type Cyclin, Promotes the Onset of S Phase in *Schizosaccharomyces pombe*. *Molecular and Cellular Biology* **16**, 1527-1533.

Moon, K. -Y., Kong, D., Lee, J. -K., Raychaudhuri, S., and Hurwitz, J. (1999). Identification and reconstitution of the origin recognition complex from *Schizosaccharomyces pombe*. *Proceedings of the National Academy of Sciences USA* **96**, 12367-12372.

Moreno, S., Klar, A., Nurse, P. (1991). Molecular genetic analysis of the fission yeast *Schizossacharomyces pombe*. *Methods in Enzymology* **194**, 795-823.

Moreno, S., Nurse, P., and Russell, P. (1989). Regulation of mitosis by cyclic accumulation of p80 ^{cdc25} mitotic inducer in fission yeast. *Nature* **344**, 549-552.

Morgan, D. O. (1999). Regulation of the APC and the exit from mitosis. *Nature Cell Biology* **1**, E47-E53.

Moser, B. A., and Russell, P. (2000). Cell cycle regulation in *Schizosaccharomyces pombe*. *Current Opinion in Microbiology* **3**, 631-636.

Nair, M., McIntosh. P. B., Frenkiel, T. A., Kelly, G., Taylor, I. A., Smerdon, S. J., and Lane, A. N. (2003). NMR Structure of the DNA-Binding Domain of the Cell Cycle Protein Mbp1 from *Saccharomyces cerevisiae*. *Biochemistry* **42**, 1266-1273.

Nakashima, N., Tanaka, K., Sturm, S., and Okayama, H. (1995). Fission yeast Rep2 is a putative transcriptional activator subunit for the cell cycle 'start' function of Res2-Cdc10. *The EMBO Journal* **14**, 4794-4802.

Nasmyth, K. (1983). Molecular analysis of a cell lineage. *Nature* **302**, 670-676.

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

Nasmyth, K., Peters, J. -M., and Uhlmann, F. (2000). Splitting the Chromosome: Cutting the Ties That Bind Sister Chromatids. *Science* **288**, 1379-1384.

Ng, S. S., Anderson, M., White, S., and McInerny, C. J. (2001). *mik1*$^{+}$ G1-S transcription regulates mitotic entry in fission yeast. *FEBS Letters* **503**, 131-134.

Nigg, E. A. (2001). Mitotic Kinases as regulators of Cell Division and its Checkpoints. *Nature Reviews Molecular Cell Biology* **2**, 21-32.

Nishitani, H., Lygerou, Z., Nishimoto, T., and Nurse, P. (2000). The Cdt1 protein is required to license DNA for replication in fission yeast. *Nature* **404**, 625-628.

Nurse, P. (1975). Genetic control of cell size at cell division in yeast. *Nature* **256**, 547-551.

Nurse, P. (1990). Universal control mechanism regulating onset of M-phase. *Nature* **344**, 503-508.

Nurse, P. (2000). A Long Twentieth Century of the Cell Cycle and Beyond. *Cell* **100**, 71-78.

Nurse, P., and Bissett, Y. (1981). Gene required in G1 for commitment to cell cycle and in G2 for control of mitosis in fission yeast. *Nature* **292**, 558-560.

Nurse, P., Masui, Y., and Hartwell, L. (1998). Understanding the cell cycle. *Nature Medicine* **4**, 1103-1106.

Nurse, P., Thuriaux, P., and Nasmyth, K. (1976). Genetic Control of the Cell Division Cycle in the Fission Yeast *Schizosaccharomyces pombe*. *Molecular and General Genetics* **146**, 167-178.

Obara-Ishihara, T., and Okayama, H. (1994). A B-type cyclin negatively regulates conjugation via interacting with cell cycle 'start' genes in fission yeast. *The EMBO Journal* **13**, 1863-1872.

Ogas, J., Andrews, B. J., and Herskowitz, I. (1991). Transcriptional Activation of *CLN1*, *CLN2*, and a Putative New G1 Cyclin (*HCS26*) by SWI4, a Positive Regulator of G1-Specific Transcription. *Cell* **66**, 1015-1026.

Okuno, Y., Satoh, H., Sekiguchi, M., and Masukata, H. (1999). Clustered Adenine/Thymine Stretches Are Essential for Function of a Fission Yeast Replication Origin. *Molecular and Cellular Biology* **19**, 6699-6709.

Pardee, A. B. (1974). A Restriction Point for Control of Normal Animal Cell Proliferation. *Proceedings of the National Academy of Sciences USA* **71**, 1286-1290.

Parker, A. E., Clyne, R., Carr, A., and Kelly, T. (1997). The *Schizosaccharomyces pombe rad11*$^{+}$ Gene Encodes the Large Subunit of Replication Protein A. *Molecular Cell Biology* **17**,

2381–2390.

Partridge, J. F., Mikesell, G. E., and Breeden, L. L. (1997). Cell Cycle-dependent Transcription of *CLN1* Involves Swi4 Binding to MCB-like Elements. *The Journal of Biological Chemistry* **272**, 9071-9077.

Pitt, A. R. (1996). Mass Spectrometric Methods. *In* : Price, N. ed. *Proteins labfax*, 174-186. BIOS Scientific Publishers.

Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**, 598-599.

Primig, M., Sockanathan, S., Auer, H. and Nasmyth, K. (1992). Anatomy of a transcription factor important for the Start of the cell cycle in *Saccharomyces cerevisiae*. *Nature* **358**, 593-597.

Provencher, S. W., and Glöckner, J. (1981). Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* **20**, 33-37.

Queralt, E., and Igual, J. C. (2003). Cell Cycle Activation of the Swi6p Transcription Factor Is Linked to Nucleocytoplasmic Shuttling. *Molecular and Cellular Biology* **23**, 3126-3140.

Ren, B., Cam, H., Takahashi, Y., Volkert, T., Terragni, J., Young, R. A., and Dynlacht, B. D. (2002). E2F integrates cell cycle progression and with DNA repair, replication, and G_2/M checkpoints. *Genes and Development* **16**, 245-256.

Reymond, A., Schmidt, S., and Simanis, V. (1992). Mutations in the *cdc10* start gene of *Schizosaccharomyces pombe* implicate the region of homology between *cdc10* and *SWI6* as important for $p85^{cdc10}$ function. *Molecular Genetics and Genomics* **234**, 449-456.

Reymond, A., and Simanis, V. (1993). Domains of $p85^{cdc10}$ required for function of the fission yeast DSC-1 factor. *Nucleic Acids Research* **21**, 3615-3621.

Rhind, N., Baber-Furnari, B. A., Lopez-Girona, A., Boddy, M. N., Brondello, J. -M., Moser, B., Shanahan, P., Blasina, A., McGowan, C., and Russell, P. (2000). DNA Damage Checkpoint

Control of Mitosis in Fission Yeast. *Cold Spring Harbor Symposia on Quantitative Biology* **65**, 353-359.

Rhind, N., and Russell, P. (2000). Chk1 and Cds1: linchpins of the DNA damage and replication checkpoint pathways. *Journal of Cell Science* **113**, 3889-3896.

Rhind, N., and Russell, P. (2001). Roles of the Mitotic Inhibitors Wee1 and Mik1 in the G2 DNA Damage and Replication Checkpoints. *Molecular and Cellular Biology* **21**, 1499-1508.

Rustici, G., Mata, J., Kivinen, K., Lió, P., Penkett, C. J., Burns, G., Hayles, J., Brazma, A., Nurse, P., and Bähler, J. (2004). Periodic gene expression program of the fission yeast cell cycle. *Nature Genetics* **36**, 809-17.

Sambrook, J., and Russell, D. W. (2001) Molecular cloning: a laboratory manual (3rd edition). Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Schultz, S. C., Shields, G. C., and Steitz, T. A. (1991). Crystal Structure of a CAP-DNA Complex: The DNA Is Bent by 90°. *Science* **253**, 1001-1007.

Schien, C. H., and Noteborn, M. H. M. (1988). Formation of soluble recombinant proteins in *Escherichia coli* is favoured by lower growth temperature. *Biol Technology* **6**, 291-294.

Schwob, E. and Nasmyth, K. (1993). CLB5 and CLB6, a new pair of B cyclins involved in S phase and mitotic spindle formation in *S. cerevisiae*. *Genes and Development* **7**, 1160-1175.

Sedgwick, S. G., and Smerdon, S. J. (1999). The ankyrin repeat: a diversity of interactions on a common structural framework. *Trends in Biochemical Sciences* **24**, 311-316.

Sedgwick, S. G., Taylor, I. A., Adam, A. C., Spanos, A., Howell, S., Morgan, B. A., Treiber, M. K., Kanuga, N., Banks, G. R., Foord, R., and Smerdon, S. J. (1998). Structural and Functional Architecture of the Yeast Cell-cycle Transcription Factor Swi6. *Journal of Molecular Biology* **281**, 763-775.

Sheehan, D. (2000). *Physical Biochemistry : Principles and Applications*. Wiley.

Sherr, C. J. (1996). Cancer Cell Cycles. *Science* **274**, 1672-1677.

Sidorova, J., and Breeden, L. L. (1993). Analysis of the SWI4/SWI6 protein Complex, Which Directs G_1/S-Specific Transcription in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **13**, 1069-1077.

Sidorova, J. M., Mikesell, G. E., and Breeden, L. L. (1995). Cell Cycle-regulated Phosphoryaltion of Swi6 Controls its Nuclear Localization. *Molecular Biology of the Cell* **6**, 1641-1658.

Siegmund, R. F., and Nasmyth, K. (1996). The *Saccharomyces cerevisiae* Start-Specific Transcription Factor Swi4 Interacts through the Ankyrin Repeats with the Mitotic Clb2/Cdc28 Kinase and through Its Conserved Carboxy Terminus with Swi6. *Molecular and Cellular Biology* **16**, 2647-2655.

Sipiczki, M. (2000). Where does fission yeast sit on the tree of life? *Genome Biology* **1**, 1011.1-1011.4.

Smith, M. C., Furman, T. C., Ingolia, T. D., and Pidgeon, C. (1988). Chelating Peptide-immobilized Metal Ion Affinity Chromatography: A new concept in affinity chromatography for recombinant proteins. *The Journal of Biochemistry* **263**, 7211-7215.

Smith, D. B., and Johnson, K. S. (1988). Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. *Gene* **67**, 31-40.

Spellman, P. T., Sherlock, G., Zhang, M. Q., Iyer, V. R., Anders, K., Eisen, M B., Brown, P. O., Botstein, D., and Futcher, B. (1998). Comprehensive Identification of Cell Cycle-regulated Genes of the Yeast *Saccharomyces cerevisiae* by Microarray Hybridization. *Molecular Biology of the Cell* **9**, 3273-3297.

Stevens, C., and LaThangue, N. B. (2003). E2F and cell cycle control: a double-edged sword. *Archives of Biochemistry and Biophysics* **412**, 157-169.

Stone, M., and Gordon, C. (2003). Ubiquitin-Dependent Proteolysis by the Proteasome. *In* : Egel, R. ed. The Molecular Biology of *Schizosaccharomyces pombe*: Genetics, Genomics and Beyond. Springer, 381-392.

Sturm, S., and Okayama, H. (1996). Domains determining the Functional Distinction of the Fission yeast Cell Cycle "Start" Molecules Res1 and Res2. *Molecular Biology of the Cell* 7, 1967-1976.

Su, S. S. Y., and Yanagida, M. (1997). Mitosis and Cytokinesis in the Fission Yeast, *Schizosaccharomyces pombe*. *In* : Pringle, J. R., Broach, J. R. and Jones, E. W., eds. The Molecular and Cellular Biology of the Yeast *Saccharomyces*: Cell Cycle and Cell Biology. Cold Spring Harbor Laboratory Press, 765-825.

Sudbery, P. E. (1996). The expression of recombinant proteins in yeasts. *Current Opinion in Biotechnology* 7, 517-524.

Sugiyama, A., Tanaka, K., Okazaki, K., Nojima, H., and Okayama, H. (1994). A zinc finger protein controls the onset of premeiotic DNA synthesis of fission yeast in a Mei2-independent cascade. *The EMBO Journal* **13**, 1881-1887.

Sreerama, N., and Woody, R. W. (1993). A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Analytical Biochemistry* **209**, 32-44.

Taba, M. R. M., Muroff, I., Lydall, D., Tebb, G., and Nasmyth, K. (1991). Changes in a SWI4,6-DNA-binding complex occur at the time of *HO* gene activation in yeast. *Genes and Development* **5**, 2000-2013.

Tahara, S., Tanaka, K., Yuasa, Y., and Okayama, H. (1998). Functional Domains of Rep2, a transcriptional Activator Subunit for Res2-Cdc10, Controlling the Cell Cycle 'Start'. *Molecular Biology of the Cell* **9**, 1577-1588.

Tanaka, K., and Okayama, H. (2000). A Pcl-like Cyclin Activates the Res2p-Cdc10p Cell Cycle "Start" Transcriptional Factor Complex in Fission Yeast. *Molecular Biology of the Cell* **11**, 2845-2862.

Tanaka, K., Okazaki, K., Okazaki, N., Ueda, T., Sugiyama, A., Nojima, H., and Okayama, H. (1992). A new cdc gene required for S phase entry of *Schizosaccharomyces pombe* encodes a protein similar to the *cdc10*$^{+}$ and *SWI4* gene products. *The EMBO Journal* **11**, 4923-4932.

Tanaka, S., and Diffley, J. F. X. (2002). Interdependent nuclear accumulation of budding yeast Cdt1 and Mcm2-7 during G1 phase. *Nature Cell Biology* **4**, 198-207.

Taylor, I. A., Treiber, M. K., Olivi, L., and Smerdon, S. J. (1997). The X-ray Structure of the DNA-binding Domain from the *Saccharomyces cerevisiae* Cell-cycle Transcription Factor Mbp1 at 2.1 Å Resolution. *Journal of Molecular Biology* **272**, 1-8.

Taylor, I., McIntosh, P. B., Pala, P., Treiber, M. K., Howell, S., Lane, A. N., and Smerdon, S. J. (2000). Characterization of the DNA-Binding Domains from the Yeast Cell-Cycle Transcription Factors Mbp1 and Swi4. *Biochemistry* **39**, 3943-3954.

Tournier, S., and Millar, J. B. A. (2000). A Role for the START Gene-specific Transcription Factor Complex in the Inactivation of Cyclin B and Cut2 Destruction. *Molecular Biology of the Cell* **11**, 3411-3424.

Trimarchi, J. M., and Lees, J. A. (2002). Sibling Rivalry in the E2F Family. *Nature Reviews Molecular Cell Biology* **31**, 11-20.

Tye, B. K. (1999). Minichromosome maintenance as a genetic assay for defects in DNA replication. *Methods* **18**, 329-334.

Uchiyama, M., Griffiths, D., Arai, K. -I., and Masai, H. (2001). Essential Role of Sna41/Cdc45 in Loading of DNA Polymerase α onto Minichromosome Maintenance Proteins in Fission Yeast. *The Journal of Biological Chemistry* **276**, 26189-26196.

Weinberg, R. A. (1995). The Retinoblastoma Protein and Cell Cycle Control. *Cell* **81**, 323-330.

Whitehall, S., Stacey, P., Dawson, K., and Jones, N. (1999). Cell Cycle-regulated Transcription in Fission Yeast: Cdc10-Res Protein Interactions during the Cell Cycle and Domains Required for Regulated Transcription. *Molecular Biology of the Cell* **10**, 3705-3714.

Wijnen, H., Landmann, A., and Futcher, B. (2002). The G_1 Cyclin Cln3 Promotes Cell Cycle Entry via the Transcription Factor Swi6. *Molecular and Cellular Biology* **22**, 4402-4418.

Wood, V., Gwilliam, R., Rajandream, M. -A., Lyne, M., Lyne, R., Stewart, A., Sgouros, J., Peat, N., Hayles, J., Baker, S., Basham, D. et al. (2002). The genome sequence of *Schizossacharomyces pombe*. *Nature* **415**, 871-880.

Wuarin, J., Buck, V., Nurse, P., and Millar, J. B. A. (2002). Stable Association of Mitotic Cyclin B/Cdc2 to Replication Origins Prevents Endoreduplication. *Cell* **111**, 419-431.

Xu, R. M., Koch, C., Liu, Y., Horton, J., Knapp, D., Nasmyth, K., and Cheng, X. (1997). Crystal structure of the DNA-binding domain of Mbp1, a transcription factor important in cell-cycle control of DNA synthesis. *Structure* **5**, 349-358.

Yanagida, M. (2000). Cell cycle mechanisms of sister chromatid separation; Roles of Cut1/separin and Cut2/securin. Genes to Cells **5**, 1-8.

Yoshida, S., Al-Amodi, H., Nakamura, T., McInerny, C. J., and Shimoda, C. (2003). The *Schizosaccharomyces pombe cdt2*$^{+}$ Gene, a Target of G1-S Phase-Specific Transcription Factor Complex DSC1, Is Required for Mitotic and Premeiotic DNA Replication. *Genetics* **164**, 881-893.

Zheng, N., Fraenkel, E., Pabo, C. O., and Pavletich, N. P. (1999). Structural basis of DNA recognition by the heterodimeric cell cycle transcription factor E2F-DP. *Genes and Development* **13**, 666-674.

Zhu, Y., Takeda, T., Nasmyth, K., and Jones, N. (1994). *pct1*$^{+}$, which encodes a new DNA-binding partner of $p85^{cdc10}$, is required for meiosis in the fission yeast *Schizosaccharomyces pombe*. *Genes and Development* **8**, 885-898.

Zhu, Y., Takeda, T., Whitehall, S., Peat, N., and Jones, N. (1997). Functional characterisation of the fission yeast Start-specific transcription factor Res2. *The EMBO Journal* **16**, 1023-1034.

Appendices

Appendix I: Bacterial and fission yeast strains

Bacterial strains

Gene	Plasmid	Purpose	Origin	Glasgow collection number
-	pET-28c	Bacterial expression vector (N-terminal His-tag)	Novagen	GB 7
-	pET-14b	Bacterial expression vector (N-terminal His-tag)	Novagen	GB 288
-	pGEX-KG	Bacterial expression vector (N-terminal GST-tag)	Lab stock	GB 159
-	pREP1	Fission yeast expression vector (nmt control)	Lab stock	GB 27
-	pREP3X	Fission yeast expression vector (nmt control)	Lab stock	GB 28
-	pCR2.1	TA cloning vector	Invitrogen	-
cdc10^{+}	pET-28c	Template for site-directed mutagenesis	Lab stock	GB 130
cdc10^{+}	pCR2.1	Intermediate cloning step	This study	GB 313
cdc10^{+}	pET-14b	Bacterial expression of *His-cdc10*$^{+}$	This study	GB 314
res1^{+}	pBC SK^{+}	For cloning *res1*$^{+}$ into pET-28c	Lab stock	GB 160
res1^{+}	pET-28c	Bacterial expression of *His-res1*$^{+}$	This study	GB 201
res1^{+}	pET-14b	Bacterial expression of *His-res1*$^{+}$	This study	GB 289
res1^{+}	pREP3X	Fission yeast expression of *His-res1*$^{+}$	This study	GB 203
res1^{+}	pCR2.1	Intermediate cloning step	This study	GB 300
res1^{+}	pGEX-KG	Bacterial expression of *GST res1*$^{+}$	This study	GB 309
res2^{+}	pBC SK^{+}	For cloning *res2*$^{+}$ into pET-28c	Lab stock	GB 164
res2^{+}	pET-28c	Bacterial expression of *His-res2*$^{+}$	This study	GB 178
res2^{+}	pREP1	Fission yeast expression of *His-res2*$^{+}$	This study	GB 195
res2^{+}	pCR2.1	Intermediate cloning step	This study	GB 302
res2^{+}	pGEX-KG	Bacterial expression of *GST res2*$^{+}$	This study	GB 318
rep1^{+}	pBC SK^{+}	For cloning *rep1*$^{+}$ into pET-28c	Lab stock	GB 121
rep1^{+}	pET-28c	Bacterial expression of *His-rep1*$^{+}$	This study	GB 177
rep1^{+}	pET-14b	Bacterial expression of *His-rep1*$^{+}$	This study	GB 289
rep1^{+}	pREP3X	Fission yeast expression of *His-rep1*$^{+}$	This study	GB 202
rep2^{+}	pBC SK^{+}	For cloning *rep2*$^{+}$ into pET-28c	Lab stock	GB 23
rep2^{+}	pET-28c	Bacterial expression of *His-rep2*$^{+}$	This study	GB 191
rep2^{+}	pET-14b	Bacterial expression of *His-rep2*$^{+}$	This study	GB 310
rep2^{+}	pREP3X	Fission yeast expression of *His-rep2*$^{+}$	This study	GB 343

Fission yeast strains

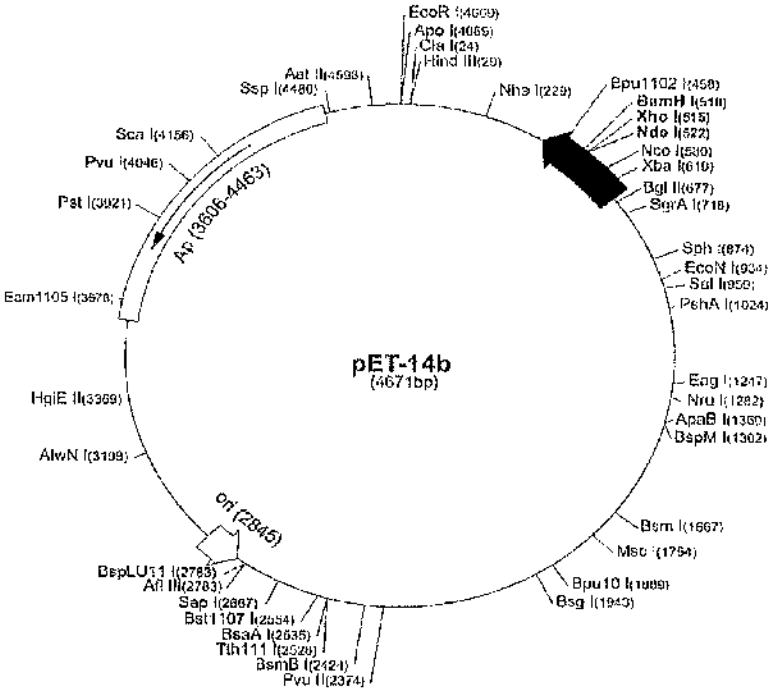
Genotype	Origin	Glasgow collection number
972 h⁻	Lab stock	GG 217
h⁻ *res1::ura4⁻ ade6-M216 ura4-D18 leu1-32*	Lab stock	GG 146
pREP3X: h⁻ *res1::ura4⁺ ade6-M216 ura4-D18 leu1-32*	This study	GG 796
His-res1⁺-pREP3X: h⁻ *res1::ura4⁻ ade6-M216 ura4-D18 leu1-32*	This study	GG 790
His-res2⁺-pREP1: h⁻ *res1::ura4⁻ ade6-M216 ura4-D18 leu1-32*	This study	GG 802
His-rep1⁺-pREP3X: h⁻ *res1::ura4⁻ ade6-M216 ura4-D18 leu1-32*	This study	GG 811
His-rep2⁺-pREP3X: h⁻ *res1::ura4⁺ ade6-M216 ura4-D18 leu1-32*	This study	GG 817
h⁻ *res2::ura4⁺ ura4-D18 leu1-32*	Lab stock	GG 156
pREP3X: h⁻ *res2::ura4⁻ ura4-D18 leu1-32*	This study	GG 801
His-res2⁺-pREP1: h⁻ *res2::ura4⁺ ura4-D18 leu1-32*	This study	GG 808

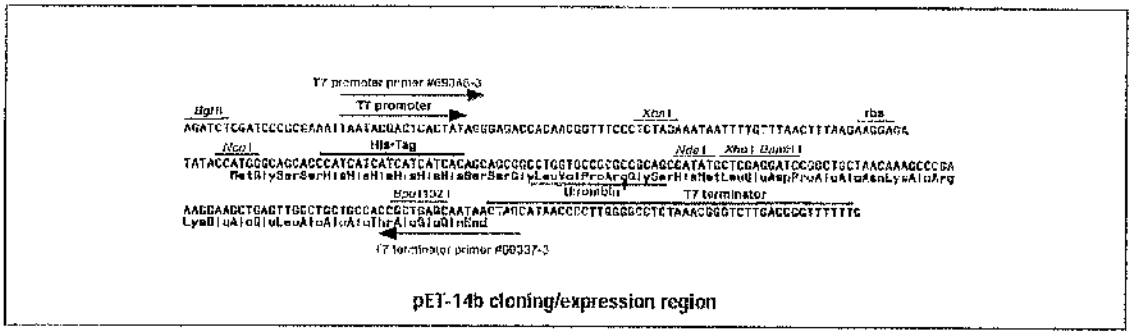
Appendix II: Oligonucleotides

Gene	Purpose	Sequence	Glasgow collection number
pET-28c forward	PCR of polylinker of GB 7 to add *Sal*I site at 5' end of His-tag	CCGGAC**GTCGAC**ATGGGCAGCAGCCATCATCATTTG	GO 458
pET-28c reverse	PCR of 3' polylinker of GB 7	TTGTCGACGGAGCTCGAATTC	GO 459
res1$^+$ forward	PCR of *res1*$^+$ (GB 201) to add *Sma*I site at 3' end	GCGC**CCCGGG**AATGTATAACGACCAAATACATAA	GO 464
res2$^+$ forward	PCR of *res2*$^+$ (GB 178) to add *Sma*I site at 3' end	GCGC**CCCGGG**AATGGCTCCACGTTCTTCC	GO 465
pET-28c reverse	PCR of *rep1*$^+$, *rep2*$^+$, *res1*$^+$ and *res2*$^+$ in GB 7 to include *Xho*I at 3' end	TGGTGGTGGTGGTG**CTCGAG**	GO 466
cdc10$^+$ forward	PCR of *cdc10*$^+$ (GB 130) to include *Nde*I site at 5' end	GATATACCATGGGCAGCAGC	GO 546
cdc10$^-$ forward	Mutagenic removal of internal stop site	TGGGTTCCA**TAC**GATCGAGCA	GO 547
cdc10$^-$ reverse	Mutagenic removal of internal stop site	TGCTCGATCGT**ATG**GAACCCA	GO 548
	Sequencing pCR2.1 plasmids	M13 forward and reverse	MWG-Biotech
	Sequencing pET-28c plasmids	T7 forward and reverse	MWG-Biotech
	Sequencing pREP plasmids 5' to 3'	GAGGAATCCTGCCATATCAT	GO 197
	Sequencing pREP plasmids 3' to 5'	GCGCGCGAATTCACTCGTTGTCGGAGATCAAGAATTTT	GO 191
	Sequencing pGEX-KG plasmids 5' to 3'	CCTTTGCAGGGCTGGCAAGCCACG	GO 471
cdc10$^-$	Sequencing *cdc10*$^-$ internal 5' to 3'	TGGGTTCCATACGATCGAGCA	GO 547
	Sequencing *cdc10*$^+$ internal 3' to 5'	TGCTCGATCGTATGGAACCCA	GO 548
res1$^-$	Sequencing *res1*$^-$ internal 5' to 3'	CCAAATAATCAGGGCCAA	GO 407
	Sequencing *res1*$^-$ internal 3' to 5'	TTGAAAGTTCGCTCAGCTGAG	GO 457
res2$^+$	Sequencing *res2*$^+$ internal 5' to 3'	TTTGCCTGGTGCAGAGGAGCA	GO 241
	Sequencing *res2*$^+$ internal 3' to 5'	GCAAGAGATGCTTCATGGTCA	GO 357
rep1$^-$	Sequencing *rep1*$^-$ internal 5' to 3'	ACTGCCTTGTCGTTGACAAGTTCAGC	GO 409
	Sequencing *rep1*$^-$ internal 3' to 5'	GCACATTATGTGATGCCAG	GO 411
rep2	Sequencing *rep2*$^-$: due to the size of the *rep2*$^-$ orf (660 bp), sufficient sequence coverage was routinely obtained using forward and reverse primers for the appropriate plasmid (e.g. T7 forward and reverse)		

Appendix III: Vector maps

pET-14b sequence landmarks	
T7 promoter	640-662
T7 transcription start	645
His•Tag coding sequence	554-571
Multiple cloning sites	
(*Nde*I - *Bam*H I)	510-526
T7 terminator	404-450
pBR322 ori/gh1	2845
bla coding sequence	3606-4465





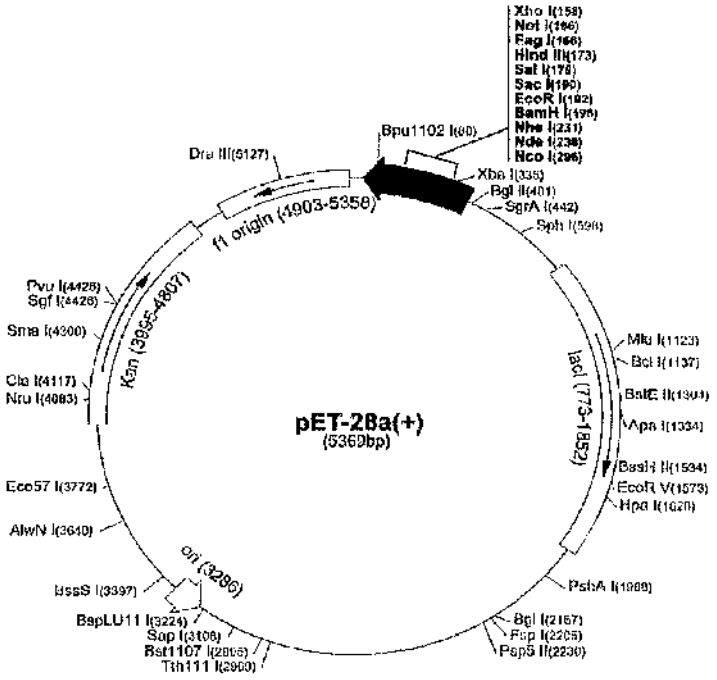
pET-14b cloning/expression region

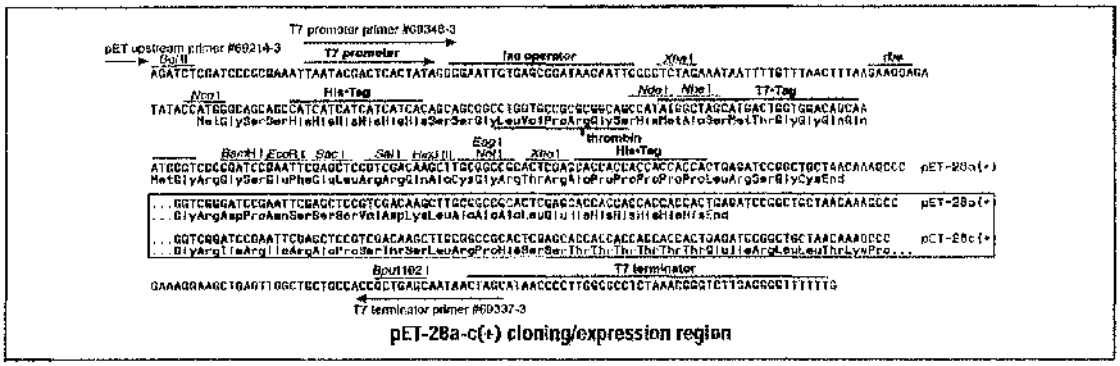
Appendix IIIa: Plasmid map of the pET-14b cloning and expression vector

The pET-14b vector carries an N-terminal His-tag sequence followed by a thrombin cleavage site and unique restriction sites are shown on the circle map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown in more detail in the box (Adapted from www.novagen.com).

pET-28a(+) sequence landmarks	
T7 promoter	370-386
T7 transcription start	369
His•Tag coding sequence	270-287
T7•Tag coding sequence	207-239
Multiple cloning sites (*BamH* I - *Xho* I)	158-203
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	773-1852
pBR322 origin	3286
Kan coding sequence	3995-4807
f1 origin	4903-5358

The maps for pET-28b(+) and pET-28c(+) are the same as pET-28a(+) (shown) with the following exceptions: pET-28b(+) is a 5358bp plasmid; subtract 1bp from each site beyond *BamH* I at 198. pET-28c(+) is a 5357bp plasmid; subtract 2bp from each site beyond *BamH* I at 198.

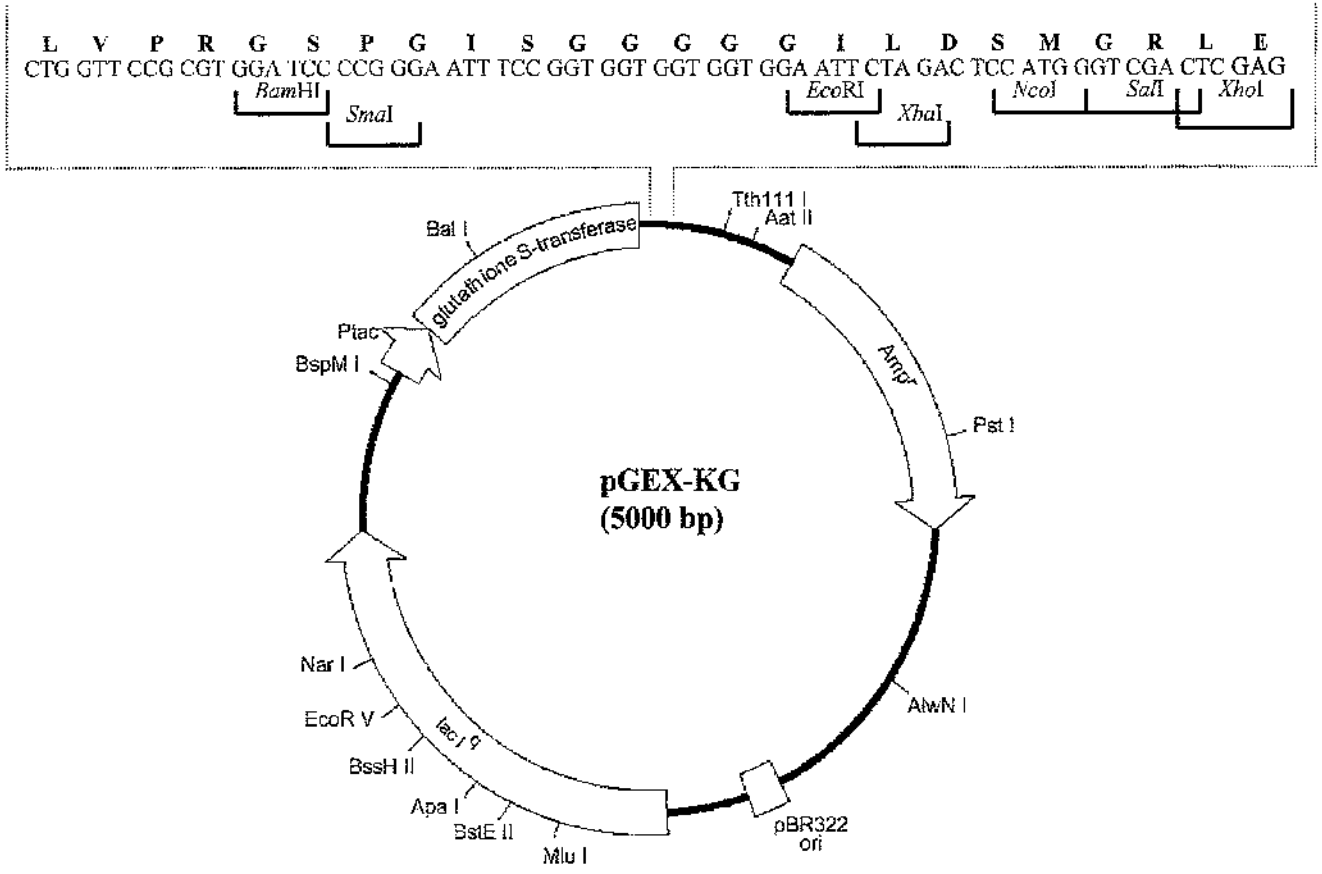




pET-28a-c(+) cloning/expression region

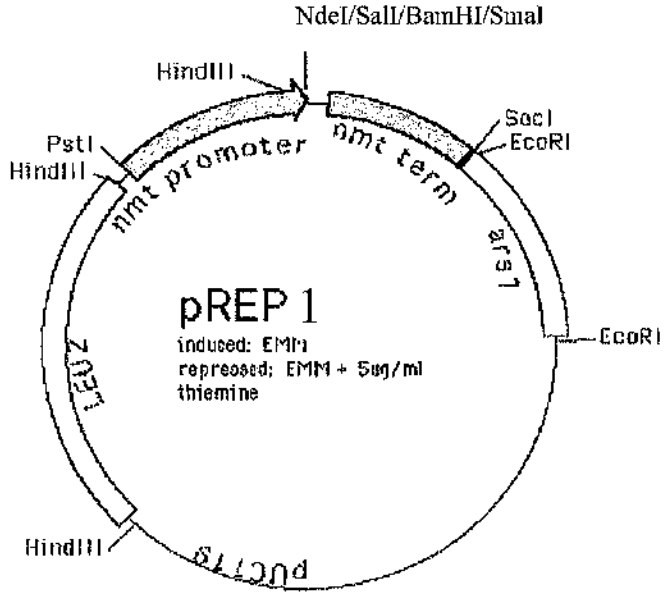
Appendix IIIb: Plasmid map of the pET-28a-c(+) cloning and expression vector

The pET-28a-c (+) vectors carry an N-terminal His-Tag®/thrombin/T7-Tag® configuration plus an optional C-terminal His-Tag sequence. Unique restriction sites are shown on the circle map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown in more detail in the box (Adapted from www.novagen.com).



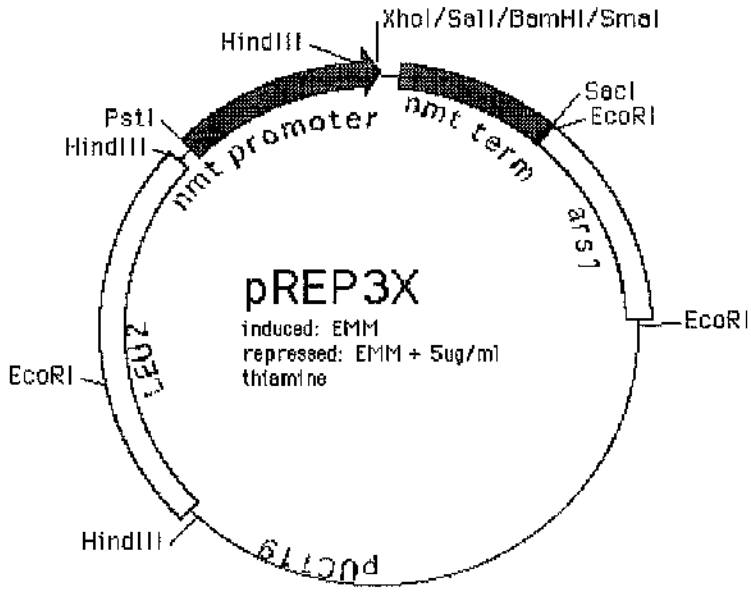
Appendix IIIc: Plasmid map of the pGEX-KG cloning and expression vector

The pGEX-KG vector carries an N-terminal GST-tag sequence followed by a thrombin cleavage site and unique restriction sites are shown on the circle map. The single-letter symbols of amino acids in-frame with the GST-tag are indicated above the nucleotide sequence together with unique restriction sites within the multiple cloning site. The locations of genes for ampicillin resistance (Amp^r) and Lac repressor ($LacI^q$) are indicated (Adapted from Guan and Dixon, 1991).



Appendix IIId: Plasmid map of the pREP1 cloning and expression vector

The pREP1 vector is derived from the original pREP series by addition of a *Nde*I site adjacent to the *Sal*I site (Adapted from Maundrell, 1993).



Appendix IIIe: Plasmid map of the pREP3X cloning and expression vector

The pREP3X vector is derived from the original pREP3 series by addition of a *Xho*I polylinker between the *Bal*I and *Sal*I sites; this deletes the ATG within the polylinker, destroys *Bal*I and recreates *Sal*I (Adapted from www-rcf.usc.edu/~forsburg).