SILENT CHROMATIN FORMATION AND HISTONE MODIFICATIONS IN FISSION YEAST

Manuela Portoso

A dissertation submitted to the University of Edinburgh in conformity with the requirements for the degree of Doctor of Philosophy.
A mio padre
ABSTRACT

The centromere is the chromosomal region which is responsible for the accurate segregation of chromosomes during mitosis and meiosis. Failure to properly segregate replicated chromosomes causes aneuploidy and contributes to cancer progression.

Fission yeast has three centromeres and their structure resembles those of metazoan. They are composed of a central domain that constitutes the kinetochore surrounded by outer repeats packaged in silent chromatin. Silent chromatin is characterized by underacetylated histone H3 and H4 and histone H3 methylated on lysine 9 by the SET-domain methyltransferase Clr4. These specific marks allow the binding of chromodomain proteins such as Swi6 and subsequent recruitment of cohesin. Non-coding transcripts have been detected from both strands of the outer repeats. The RNAi pathway processes these transcripts and the resulting siRNAs direct silent chromatin formation over these repeats causing transcriptional repression of an inserted marker gene.

In order to find new components involved in the assembly of a functional kinetochore, the csp temperature sensitive mutants, previously shown to alleviate silencing of marker genes inserted at the outer repeats of centromere, were analyzed. Specifically components involved in RNA transcription and processing such as a subunit of RNA pol II (Csp3/Rpb7) and splicing factors (Csp4/Cwf10 and Csp5/Prp39) were required for the formation of silent chromatin by affecting the production of centromeric transcripts.

Fission yeast silent chromatin formation is dependent on methylation of histone H3 on lysine 9. In order to identify additional factors that might directly affect modification of histone tails, four genes encoding for SET domain proteins, which may represent putative histone methyltransferases, were deleted (set3, set6, set7 and set9). No effect on silent chromatin formation was detected. In a collaborative study, Set9 was found to specifically methylate lysine 20 on histone H4 and to be involved in general genome integrity. Methylation on lysine 20 of histone H4 together with phosphorylation of serine 129 of histone H2A is necessary to maintain the DNA damage checkpoint activity through the recruitment of Crb2 at the sites of DNA damage.
# TABLE OF CONTENTS

Abstract

Abbreviations

List of figures

## CHAPTER 1

### INTRODUCTION

Chromatin organization

Nucleosome structure

Histone modification and histone code hypothesis

Histone acetylation

Histone phosphorylation

Histone methylation

Lysine methyltransferases

Heterochromatin

RNA and RNAi involvement in heterochromatin formation

Centromere organization

Budding yeasts centromeres: DNA and protein components

*Drosophila melanogaster* centromeres

*Caenorhabditis elegans* holocentric chromosomes

Mammalian centromeres

Centromere inactivation and neocentromere formation

*Arabidopsis thaliana* centromeres

**Components of fission yeast centromeres**

Centromere structure

Epigenetic features of fission yeast centromeres

Domain structure of fission yeast centromeres

Histone modifications at *S. pombe* centromeres

RNA interference and centromere function in *S. pombe*

**Aims of the project**

## CHAPTER 2

### MATERIALS AND METHODS

2.1 *S. pombe* growth and classical genetics

2.2 Molecular genetics

2.3 *S. pombe* physiology

2.4 *S. pombe* molecular biology

2.5 Bacterial methods

2.6 General solutions

Strains used in this thesis

Primers used in this thesis

## CHAPTER 3

**A ROLE FOR cspts MUTANTS IN SILENT CHROMATIN FORMATION**

Introduction

Results
3.1 Analysis of silencing phenotype. Alleviation of silencing
3.2 H3K9me2 and Swi6 binding over the centromeric repeats
3.3 Rad21 recruitment over the centromeric repeats
3.4 Centromeric repeat transcript analyses in csp ts mutants
3.5 Production of siRNAs from centromeric repeats
Discussion

CHAPTER 4
Rpb7, A SUBUNIT OF RNA POL II THAT IS INVOLVED IN CENTROMERIC siRNAs PRODUCTION
Introduction
Results
4.1 Cloning, structural and functional characterization of rpb7-G150D mutation
4.2 Silencing defect analysis
4.3 Rpb7 functions upstream of Dcr1
4.4 RNA pol II binds dg-dh repeats
4.5 Rpb7 is not defective in general transcription at permissive temperature
4.6 Rpb7 specifically promotes transcription of the bottom dg-dh strand
4.7 Localization of RITS complex
Discussion

CHAPTER 5
SPlicing MUTANTS AFFECT SILENT CHROMATIN FORMATION AT FISSION YEAST CENTROMERES
Introduction
Results
5.1 Sequencing of csp4/cwf10 and structural analysis
5.2 Cloning of csp5^ and identification of csp5-39 mutation
5.3 Other splicing mutants affect centromere silencing
Discussion

CHAPTER 6
csp6 CLONING
Introduction
Results
6.1 Attempts to clone csp6 using a genomic library
6.2 Mapping csp6 to chromosome III
Discussion

CHAPTER 7
H4K20 METHYLATION AND SET9 FUNCTION IN GENOME INTEGRITY
Introduction
Results
7.1 Deletion of four genes encoding SET domain proteins
7.2 Viability
7.3 Monitoring chromatin silencing defects
7.4 Analysis of Set9. H4-K20 is methylated in fission yeast and Set9 is the methyltransferase responsible for this modification
7.5 H4K20me3 localization
7.6 Set9 immunolocalization
7.7 Set9 and H4K20me3 have a role in genome integrity
Discussion

CONCLUSION AND PERSPECTIVES

NEW COMPONENTS OF THE RNAi PATHWAY
Rpb7, a component of RNA pol II, is involved in centromeric transcription
csp mutants: a link between general RNA metabolism and RNAi pathway
CHROMATIN STRUCTURE AND DNA REPAIR
SET domain proteins and the role of Set9 in DNA damage checkpoint
LIST OF FIGURES

Figure 1.1
Histone modifications and chromatin states.

Figure 1-2
The DNA sequence of centromeres differs between species.

Figure 1-3
Fission yeast centromere organization.

Figure 1-4
Features of fission yeast chromatin at centromere 1.

Figure 1-5
Centromeric siRNAs direct heterochromatin formation.

Figure 3-1
Schematic representation of centromere 1. Primers and siRNAs map.

Figure 3-2
Monitoring silencing defect on outer repeats of csp ts mutants.

Figure 3-3
Silent chromatin is retained over the centromeric repeats of csp ts mutants except in a csp6 mutant.

Figure 3-4
Swi6-GFP is still localized in the csp ts mutants.

Figure 3-5
The cohesin subunit Rad21 is recruited less efficiently to centromeric repeats in csp ts mutants.

Figure 3-6A
RT-PCR procedure to detect centromeric transcripts.

Figure 3-6B
Model for establishment of silent chromatin at centromeres.

Figure 3-6C
Analysis of centromeric transcripts in csp ts mutants.

Figure 3-7
Analysis of siRNAs from centromeric repeats.

Figure 3-8
Chromatin and RNAi: possible role of csp gene products.

Figure 4-1
Rpb7 from budding yeast, fission yeast and human.

Figure 4-2
rpb7-G150D cells lose H3K9me2 over a marker gene inserted in the dg-dh repeats.

Figure 4-3
rpb7-G150D cells lose Swi6 binding over a marker gene inserted in the dg-dh repeats.

Figure 4-4
Rpb7 acts upstream of Dcr1 in the production of the dg-dh bottom strand.

Figure 4-5
RNA pol II is associated with the dg-dh repeats.

Figure 4-6
rpb7-G150D cells have dramatically reduced dg-dh siRNAs.

Figure 4-7A
Rpb7-G150D cells are proficient in general transcription termination.

Figure 4-7B
Diagram of ura4 engineered strains and primers used for RT-PCR.

Figure 4-7C
rpb7-G150D cells are proficient in general transcription termination.

Figure 4-8
Rpb7 is specifically required to promote initiation of centromere repeat transcription.
Figure 4-9
The RITS complex is localized at the dg-dh repeats in rpb7-G150D cells.

Figure 4-10
Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing.

Figure 5-1
ScSnu114p interactions in the pre-spliceosome.

Figure 5-2
Alignment of ScEF-2, SpCwf10 and ScSnu114p.

Figure 5-3
csp4-39 is an allele of cwf10, encoding for Snu114p homologue.

Figure 5-4
Strategy to clone csp5*.

Figure 5-5
csp5-39 temperature and silencing defects are rescued by five plasmids.

Figure 5-6
All the five plasmids contain prp39* and ght2* genes.

Figure 5-7
csp5-39 is an allele of prp39*.

Figure 5-8
Alignment of Prp39 from S. pombe, human and S. cerevisiae.

Figure 5-9
Prp39 contains a HAT motif in the N-terminus.

Figure 5-10
Strategy to integrate the plasmid linearized at prp39 locus.

Figure 5-11A
Splicing mutants do not affect silent chromatin formation over the dg-dh repeats.

Figure 5-11B
Splicing mutants affect silent chromatin formation over a marker gene inserted at the outer repeats.

Figure 5-12
Table 1. Summary of the features relative to RNAi in splicing mutants.

Figure 5-13
prp2, prp3, prp5, prp12 and csp4/cwf10 cells are not defective in splicing at 25°C.

Figure 5-14
Role for splicing factors at fission yeast centromeres.

Figure 6-1
First strategy to clone csp6 by complementation with a high copy genomic library.

Figure 6-2
Plasmids that suppressed csp6ts phenotype were grouped by restriction pattern.

Figure 6-3
Most of the plasmids that suppress csp6-75/csp6-95 mutation, contain ssa1 gene.

Figure 6-4
Second strategy to clone csp6 by complementation with a high copy genomic library.

Figure 6-5
Diploid breakdown experiment to map csp6 to a chromosome.

Figure 6-6A
csp6 maps on chromosome 3.

Figure 6-6B
csp6 maps on chromosome 3.

Figure 6-7
Narrowing down csp6 loci.

Figure 7-1
Structure of SET domain proteins.

Figure 7-2A
Diagram of the SET domain proteins in S. pombe.
Figure 7-2B
Table 1. Fission yeast SET domain protein and related factors.

Figure 7-3
Deletion of set3, set10, set7 and set9 genes with kanamycin (KAN) cassette.

Figure 7-4
Testing cold and temperature sensitivity of set3, set10, set7 and set9 deletion strains.

Figure 7-5A
Silencing at centromere 1 outer repeats in set3Δ, set10Δ, set7Δ and set9Δ cells remains intact and they are not sensitive to TBZ.

Figure 7-5B
set3, set10, set7 and set9 are not required for silencing of ura4 in outer repeats of the centromere 1.

Figure 7-6
Silencing in the imr region of cen1 remains intact in set3Δ, set10Δ, set7Δ and set9Δ cells.

Figure 7-7
Silencing at the central core remains intact in set3Δ, set10Δ, set7Δ and set9Δ cells.

Figure 7-8
set3Δ, set10Δ, set7Δ and set9Δ cells maintain silencing at the mating type locus and at telomeres.

Figure 7-9
Histone H4 is methylated on K20 in fission yeast and Set9 is responsible for this modification.

Figure 7-10A
H4K20me3 localizes to the nucleus and its nuclear localization is lost in set9Δ cells.

Figure 7-10B
H4K20me3 staining is lost from chromosomes in H4K20R mutant cells.

Figure 7-11A
H4K20me3 does not colocalize with Cnp1 and Cnp1 is still localized in set9Δ cells.

Figure 7-11B
H4K20me3 remains localized with chromatin in swi6Δ and clr4Δ cells.

Figure 7-12
The H4K20me3 modification occurs on euchromatin and is not a mark for centromeric sequences. H4K20me3 does not occur on rDNA.

Figure 7-13
Set9 localizes to the nucleus.

Figure 7-14
Set9 localizes to the nucleolus.

Figure 7-15
Set9 does not localize at centromeres.

Figure 7-16
Diagram of H4K20me3 function in the DNA damage checkpoint.
### ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>bp</td>
<td>base pairs</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>ChIP</td>
<td>Chromatin Immunoprecipitation</td>
</tr>
<tr>
<td>CENP</td>
<td>Centromere Protein</td>
</tr>
<tr>
<td>CID1</td>
<td>Centromere IDentifier 1</td>
</tr>
<tr>
<td>C-terminal</td>
<td>Carboxyl-Terminal</td>
</tr>
<tr>
<td>CTD</td>
<td>Carboxyl-Terminal Domain</td>
</tr>
<tr>
<td>DAPI</td>
<td>4,6-diamino-2-phenylindole</td>
</tr>
<tr>
<td>dH2O</td>
<td>distilled water</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxynucleotide triphosphate</td>
</tr>
<tr>
<td>DSB</td>
<td>double strand breaks</td>
</tr>
<tr>
<td>dsRNA</td>
<td>double strand RNA</td>
</tr>
<tr>
<td>ECL</td>
<td>enhanced chemiluminescence</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylene-Diamine-Tetraacetic Acid</td>
</tr>
<tr>
<td>EMM</td>
<td>Edinburgh Minimal Media</td>
</tr>
<tr>
<td>FITC</td>
<td>fluorescein isothiocyanate</td>
</tr>
<tr>
<td>5-FOA</td>
<td>5-fluoro-orotic acid</td>
</tr>
<tr>
<td>GFP</td>
<td>Green Fluorescence Protein</td>
</tr>
<tr>
<td>HA</td>
<td>haemagglutinin</td>
</tr>
<tr>
<td>HAT</td>
<td>histone acetyltransferase &amp; Half A TPR</td>
</tr>
<tr>
<td>HDAC</td>
<td>histone deacetylase</td>
</tr>
<tr>
<td>HMTase</td>
<td>histone H3 methyltransferase</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin G</td>
</tr>
<tr>
<td>IP</td>
<td>Immunoprecipitation</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>kb</td>
<td>kilobase pairs</td>
</tr>
<tr>
<td>kDa</td>
<td>kilo Daltons</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani broth</td>
</tr>
<tr>
<td>MAC</td>
<td>Mammalian Artificial Chromosome</td>
</tr>
<tr>
<td>MAP</td>
<td>Microtubule Associated Protein</td>
</tr>
<tr>
<td>ME</td>
<td>Malt Extract</td>
</tr>
<tr>
<td>MNase</td>
<td>Micrococcal nuclease</td>
</tr>
<tr>
<td>NAD</td>
<td>Nicotinamide Adenine Dinucleotide</td>
</tr>
<tr>
<td>nmt</td>
<td>no message in thiamine</td>
</tr>
<tr>
<td>N-terminal</td>
<td>amino-terminal</td>
</tr>
<tr>
<td>ORF</td>
<td>Open Reading Frame</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel elecrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffer Saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>PMG</td>
<td>Pombe Media Glutamate</td>
</tr>
<tr>
<td>PTGS</td>
<td>Post Transcriptional Gene Silencing</td>
</tr>
<tr>
<td>RDRC</td>
<td>RNAi directed RNA polymerase complex</td>
</tr>
<tr>
<td>RISC</td>
<td>RNA induced silencing complex</td>
</tr>
<tr>
<td>RITS</td>
<td>RNA induced initiation of transcriptional silencing</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNAi</td>
<td>RNA interference</td>
</tr>
<tr>
<td>RT</td>
<td>room temperature</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl Sulphate</td>
</tr>
<tr>
<td>SET</td>
<td>Suvar3-9, Enhancer of zeste, Trithorax</td>
</tr>
<tr>
<td>siRNA</td>
<td>small intereference RNA</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris-Borate EDTA</td>
</tr>
<tr>
<td>TE</td>
<td>Transposable Elements</td>
</tr>
<tr>
<td>TPR</td>
<td>Tetra trico Peptide Repeat</td>
</tr>
<tr>
<td>TBZ</td>
<td>Thiabendazole</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>------------</td>
</tr>
<tr>
<td>TSA</td>
<td>Trichostatin A</td>
</tr>
<tr>
<td>Tween</td>
<td>polyoxyethylenesorbitan monolaurate</td>
</tr>
<tr>
<td>YAC</td>
<td>Yeast Artificial Chromosome</td>
</tr>
<tr>
<td>YES</td>
<td>Yeast Extract Supplemented</td>
</tr>
</tbody>
</table>

**STANDARD NOMENCLATURE FOR HISTONE MODIFICATIONS USED IN THIS THESIS:**

*Histone - residue* (single letter code plus number)-*modification* (lower case)

eg. H3K9ac

- **Acetylation**, ac
- **Methylation** (lysine, mono-, di-, tri-respectively), me1, me2, me3
- **Phosphorylation**, ph
CHAPTER 1
INTRODUCTION

The genetic information that determines cellular phenotypes is transmitted from generation to generation during the cell division. The genome must be replicated and then segregate correctly to guarantee that the same number of chromosomes (genetic information) in each cell is maintained. Loss (aneuploidy) or gain of genetic information may result in diseases, like Down's syndrome or cancer. During DNA replication and cell division, a lot of mechanisms monitor for genome stability (cell cycle checkpoints). Another important role in genome stability is determined by the assembly of correct centromeres.

Centromeres are the cis-acting regions on chromosomes that determine the site where the protein complex responsible for the chromosomes segregation (kinetochore) assembles. Centromeres ensure the accurate segregation of chromosomes during both meiotic and mitotic cell division. In many organisms DNA sequences at the centromere are composed of repetitive sequences that are packaged in heterochromatin. The kinetochore assembles on a subset of these repeats surrounded by heterochromatin. This heterochromatin appears to play a role in maintaining sister chromatid cohesin during cell division. This Chapter will mainly focus on factors that determine chromatin organization and in particular centromere structure and organization.

CHROMATIN ORGANIZATION.

NUCLEOSOME STRUCTURE.
The genetic information of eukaryotic cells is organized into chromatin. Chromatin is dynamic and undergoes different changes during many biological processes: DNA replication, cell cycle progression, gene expression, DNA repair and recombination. Chromatin consists of nucleosomes; each nucleosome represents a core particle, very well conserved in different species. Each core particles, joined to the others by a linker DNA, is made of a protein octamer of two of each histone, H2A, H2B, H3 and H4 wrapped by 146 base pairs length DNA (Luger et al., 1997). These core histones are highly conserved small basic proteins: the most conserved part is their central region that constitutes the "histone fold domain". The linker DNA between the core particles associates with the linker histone, H1. This histone is not so conserved as core histones and some species may even lack it (such as yeast).

Nucleosomes allow chromatin to fold into higher order structures. The first level is the 10 nm fibre, that then is packaged into a spiral of six nucleosomes per turn or solenoid arrangement
(Kornberg & Lorch, 1999; Hayes and Hanse, 2001) to reach the "compaction" of a 30 nm fibre.

Inside the nucleus the 10 nm filament can be arranged in a higher order structure due to three models of organization: nucleosomes may be stacked into more compact structures (Dorido et al., 2004); the chromatin fibre may form loops allowing dispersed chromosomal regions to associate (Derrek et al., 2004; Chambeyron & Bickmore, 2004); particular regions of the genome may be tethered to scaffolding structures in the nucleus, such as nuclear envelope (Taddei et al., 2004) and the nucleolus (Yusufzai et al., 2004).

Once assembled a particular nucleosome can differ from its neighbours due to DNA modifications or post-translational modification of the histones or by substitution of H3 for H3.3 or CENP-A variants or H2A by H2A2 (Sullivan et al., 1994). These different modifications and variants determine different structure on particular nucleosome or chromosomal regions. With respect to post-translational modifications, histones and in particular their N-terminal tails can carry many different modifications: acetylation, phosphorylation, methylation, ubiquitination, polyADP-ribosylation (Van Holde, 1988). All these modifications are linked with specific biological processes and this has led to propose the existence of a language, referred to as the "histone code", that marks genomic regions (Strahl & Allis, 2000); the code is read by other proteins or protein complexes that are capable of understanding and interpreting the profiles of specific modifications (Strahl & Allis, 2000; Jenuwein, 2001).

Chromatin structure and thus DNA accessibility can be determined by two kind of enzymatic activity. One involves ATP hydrolysing enzymes that re-model chromatin by "shuffling" nucleosomes such as Swi2/snf2/ISWI and are conserved in yeast, human and Xenopus (Kingston & Narlikar, 1999; Vignali et al., 2000). Another enzymatic activity includes a set of enzymes that are able to modify histones covalently, at specific residues, located most commonly at the histone tails (Grunstein, 1995).

**HISTONE MODIFICATIONS AND HISTONE CODE HYPOTHESIS.**

Thus, chromatin is composed of nucleosomes that are characterized by highly conserved histone proteins (H3, H4, H2A, H2B and H1) allowing eukaryotic DNA to be packaged and folded into higher-order chromatin fibres (Luger & Richmond, 1998; Kornberg, 1999). Histones are integral components of the machinery responsible for regulating gene transcription and for other DNA associated processes such as replication, repair, recombination and chromosome segregation. The N-terminal histone tails, which protrude from the surface of chromatin polymer, are protease sensitive and provide an exposed surface possibly for interactions with other proteins. The many post-translational modifications of these tails play an important role in regulating gene expression: binding factors can differently access the DNA depending on the histone modifications. This may then be important to determine the interactions with other factors that regulate transcription. In 2000, Strahl and Allis proposed that a 'language' might be encoded on the histone tail domains. This language then is read by other proteins or
combinations. They referred to this language as the "histone code". The identification and characterization of histone acetyltransferase HAT (Brownell et al., 1996), histone H3 kinase (Hsu et al., 2000), histone arginine and histone lysine methyltransferases (Chen et al., 1999; Rea et al., 2000) reinforce the idea of a histone code hypothesis: the code can be read and interpreted by non-histone proteins that bind chromatin in a modification-sensitive manner, which regulates chromatin organization, transcription, DNA replication and repair.

HISTONE ACETYLATION.

One of the most studied and understood examples of the "histone code" language, is histone acetylation. This modification is the most studied and characterized (Grunstein, 1999) because of the discovery of enzymes responsible not only for this modification, the histone acetyltransferases (HAT) but also for its removal, the histone deacetylases (HDACs). Acetylation of specific lysine residues in H3 (K9, K14, K18, K23) is associated with biological processes such as transcription and DNA replication. Many proteins that regulate transcription have intrinsic HAT activity such as GCN5, PCAF, TAFII250 and the p160 family of nuclear receptor coactivators (Brownell et al., 1996; Mizzen et al., 1996; Sterner & Berger, 2000; Roth et al., 2001). These proteins have a specific structural domain, the bromo-domain whose function is to recognize acetylated lysines (Dhalluin et al., 1999). Similarly, HDACs have been found as components of multi-protein complexes associated with repressive chromatin. Thus histone acetylation is a reversible modification that can determine gene activation. In addition, H3 and H4 N-tails are consistently underacetylated in silent chromatin regions of diverse organisms from budding and fission yeast (Ekwall et al., 1997; Grunstein, 1998; Mellone et al., 2003) to mammals (Jeppesen & Turner, 1993; Taddei et al., 2001).

Upon DNA damage it is possible that specific histone modifications are required to signal and recruit repair proteins. However until recently the evidence for a DNA damage-induced histone code was limited to the phosphorylation of the C-terminal tail of the histone variant H2AX (H2A in S. cerevisiae)(Paull et al., 2000), there is a growing body of evidence supporting a role for histone acetylation in double-strand DNA repair. Mutation of all the four reversibly acetylatable lysines of histone H4 (K5, K8, K12 and K16) causes sensitivity to double-strand DNA-damaging agents (Fisher-Adams & Grunstein, 1995). Mutation of both K14 and K23 on H3 increases sensitivity to expression of the HO endonuclease that cuts dsDNA causing breaks (Qinn & Parthun, 2002). Also the enzymes that alter histone acetylation have been implicated in the repair of DSBs (Chon & Kroy, 2002).

Two recent reports in Saccharomyces cerevisiae have shown that histone acetylation levels actually change around a DSB. H4K16Ac is reduced 40 to 50% following induction of a DSB that can be repaired only by non-homologous end joining (NHEJ) (Jazayeri et al., 2004), and H4K8Ac transiently increases in chromatin flanking the double strand DNA breaks that can be only repaired by NHEJ (Downs et al., 2004). More recently, Tamburini & Tyler (2005) performed a comprehensive analysis of acetylation levels on all nine acetylatable residues of the N-
terminal tails of histones H3 and H4 in budding yeast before, during and after homologous recombinational repair of an HO lesion. They found that localized histone acetylation, followed by histone deacetylation occurs in a manner dependent on homologous recombinatorial repair. These modifications are important for cell viability following DNA repair and are probably mediated by the recruitment of the histone acetylases Gcn5 and Esa1 and the histone deacetylases Sir2, Hstl and Rpd3 during homologous recombination. Thus histone acetylation is involved not only in transcriptional regulation but also in genome integrity.

HISTONE PHOSPHORYLATION.
Histone phosphorylation plays different roles, from chromosome condensation during mitosis, (Bradbury, 1992; Koshland and Strunnikov, 1996) to transcription activation and DNA repair. In particular H3S10 that is involved in transcription activation, becomes phosphorylated by MSK kinase during the induction of immediate-early genes such as c-jun, c-fos and c-myc (Mahadevan et al., 1991; Thomson et al., 1999, 2001; Chadee et al., 1999; Soloaga et al., 2003). Histone H3S10 phosphorylation also influences DNA repair, in which the phosphorylation is catalyzed by the Fyn kinase (He et al., 2005). A very recent role of H3S10 phosphorylation has been found in mammals: when in mitosis H3S10 becomes phosphorylated near a locus enriched for H3K9me3, the chromodomain protein, HP1 is released from its usual binding sites: this suggests the possibility of combinatorial readouts of two near histone post-translational modifications, being one stable (methylation) and the other dynamic (phosphorylation)(Fischle et al., 2005; Hirota et al., 2005) and reinforces the concepts of binary switches and “modification cassette” proposed by Fischle et al., 2003.

In mammalian cells the phosphorylation of H3 on threonine 3 occurs by haspin kinase during metaphase chromosome alignment (Dai et al., 2005) while in plant cells the phosphorylation of threonine 11 of histone H3, is required for mitotic and meiotic chromosome condensation (Houben et al., 2005).

Phosphorylation of the C-terminal tail of the histone variant H2AX (H2A in S. cerevisiae) is a very early response to double-strand DNA breaks. In yeast, H2A phosphorylation is required for the recruitment of chromatin-modifying complexes to the vicinity of DNA damage. The ATP-dependent chromatin-remodelling complex INO80 in the SWI/SNF superfamily (Mizuguchi et al., 2004; Shen et al., 2000) is recruited to dsDNA breaks by the interaction with γ-H2AX (Morrison et al., 2004; Van Attikum et al., 2004). In addition, H2A phosphorylation in yeast is required for the recruitment of the NuA4 histone acetyltransferase complex to a region proximal to a double-strand break where, as previously mentioned, H4K8Ac increases (Downs et al., 2004) and inducing recruitment of cohesins to double strand breaks (Strom et al., 2004; Unal et al., 2004). Thus, not only gene activation but also DNA repair involves combinations of histone modifications and factor interactions that support the idea of the existence of a particular set of histone modifications that recruits DNA repair proteins.
CHAPTER 1  INTRODUCTION

HISTONE METHYLATION.

Together with acetylation and phosphorylation, lysine methylation plays an important role in determining changes in chromatin structure. Methylation occurs on histones H3 and H4 on arginine and lysine residues (van Holde, 1988; von Holt et al., 1989; Strahl et al., 1999). Arginine can be either mono or dimethylated and the dimethylation can occur in symmetric or asymmetric configurations. Arginine methylation is involved in transcriptional activation: a nuclear receptor that interacts with an activator factor, CARM1, has histone H3 specific arginine methyltransferase activity. This activity is required for transcriptional activation and it works in association with co-activators containing HAT activity (Chen et al., 1999).

Methylation of lysines is known to occur on histone H3 (K4, K9, K27, K36, K79) and H4 (K20). Moreover, lysine residues can be mono-, di- or trimethylated and this add further complexity to this histone mark. In contrast to acetylation and phosphorylation, lysine methylation appeared to be a relatively stable modification. However, recently, the idea of the existence of histone demethylating enzyme has arisen: arginine is susceptible to enzymatic conversion forming a non-standard amino acid. Human peptidylarginine deiminase 4 (PAD4/PADI4) converts monomethylated arginine in histone H3 and H4 to citrulline by demethylimination. The fate of the citrulline residue remains unknown and PAD4/PADI4 also acts on unmethylated arginine to create citrulline, which can no longer be methylated by methyltransferase suggesting that the role of PAD4/PADI4 may be to deplete the histone substrate for methyltransferases (Cuthbert et al., 2004; Wang et al., 2004). Recently the LSD1 protein has been characterized as lysine-specific demethylase 1 and shown to be a flavin-dependent amine oxidase that is able to specifically demethylate mono and dimethylated histone H3K4 (Shi et al., 2004).

Active transcription is usually marked by H3K4 methylation in mammalian cells and budding yeast (Strahl et al., 1999): H3K4me2 and H3K4me3 correlate with active transcription (Litt et al., 2001; Noma et al., 2001) and H3K4me3 is enriched on fully activated promoters (Santos-Rosa et al., 2002). However, in budding yeast H3K4me2 is also involved in rDNA silencing (Briggs et al., 2001; Bryk et al., 2002).

H3K36 methylation also correlates with transcription and it is coupled with transcription elongation (Howe et al., 2003; Schaft et al., 2003; Bannister et al., 2005; Kizer et al., 2005).

Marks for silent chromatin are typically H3K27 and H3K9 methylation in mammalian cells, Drosophila and plants (Cao et al., 2002; Czermin et al., 2002; Peters et al., 2003; Lindroth et al., 2004). In budding yeast these two modifications on H3 have not been detected, whilst in fission yeast only K9 is dimethylated at heterochromatic loci (Nakayama et al., 2001; Tamaru et al., 2003). While bromo-domain proteins preferentially bind acetylated lysine on histone tails (Dhalluin et al., 1999; Jacobson et al., 2000), chromo-domain proteins, such as heterochromatin-associated protein 1 (HP1 in metazoans and Swi6 in fission yeast) preferentially bind to methylated H3K9 (Bannister et al., 2001; Lachner et al., 2001; Nakayama et al., 2001). This further supports the histone code hypothesis. In addition to heterochromatin silencing, H3K9me/HP1 is also involved in silencing of specific genes in euchromatin. In this
case the SUV39H1 HMTase is recruited, through interaction with HP1, by the tumour suppressor retinoblastoma protein (Nielsen et al., 2001; Vandel et al., 2001). However, H3K9me2, H3K9me3 and HP1γ have also been found to be present on actively transcribed genes in mammalian cells (Vakoc et al., 2005). This prompts one to think that not only the histone modifications but also all the surrounding environment allows the underlying code to be read, in this case, by the transcription machinery. In fact both the presence of HP1γ and H3K9 methylation are dependent upon elongation by RNA pol II (Vakoc et al., 2005).

H3K79 methylation in budding yeast is involved in telomeric silencing and Sir (silent information regulatory proteins) association (Ng et al., 2002; Feng et al., 2002; Lacoste et al., 2002; van Leeuwen et al., 2002). Many non-transcriptional processes are also regulated or influenced by histone methylation such as DNA repair and recombination.

Methylation of H3K79 together with ubiquitination of H2BK123 in budding yeast and human cells whilst methylation of H4K20 in fission yeast (as discussed in Chapter 7 of this thesis), have been shown to play an important role in the signalling of double-strand DNA damage to the cell (Giannattasio et al., 2005; Huyen et al., 2004; Sanders et al., 2004). It is thought that tandem Tudor domains, present in the DNA damage checkpoint adaptor proteins (53BP1 in mammals, Rad9 in S. cerevisiae, Crb2 in S. pombe), may bind the methylated histone (Huyen et al., 2004) and somehow influence the formation of damage-induced foci.

**Lysine methyltransferases.**

Most of the histone lysine methyltransferases are characterized by a particular domain, the SET domain. This domain is evolutionary conserved and it was initially identified in the *Drosophila* position effect variegation (PEV) suppressor SU(VAR)3-9 (Tschiersch et al., 1994), the Polycomb-group protein Enhancer of zeste (Jones and Gelbart, 1993) and the trithorax-group protein Trithorax (Stassen et al., 1995). The conserved motif is characterized by 4 amino acids, NHSC and it is necessary for the catalytic activity. There are 73 SET domain protein in human, 7 in S. cerevisiae, 13 in S. pombe, 41 in Drosophila and 37 in Caenorhabditis elegans. The SET domain proteins can be classified in four families, from the homology within their SET domain: the SUV39, the SET1, the SET2 and RIZ SET (Kouzarides, 2002).

The SET domain of SUV39H1 was found by Jenuwein and colleagues (Rea et al., 2000) to have a very high similarity with some plant SET-domain-containing methyltransferases (Klein and Houtz, 1995; Zheng et al., 1998). They tested SUV39H1 in a methyltransferase activity assay and shown that in fact it specifically methylates histone H3 on lysine 9. The *S. pombe* SUV39 homolog, known as CIR4, was also shown to have histone methyltransferase activity and specifically methylates histone H3 on lysine 9 (Nakayama et al., 2001; Bannister et al., 2001). The sequences within the SET domain are not however sufficient for enzymatic activity: two cysteine-rich regions PRE-SET and POST-SET domains are necessary for SUV39 enzymatic activity and are probably important to guarantee the specificity for this SET domain to methylate H3K9 (Rea et al., 2000). The SUV39 family comprises SUV39H1, SUV39H2 (O'Carroll et al., 2000).
2000), G9A (Tachibana et al., 2001) and ESET (Yang et al., 2002). SUV39H1 and SUV39H2 also have a chromo-domain that does not bind H3K9 methylated. G9A has an additional domain, the Ankyrin repeats, probably involved in protein:protein interactions. It can also methylate K27 on H3 in addition to K9. ESET and CLLL8 contain an expanded SET domain. In addition both these proteins also contain an MBD domain, a structure that is potentially capable of recognizing methylated DNA (Ng & Bird, 1999).

Two human SET1 proteins, hSET1A and hSET1B, are highly similar in their SET domain to budding yeast SET1 that methylates K4 on H3 (Santos-Rosa et al., 2002). Also proteins of the polycomb and trithorax (trx) group of genes whose function is connected to the epigenetic regulation of developmentally regulated genes are comprised in this family. They include the polycomb proteins EZH1 and EZH2, the TRX proteins MLL, MLL2 and the related protein ALR. This family is characterized by the presence of the SET domain at the carboxyl terminus of the protein, followed by a POST-SET domain. Only EZH1 and EZH2 do not have a POST-SET domain. EHZ2 is part of a deacetylase complex (Van der Vlag, 1999) and it binds a protein homologous to the ATPase SNF2 (Cardoso et al., 1998). The TRX-related MLL proteins are linked to cancer, in particular to leukaemias (Caldas & Aparicio, 1999; Okada et al., 2005). Interestingly, MLL has a bromo-domain (an acetyl binding structure). It also has a MBD domain that may allow it to recognize methylated DNA (Caldas et al., 1999).

The SET2 family has a SET domain homologous to budding yeast SET2 that methylates K36 on H3 (Strahl et al., 2002). Members of this class have a SET domain that is always followed by a POST-SET and preceded by a cysteine rich sequence that may play a role in enzymatic activity, the AWS domain. Recently it has been discovered at the C-terminus of budding yeast Set2 a conserved domain called SRI for Set2 Rpb1 interacting domain. This domain mediates the interaction between Set2 and RNA pol II and coupled H3K36 methylation with transcript elongation (Kizer et al., 2005).

The RIZ (retinoblastoma-interacting zinc finger) protein has a SET domain in the amino terminus, it does not have a PRE or POST-SET and presents zinc finger motifs clustered at the carboxyl terminus. Members of this family exhibit changes in the highly conserved motif (NHSC) (Rea et al., 2000). The RIZ family comprises transcriptional regulators controlling differentiation and somehow they are involved in cancer (Steele-Perkins et al., 2001).

Not all histone methyltransferases contain a SET domain. The S. cerevisiae Dot1 protein methylates H3K79. Homologs have been identified in humans (hDOT1) and Drosophila (Grappa) but it appears to be absent from S. pombe (Feng et al., 2002; Shamower et al., 2005). Dot1 does share one notable structural feature with the SET domain-containing methyltransferases. Crystallographic studies of Dot1 and members of the SET family show that
Figure 1-1 Histone modifications and chromatin states.
A) Modifications on histone H3 and H4 N-termini that are known to be important for active or inactive chromatin.
B) Representation of euchromatin and heterochromatin as accessible or condensed nucleosome fibers containing acetylated (Ac) and methylated (Me) histone N-termini recognized respectively by Bromodomain and Chromodomain (HP1) proteins.
they both possess a narrow channel that passes through these enzymes; this channel connects the target lysine side chains, which enters through one side of the enzyme, to the AdoMet cofactor, which enters through the opposite side (Min et al., 2003; Trievel et al., 2002). The particular structure of this channel determines the mono, di or trimethyltransferase specificity of the enzyme.

HETEROCHROMATIN.

Euchromatin and heterochromatin are domains that characterize different portions of the genome. Euchromatin is generally open chromatin in actively transcribed parts of the genome while heterochromatic parts of the genome tend to be inaccessible to DNA binding factors and are transcriptionally silent. Large blocks of heterochromatin are found at centromeres and telomeres, while smaller heterochromatic domains are interspersed along chromosomes arms. Another function of heterochromatin may be to stabilize repetitive DNA sequences at centromeres, telomeres, and elsewhere in the genome by inhibiting recombination between homologous repeats (Grewal et al., 1997; Guarente et al., 2000). Heterochromatin not only has this important role in maintaining genome integrity, but regulates also gene expression during development and cellular differentiation. Heterochromatin domains are found to stably inactivate developmental regulators such as the homeotic gene clusters in Drosophila and mammals, and the mating type genes in fungi (reviewed in Cavalli, 2002). Also, dosage compensation in female mammals involves the heterochromatic inactivation of one of the two X chromosome in somatic cells (Avner and Heard, 2001).

Heterochromatin consists of densely packed nucleosomes, kept together by specialized proteins (Figure 1-1). In fact, DNA in heterochromatin is less accessible to nuclease and reporter genes inserted into heterochromatic regions of the genome tend to be silenced (Allshire et al., 1995; Henikoff, 2000; Richards & Elgin, 2002; Maison & Almouni, 2004). Histones and their post-translational modifications determine assembly of heterochromatin. Generally heterochromatin is associated with hypoacetylated histones (reviewed in Moazed et al., 2001) and enzymes that directly modify histones and factors that bind to histones are required for heterochromatin assembly (reviewed in Richards, 2002)). For example in budding yeast, the silent information regulator (SIR) genes encode for proteins like Sir2, Sir3 and Sir4, that are required for assembly of heterochromatin at the mating type loci and telomere DNA (Aparicio et al., 1991). These proteins bind to specific loci of budding yeast genome where histones are deacetylated. Sir3 and Sir4 bind to deacetylated histones in vitro and Sir2 is a NAD dependent histone deacetylase necessary for heterochromatin assembly (Hoppe et al. 2002; Carmen et al., 2002; Gasser, 2001). Deacetylated histones are also a heterochromatic mark in fission yeast and metazoa: several histone deacetylases, including Sir2-like NAD-dependent and
Hda1/Rdp3-like NAD-independent deacetylases are required for silencing (Shankaranarayana et al., 2003; Sekinger and Gross, 2001).

Methylation of H3K9 marks heterochromatin in fission yeast, *Drosophila* and mammals. This modification is carried out by the methyltransferase Su(var)3-9 in *Drosophila*, SUV39H1 in human and Clr4 in fission yeast (Rea et al., 2000; Nakayama et al., 2001). In heterochromatic regions, not only is H3 methylated on K9 but also a conserved protein HP1/Swi6 associates by binding this methylated histone (Bannister et al., 2001; Lachner et al., 2001). Interestingly, HP1 as well as Swi6, can dimerize and this dimerization may hold nucleosomes together and facilitate compaction of the chromatin fibre (Jones et al., 2000; Jenuwein, 2001; Cowieson et al., 2000).

In *Drosophila*, HP1 and Su(var)3-9 coincide predominately at Transposable Elements (TEs) and pericentric genes (Greil et al., 2003). Su(var)3-9 also binds independently of H3K9 methylation at genes scattered along the chromosome arms. HP1 in turn binds independently of Su(var)3-9 to genes on chromosome 4, a small chromosome almost entirely packaged into heterochromatin. This suggests that another methyltransferase or another mechanism could be responsible for recruiting HP1 to these loci (Greil et al., 2003). On average, target genes of HP1 and Su(var)3-9 show reduced expression levels. However, especially at pericentric regions and on chromosome 4, many genes bound by both these proteins are expressed in many cases at high levels (Greil et al., 2003). This shows that heterochromatin can be heterogeneous.

In mammalian cells and *Drosophila*, heterochromatin is characterized by H3K9me2 and HP1 binding and also by another modification, H4K20me3. It has been proposed that a sequential mechanism induces H3K9 and H4K20 methylation at constitutive heterochromatin (Schotta et al., 2004) (see Chapter 7 for further discussion).

DNA methylation is also believed to contribute to the stability of silenced chromatin states in higher eukaryotes with complex genomes. There is a feedback type mechanism between DNA and histone methylation such that one promotes maintenance of the other (Tamaru et al., 2001; Johnson et al., 2002; Soppe et al., 2002). Indeed, in *Arabidopsis thaliana*, the left arm of chromosome 4 contains a large heterochromatic region, which is also dense in TEs. Lippman et al., (2004) performed detailed mapping of histone modifications and DNA methylation in this region: the results indicate that H3K9 methylated and methylated DNA, which are both thought to be repressive chromatin marks, are primarily concentrated on TEs and related tandem repeats.

**RNA and RNAi INVOLVEMENT IN HETEROCHROMATIN FORMATION.**

At the DNA level, heterochromatin is composed of degenerate retrotransposon sequences, tandem arrays of a simple repeat unit, or a combination of both (Volpe et al., 2002; Weller et al., 1995). In addition to repetitive DNA elements, noncoding RNAs are involved in targeting to specific region of the genome, heterochromatin complexes (Henikoff, 2000; Hsieh, 2000). For
example RNA can direct chromatin-modifying activities required for dosage compensation in Drosophila and mammals (Park, 2001) and also in genomic imprinting in mammals (Sleutels, 2002). Also X-inactivation center originates from Xist RNA that initiates the silent state. Xist RNA is not required for inheritance of the silent state (Avner, 2001). In addition, silencing is also regulated by Tsix, a Xist antisense transcript (Cohen et al., 2002). In Drosophila another RNA, rox RNA assembles with the MSL (male specific lethal) complex. This complex associates with the male X chromosome along its length in an early step during dosage compensation (Park et al., 2001).

A growing body of evidence suggest that heterochromatin formation, in many cases is regulated by RNA interference (RNAi) or involves homolog non coding RNA and possibly RNAi (Volpe et al., 2002). RNAi is initiated by the ribonuclease III Dicer that generates small interfering RNAs, siRNAs from dsRNA templates. siRNAs are then loaded into the RISC complex (RNA induced silencing complex) which then acts as guide to target complementary mRNA for degradation. The key component of RISC is a PIWI domain protein, Argonaute, that can bind small RNAs and in some cases execute target-mRNA cleavage (Ya et al., 2003; Liu et al., 2004; Meister et al., 2004). Another component, of the RNAi machinery in some organisms is RNA-dependent RNA polymerase (RdRp), which synthesizes dsRNA from ssRNA template to initiate or amplify the RNAi reaction (Sijen et al., 2001; Martienssen, 2004).

Proteins involved in the RNAi machinery (Dicer, Argonaute and RdRP) and the siRNAs themselves, can direct epigenetic alterations at DNA or chromatin level. This phenomenon known as transcriptional gene silencing (TGS), was first observed in plants: dsRNAs homologous to promoter DNA results in silencing that can be propagated even when the trigger is subsequently removed (Mette et al., 2000). This silencing is associated with de novo methylation of the gene homologous to the dsRNA.

TGS mediated by RNAi is also involved in silencing retrotransposable elements in the Arabidopsis thaliana genome (Lippman et al., 2003; Lippman et al., 2004).

In Tetrahymena, small RNAs direct the elimination of particular DNA sequences from the macronucleus (Mochizuki et al., 2004). In fact an Argonaute homolog, TWI1 and 28 nt long siRNAs are required for the recognition of the DNA to be eliminated that is usually marked by H3K9 methylation (Mochizuki et al., 2004; Taverna et al., 2002). The rest of the sequences are amplified and become transcriptionally active.

In fission yeast S. pombe, RNAi is required to establish and maintain silencing at centromere and only for maintenance of silencing at mating type locus (Volpe et al., 2002; Hall et al., 2002). The current model suggests that both strands of the outer repeats at centromere (dg-dh/K-L) are transcribed to produce dsRNAs. These dsRNAs are processed by Dicer producing complementary siRNAs. Incorporation of these siRNAs into RITS/RISC (RNA induced transcriptional silencing complex), which contains Ago1 (Argonaute), Chp1 (chromo-domain protein) and Tas3, directs H3K9me2 over homologous chromatin (Verdel et al., 2004). This process also requires the histone methyltransferase Ctr4 (SET domain protein related to
mammalian Suv39) that leads to the formation of a binding site for the chromodomain protein Swi6. Cohesin subunit Rad21 is then recruited and correct chromosome segregation is ensured (Bernard et al., 2001; Nonaka et al., 2002).

In Drosophila piwi and aubergine mutations (argonaute homologs) and the helicase spindle-E (homeless) alleviate silencing of heterochromatic mini-white arrays (Pal Bhadra et al., 2004; Dorer & Henikoff, 1994). However no dsRNA or siRNAs homologous to the mini-white array has been found so far. By small RNA profiling during Drosophila development, it has been found that there is a population of repeat-associated siRNAs (rasiRNAs) that target sequences such as TE, satellite and microsatellite DNA (Aravin et al., 2003).

Studies in vertebrate cells indicate that the RNAi also impinges on chromatin status in vertebrate cells: in a Dicer depleted chicken-human hybrid system, defective sister chromatid cohesion of human chromosome 21 in the DT40 chicken cell line was observed due to displacement of cohesion and HP1 homolog. In addition, small RNAs homologous to human alpha-satellite DNA were detected (Kanellopoulou et al., 2005). In human cells, siRNAs directed against the promoter of an integrated GFP reporter construct can induce transcriptional silencing of GFP by DNA methylation (Morris et al., 2004). However, there are also data that suggest that RNAi does not result in DNA methylation of homologues sequence (Park, 2004; Svoboda et al., 2004).

As in S. pombe, centromeric transcripts have been detected from mouse and human satellite repeats (Lenhertz et al., 2003; Saffery et al., 2003) and RNA has a role in the higher order structure of pericentromeric heterochromatin (Maison et al., 2002).

In sum, the RNAi machinery is both present and functional in fungi, plants and metazoans and it seems to have a conserved role in silent chromatin formation directed by small RNAs.

CENTROMERE ORGANIZATION.

The centromere represents the part of the chromosome that determines the chromosome movement during mitotic and meiotic cell division where the kinetochore is assembled and the spindle microtubules are attached (Clarke, 1998; Nasmyth, 2002).

Its functionality is essential for normal sister-chromatid cohesion and separation (Miyazaki & Orr-Weaver, 1994; Karpen and Allshire, 1997) and it depends on heterochromatin formation. Heterochromatin proteins are associated with centromeric DNA, usually repeat sequences that surround the kinetochore domain and are required for proper sister-chromatid cohesion and chromosome segregation (Bernard et al., 2001; Nonaka et al., 2002; Peters et al. 2001; Hall et al., 2003). In addition centromeres are not merely specified by their DNA sequence: they exhibit properties of an epigenetic locus and behave as a self-replicating protein complex that resides on CEN DNA but is not determined by it (Sullivan et al., 2002).

Centromeres/kinetochores varies in sizes and forms (Sullivan et al., 2001): most eukaryotes have monocentric chromosomes. Holocentric chromosomes are present in nematodes and
**Figure 1-2** The DNA sequence of centromeres differs between species.

a) *Saccharomyces cerevisiae* centromere function depends on a small region containing the three conserved elements CDEI, II and III.

b) *Schizosaccharomyces pombe* contains a unique central core, which Cnp1 is localized to, flanked by inverted inner and outer repeats.

c) The minichromosome *Dp 1187*, the only defined *Drosophila melanogaster* centromere, consists of a core of 5-bp satellites and transposons, flanked by other repetitive DNA.

d) Human centromeres consist of alpha-satellite DNA tandemly arranged into higher-order arrays, which extends over megabases.

e) *C. elegans* kinetochores assemble along the length of each chromosome.

(Sullivan B. et al., 2001)
crayfish. Centipedes have both holocentric and monocentric chromosomes in the same nuclei (Pimpinelli and Goday, 1989). Below are summarized the details of centromere structure and components in several model organisms (Figure 1-2). A detailed section about fission yeast centromeres where kinetochore assembly relies on epigenetic regulation follows next.

**BUDDING YEAST CENTROMERES: DNA AND PROTEIN COMPONENTS.**

The best-defined centromeric DNAs are those of *Saccharomyces cerevisiae*. Budding yeast has 16 chromosomes and each centromere consists of an essential conserved 125-bp sequence with three functional elements that recruit centromere proteins (Hyman and Sorger, 1995). Despite budding yeast centromeres display some functional and structural characteristics conserved in other centromere eukaryotes, they are the only ones to be strictly dependent on their sequence to be functional. However they do not lack complexity and lots of proteins that associate, are conserved also in other eukaryotes (reviewed in Pidoux and Allshire, 2000; Clarke, 1998; Kitagawa & Hieter, 2001). The DNA-inner kinetochore protein interface is composed of protein components interacting with the three centromeric DNA elements: CDE I, II and III.

The CDE I 8 bp sequence is associated with the centromere-binding factor 1, Cbf1/Cpf1 or Cep1 that contains a helix-loop-helix DNA binding domain, and mediates both transcriptional regulation and chromosome segregation. Both CDEI and Cbf1 are not essential for viability. The 78-bp region of CDE II is made of AT-rich DNA and it acts like a spacer between the conserved CDE I and CDE III DNA elements. Mutations in the gene encoding Cse4, the centromere specific histone-H3 like protein, have genetic interaction with mutations in the CDE II and CDE I DNA regions. This indicates that this specialized centromere nucleosome containing Cse4, interacts with the DNA in these regions (Stoler et al., 1995; Keith & Fitzgerald-Hayes, 2000).

Only CDE III is essential and point mutations in this region destroys centromere function. The main CDE III DNA-binding complex is CBF3, which is composed of Ndc10, Cep3, Ctf13 and Skp1 and Sgt1 (Kaplan et al., 1997; Kitagawa et al., 1999). Three other centromere proteins, Ctf19, Mcm21 and Okp1, are part of a complex that links CBF3 with other components of budding yeast centromere such as Cbf1, Mif2 and Cse4 (Ortiz et al., 1999). Mif2 shares 38% similarity over 558 aminoacids with the conserved kinetochore component CENP-C and it binds centromeres as judged by Chromatin Immunoprecipitation (Meluh & Koshland, 1997). Bir1 is part of the Bir1/Cut17/Survivin family that is involved in Aurora kinase function and it interacts by yeast-two hybrid with Ndc10, an inner kinetochore protein (Yoon & Carbon, 1999). The Aurora kinase homolog in budding yeast, Ipl1, interacts with Sli15, the budding yeast homolog of INCENP passenger protein (Kang et al., 2001). Both these proteins bind microtubules *in vitro* and are associated with the centromeric DNA *in vivo* (Kang et al., 2001).
Spindle Pole Body proteins (SPB), Ndc80, Nuf2, Spc24 and Spc25 have been shown to bind centromeres in vivo (Wigge & Kilmartin, 2001). Homologues of all four are found in fission yeast and Ndc80 and Nuf2 have human counterparts. Other proteins that may provide a link between kinetochore and spindle microtubules are Slik19 and Mtw1, this last homologue of fission yeast Mis12, and Dam1, a microtubule binding protein that is physically and functionally associated with kinetochore (Jones et al., 2001).

The centromeres of all eukaryotic organisms mediate the same basic mechanism to coordinate the events that lead to chromosome segregation but, unlike budding yeast, their activity is not dependent on specific cis-acting DNA sequences. Several studies have led to the conclusion that kinetochores may be assembled on an atypical DNA/chromatin region and subsequently propagated in an epigenetic manner (Karpen & Allshire, 1997).

**Drosophila melanogaster CENTROMERES.**

*Drosophila* centromeres are characterized by repetitive DNA. The only characterized *Drosophila* centromere derives from a minichromosome *Dp1187* and it is contained within a specific 420 kb region of the centric heterochromatin. The essential core of 220 kb, composed predominantly of the complex island *BoraBora*, guarantees stability but full transmission requires also 200 kb of DNA flanking both sides (Murphy & Karpen, 1995). The DNA sequence of this centromere is characterized by AATAT and AAGAG satellites, 85% (370 kb), whereas transposable elements comprise about 10% of it. These are found interspersed in the AATAT but not in the AAGAG satellite and also clustered in the terminal region of the centromere called *Maupiti* (Sun et al., 1997). It is thought that *BoraBora* includes the sites of kinetochore formation, based on the absolute requirement for this region, whereas the flanking satellite DNA is responsible for sister-chromatid cohesion and the assembly/stability of a fully functional kinetochore (Murphy & Karpen, 1995). This model is supported by cytological studies: the domain containing the CID1/CENP-A chromatin is flanked by, but does not overlap with, chromatin containing proteins involved in centric heterochromatin function, such as Su(var)2-5/HP1 and Prod (Proliferatio disrupter) (reviewed in Sullivan et al., 2001). In addition, Mei-S332, a protein required for the maintenance of sister-chromatid cohesion in meiosis is located near but not in the CID1-containing chromatin (Blower & Karpen, 2001).

By cytological studies the centromeric chromatin in humans and fruit flies, has been shown to contain interspersed blocks of CENP-A/CID and H3 nucleosomes, called CEN chromatin (Blower et al., 2002). A very recent study by Sullivan & Karpen, 2004, defined 'centrochromatin' the CEN chromatin in humans and *Drosophila* that is distinct for the histone modifications it presents, from euchromatin and heterochromatin. It contains the histone H3 variant CENP-A and a distinct pattern of H3 modification. H3K4me2, usually present in euchromatin, is interspersed with the CENP-A/CID1 on extended chromatin fibers (Sullivan & Karpen, 2004). However the centrochromatin can be distinguished from active euchromatin by the absence of acetylated H3
and H4. The function of this centromere-associated chromatin is still to be determined. It is possible that centrochromatin somehow serves to promote the assembly of CENP-A/CID1 nucleosomes and thus it contributes to the epigenetic inheritance of centromeric chromatin (Dunleavy et al., 2005).

**Caenorhabditis elegans HOLOCENTRIC CHROMOSOMES.**

In most eukaryotes each chromosome contains one centromere (monocentric) whose location does not appear to change from one division to the next. However, nematodes and crayfish are holocentric and their kinetochore is formed along the entire length of chromosomes (Pimpinelli & Goday, 1989).

Holocentric chromosomes bind to microtubules along their entire length and during mitosis move broadside to the pole from the metaphase plate. The most studied holocentric chromosome is from *C. elegans*. A lot of the proteins that bind to *C. elegans* kinetochores are homologues to components identified in monocentric organisms. In *C. elegans* the HCP-3 (CENP-A homolog) antibody staining shows a punctuate pattern within interphase nuclei and bands along mitotic chromosomes, indicating that not all genomic sequences are associated with this centromeric protein (reviewed by Dernburg, 2001). *C. elegans* holocentric chromosomes might contain more repeat units at multiple sites, covering most of the chromosome (Sullivan et al., 2001). It seems that specific DNA sequences are not required for nematodes centromere: concatamers of phage lambda DNA and many other types of DNA can be stably transmitted (Stinchcomb et al., 1985).

**MAMMALIAN CENTROMERES.**

Human centromeres are composed of alpha-satellite DNA, a 171 bp monomer tandemly arranged into higher-order arrays that extend for 100 kb to several megabases (Willard, 1998). Not all this region is required to maintain functional human centromeres: functional kinetochores have been made where much of this array was deleted (Mills et al., 1999; Yang et al., 2000). However there must be some peculiarities in this arrays (sequence or factors bound), since not all alpha-satellite arrays in human artificial chromosomes form centromeres *de novo* (Harrington et al., 1997; Ikeno et al., 1998). The kinetochore protein CENP-A *in vivo* associates alpha-satellite repeats by ChIP (Vafa & Sullivan, 1997). However, not all the entire alpha-satellite is involved in kinetochore assembly because antibodies to kinetochore proteins such as CENP-A localized to only a portion of the alpha satellite DNA (Warburton et al., 1997). Moreover, chromosomes that are deleted for much of this array can still assemble a kinetochore and segregate normally (Wevrick et al., 1990; Yang et al., 2000). Alpha satellite arrays appear to be mostly homogeneous, but some interspersed sequences such as LINEs, Alu repeats and other satellites are also present; however these sequences are also present in other parts of the genome that never associate with kinetochore, suggesting that they themselves are not sufficient
for centromere function (Lee et al., 1997). Transfection of naked alpha satellite DNA can result in *de novo* forms of centromeres that bind the active kinetochore proteins CENP-C and CENP-E to create mammalian artificial chromosomes (MAC) (Harrington et al., 1997; Ikeno et al., 1998). Usually the stable MACs were substantially larger than the input DNA and the nature of events that led to these rearrangements are not known. This suggests that additional epigenetic processes are necessary to allow centromere formation and mitotic propagation.

Finally, as discussed below, identical alphoid sequences can differ in their ability to bind active kinetochore proteins like in dicentric chromosomes and centromeres can assemble also in the absence of alphoid DNA.

**CENTROMERE INACTIVATION AND NEOCENTROMERE FORMATION.**

Differently from budding yeast, the simple presence of centromeric DNA in a cell or on a chromosome does not correlate with centromere function in metazoans. For example, dicentric chromosomes exist and both are able to function as centromeres (Sullivan & Willard, 1998; Higgins et al., 1999). Moreover, in human and flies there must be a mechanism for functional inactivation of one centromere in presence of dicentric chromosomes that can be in fact stably transmitted (Sullivan et al., 1995; Warburton et al., 1997; Agudo et al., 2000).

The idea that something else other than the simple sequence specifies where kinetochore assemblies is further supported by examples where centromere function is exhibited by noncentromeric DNA. Rearranged human marker chromosomes, the best studied of which is the 10q25 region of a rearranged marred (10) chromosome, have been identified that contain no detectable alphoid-containing regions (Depinet et al., 1997; Du Sart et al., 1997; Williams et al., 1998; Warburton, 2000). They are functionally identical to normal centromeres in mitotic activity and their binding to essential centromeric and pericentromeric proteins (Aagaard et al., 2000; Saffery et al., 2000). Most reported human neocentromeres originate from regions of euchromatin (Choo, 2001; Warburton et al., 2000). The absence of any difference in DNA sequence between the precursor 10q25 genomic region and the neocentromere implies an epigenetic process of neocentromerization (Choo, 2000). By Chromatin immunoprecipitation and genomic DNA array (CIA) analysis, a binding domain of 330 kb has been defined for the centromere-specific histone H3 homolog CENP-A at the 10q25 neocentromere (Lo et al., 2001). Similar results have been obtained by irradiation of *Drosophila Dp1187* minichromosome: acentric derivatives lack the collection of satellites and transposable elements associated with the *Dp1187* centromere and are composed of only 225-290 Kb of subtelomeric heterochromatin and euchromatin (Murphy & Karpen, 1995). *Dp1187* acentric derivative are transmitted by a microtubule-based mechanism and exhibit 'neocentromere' activity.

How can noncentromeric DNA acquire and propagate centromeric function in *Drosophila* and humans? Neocentromere activation in flies and other eukaryotes might occur by the spreading of centromere function or proteins from the 'normal' centromere into adjacent regions (Allshire & Karpen, 1997; Maggert & Karpen, 2001). Neocentromeric activation cannot be a frequent
occurrence, otherwise monocentric chromosomes would become di- or multicentric and would be lost or become holocentric. This indicates that centromere identity might be established and propagated at a specific and consistent chromosomal site by epigenetic mechanisms (Karpen & Allshire, 1997; Murphy & Karpen, 1998).

**Arabidopsis thaliana CENTROMERES.**

*Arabidopsis thaliana* centromeres are localized to 500-1000 kb regions that contain 180-bp repeats, bounded by ribosomal DNA arrays and other repetitive sequences (Murata et al., 1994; Round et al., 1997; Copenhaver et al., 1999; Heslop-Harrison et al., 1999). Their localization have been done by genetic mapping and recombination suppression rather than kinetochore function: it could be that their size is overstimated. Recent studies reveal that plant centromeric regions consist of satellite DNA and transposable elements (Gindullis et al., 2001; Zhong et al., 2002).

**COMPONENTS OF FISSION YEAST CENTROMERES.**

Fission yeast centromeres are composed of two distinct chromatin domains. The central domain where the kinetochore is assembled, is occupied by the histone H3-like protein CENP-A_Cnp1. In contrast, the flanking repeats are coated with silent chromatin, characterized by H3K9me2. This modification, induced by the methyltransferase Clr4 and somehow by the action of RNAi pathway on non-coding centromeirc transcripts, allows the subsequent binding of Swi6 (HP1). Their structure is similar to that of metazoan centromeres where the kinetochore is surrounded by heterochromatin. The flanking heterochromatin occupied mainly by the chromodomains protein Swi6, is important for the recruitment of cohesin that ensures tight physical cohesion between sister centromeres (Bernard et al., 2001). Like their more complex metazoan counterparts, fission yeast kinetochores interact with multiple microtubules. Thus fission yeast represents an excellent model system in which to study centromere structure and function thanks to its genetic tractability and complex structure. In this section I will focus on two aspects of fission yeast centromeres:
- their domain organization and protein composition,
- the role of histone modifications and the RNAi machinery in their function.

**CENTROMERE STRUCTURE.**

Fission yeast has only three chromosomes. Centromere I (*cen1*) occupies 35 kb and it is on the largest chromosome (5.6 Mb), *cen2* is 65 kb (4.4 Mb) and *cen3* is 110 kb (2.5 Mb) (Takahashi et al., 1992; Steiner et al., 1993; Wood et al., 2002). The central cores (*cnt*) of all the three centromeres are made of 4-7 kb of non-repetitive sequence. *cen1* and *cen3* share a 3.3 kb element *'tm'* which is 99% identical between the two centromeres, *cen2* has a ~1.5 kb element in
its central core that is 48% identical to 'tm' (Wood et al., 2002). Each cnt is surrounded by inverted 'innermost' (imr) repeats that are unique to each centromere. Thus the central core and the inner part of inverted 'innermost' repeats constitute the central domain, that as previously described, is structurally and functionally distinct from the outer repeats (Partridge et al., 2000) and it is where the kinetochore assemblies. The outer repeat regions are composed of 'dg' and 'dh' elements (also known as K and L repeats) and part of the 'innermost' repeats. Between the different centromeres, the dg repeats are 97% identical, whilst dh are 48% identical and they share some homology also with elements at the mating type locus and telomeres (Grewal & Klar, 1997). The difference length of the three centromeres, results from the number of outer repeats elements that they content. Moreover the organization and orientation of the repetitive elements relative to the central core at each centromere varies.

The central core DNA is essential for centromere function but is not active on its own. Nevertheless all the central core with limited inverted repeat sequences is enough for centromere function: the minimal requirement for a functional S. pombe centromere is the central core sequences, together with at least the K-type repeats (Clarke et al., 1986; Chikashige et al., 1989; Carbon & Clarke, 1990; Clarke & Baum, 1990; Hahnenberger et al., 1991; Baum et al., 1994).

A small cluster of tRNA genes located within the imr repeats and within the central core-associated repeats and the central core itself of cen3 (Kuhn et al., 1991; Takahashi et al., 1991, 1992) take part to the composition of the three centromeres. It is not clear their function and the explanations for their presence in such a particular genomic locus are different. It is possible that they may be involved in centromere function such as forming a nucleosome-disrupting transition zone (barrier) between the central core (which lacks normal nucleosomes) and sequences within the inverted repeats (which are packed into standard nucleosomal arrays) (Figure 1-3).

EPIGENETIC FEATURES OF FISSION YEAST CENTROMERES.

Fission yeast centromeres are epigenetically regulated (Karpen & Allshire, 1997; Sullivan et al., 2001; Amoe & Choo, 2002). This is suggested by different observations. First, marker genes inserted at the outer repeats, are silenced but variegation can occur: for example, the ade6+ gene inserted at the central core produces red (repressed state), white (expressed state), sectored and pink colonies on indicator plates (Allshire et al., 1994) indicating that distinct metastable states can be adopted by normal centromeres and propagated in genetically identical cells.

Another evidence of fission yeast centromere being epigenetically regulated comes from the following experiment. The outer repeats of fission yeast centromeres are associated with nucleosomes that are underacetylated on histone H3 and H4 (Ekwall et al., 1997; Mellone et al., 2003). Under transient treatment of cells with the histone deacetylase inhibitor TSA, these cells
Figure 1-3 Fission yeast centromere organization.

*S. pombe* is organized into 3 chromosomes 1, 2, and 3. The largest chromosome (Ch1) contains the smallest centromere (35 Kb) whereas the smallest chromosome (Ch3) contains the largest centromere (110 Kb). Centromere 2 occupies 60 Kb. The outer repeats (otr) are composed by *dg* (gray arrow) and *dh* (cross hatched arrow) and a small part of the innermost repeats. *cen3* contains 9 additional copies represented by x3 on the left and x8 on the right side. The central core domain (cc) of *cen1* and *cen3* shares an element, TM/cnt (checked rectangle). Inner repeats at all three centromeres have a distinct sequence represented by the differently shaped arrowheads, *imr1* (white) *imr2* (stippled) *imr3* (striped) and most of their sequences are part of the central core domain. The multiple tRNA genes are denoted by white boxes and are thought to mediate the transition zone between the central core domain and the outer repeat domain (Allshire, 2001).
gain a heritable increase in acetylation of outer repeat histones, lose marker gene silencing and become defective in centromere function (Ekwall et al., 1997). However, at a low frequency these defective acetylated lineages are able to regain the functional state. Moreover, when transformation is performed into S. pombe of minimal centromere constructs, particular plasmids can adopt ‘stable’ or ‘unstable’ states (Steiner & Clarke, 1994). The difference between the two states is neither due to mutation and rearrangement of DNA sequences, nor to recovery and passage of the ‘stable’ and ‘unstable’ form upon retransformation into S. pombe. This suggests that, naked plasmids transformed into fission yeast, stochastically can adopt a functional centromere conformation, possibly through the formation of a specific chromatin architecture. Moreover if plasmid with ‘non-functional’ centromeres are propagated selectively in S. pombe for many generations, sometimes they can gain centromere activity. Together all these studies support the idea that fission yeast centromeres are subject to epigenetic regulation, with the propagation of meta-stable chromatin states without changes in DNA sequence. These events may relate to neocentromere formation in humans.

**DOMAIN STRUCTURE OF FISSION YEAST CENTROMERE.**

The idea of an organization of fission yeast centromere in two distinct domains also with different functions is supported by cytological analysis. During interphase, S. pombe centromeres are clustered at the nuclear periphery, adjacent to the spindle pole body (SPB) (Funabiki et al., 1993) but by electron microscopy analysis, the central core/kinetochore and heterochromatin domains are cytologically distinct (Kniola et al., 2001). Also in mitotic cells, cytological analysis revealed two distinguishable domains and supported the idea that the kinetochore/central core and heterochromatin domains have distinct functions (Appelgren et al., 2003).

The proteins that bind to these two different domains, are different and mutations in their genes display also completely different phenotypes: conditional mutants that affect the central core chromatin structure, such as mis6, cnp1CENPA and sim4, display severe mitotic phenotypes and their null mutants are lethal (Saitoh et al., 1997; Takahashi et al., 2000; Pidoux et al., 2003) whereas, apart from a small class of mutants, the csp ts that compromise heterochromatin structure and are essential genes (Ekwall et al., 1999), null mutants that compromise the flanking heterochromatin, such as clr4, rik1 and swi6 are viable. However they display lagging chromosome in anaphase and sensitivity to microtubule inhibitors due to a sister centromere cohesion defect due to loss of Swi6 (Allshire et al., 1995; Ekwall et al., 1995; Ekwall et al., 1996). Thus the centromere in fission yeast is divided in two domains: the kinetochore, which is anchored via the central core region and which mediates interactions with spindle microtubules and the flanking heterochromatin domain with a specific functions in sister centromere cohesion. Both domains are epigenetically regulated.
Also the pattern of protein binding established by Chromatin immunoprecipitation (ChIP) analysis indicates that the fission yeast centromeres contain two distinct protein domains. The central domain associated with the histone H3 variant Cnp1\textsuperscript{CENP-A}, Mis6, Mis12, Dis1, Mtc1/Alp14, Mal2 and Sim4 (Saitoh et al., 1997; Takahashi et al., 2000; Garcia et al., 2001; Nakaseko et al., 2001; Jin et al., 2002, Pidoux et al., 2003). The outer repeat domain is occupied by the chromodomain proteins, Swi6, Chp1, the CENP-B homolog Abp1, and the passenger protein Bir1 (Partridge et al., 2000; Morishita et al., 2001; Nakagawa et al., 2002). The set of proteins that binds to the outer repeats does not bind to the central core and vice versa (Figure 1-4).

CENTRAL CORE CHROMATIN AND THE KINETOCHORE.

A unique chromatin structure characterized the central core where the kinetochore assembled: upon micrococcal nuclease digestion, the central core produces a smear pattern (Polizzi & Clarke, 1991; Takahashi et al., 1992), whilst all other regions of the genome, including the outer repeats, produce a typical nucleosome ladder pattern. It is not known what this smear pattern means at the chromatin level. However it is not simply due to the primary DNA sequence as central core sequences assemble into ordinary nucleosomal arrays when introduced into budding yeast (Polizzi & Clarke, 1991). The unique structure also correlates with function: only functional centromeric constructs assemble this chromatin (Marschall & Clarke, 1995). Another peculiarity of the central core chromatin compared to the rest of the genome is the presence of the histone H3-variant CENP-A\textsuperscript{Cnp3} that replaces histone H3 (Takahashi et al., 2000).

Marker genes inserted at the central core are transcriptionally silenced: the assembly of a functional kinetochore complex ensures silencing. This has been a useful tool to search for novel factors important for the kinetochore structure and architecture. In fact, despite the level of transcriptional silencing is very modest, in a screen performed using strains with \textit{arg3}\textsuperscript{+} marker gene driven by a crippled promoter inserted at the central core, the Sim (silencing in the middle of the centromere) mutants, CENP-A\textsuperscript{Cnp1} itself and the kinetochore protein Sim4 which is probably the homologue of CENP-H were isolated (Pidoux et al., 2003). Other central core associated proteins, like Mal2, Mis6 and Mis12 were identified in minichromosome loss screens (Fleig et al., 1996; Takahashi et al., 1994). All these three proteins associated with the central core region by ChIP (Saitoh et al., 1996; Goshima et al., 1999; Jin et al., 2002) and both \textit{mal2} and \textit{mis6} mutants alleviate central core silencing (Partridge et al., 2000; Jin et al., 2002). Moreover \textit{mis6}, \textit{mis12}, \textit{cnp1}, \textit{mal2}, \textit{sim4} and \textit{ams2} mutants alter the unique chromatin structure of the central core: the smear pattern is lost and replaced by a ladder pattern typical of bulk chromatin (Saitoh et al., 1996; Goshima et al., 1999; Jin et al., 2002; Chen et al., 2003; Pidoux et al., 2003).

It is still unknown how CENP-A\textsuperscript{Cnp1} nucleosome are incorporated only at the central core. This is not only due to a particular DNA sequence: in fact a mutation in a protein that associates with the central core, Mis6, fail to incorporate newly synthesized GFP-Cnp1, suggesting that Mis6 could be a possible loading factor for CENP-A\textsuperscript{Cnp1} (Takahashi et al., 2000). However, a similar
Figure 1-4 Features of fission yeast chromatin at centromere 1. The centromeric chromatin is composed of two defined regions: a central domain (central core and central core associated repeats or imr) with which Mis6 and Mis12 are associated and Cnpl localizes and flanking outer repeats, coated with Swi6 and Chpl. Outer repeat elements contain underacetylated histone H3 and H4 N-tails, histone H3 methylated on lysine9 by C1r4 that determines the recruitment of Swi6.
mechanism does not seem to be conserved since neither the mammalian Mis6 homologue CENP-I nor the budding yeast homologue Ctf3, are required for centromere localization of their respective Cnp1 counterpart, CENP-A and Cse4p (Measday et al., 2002; Nishihashi et al., 2002). Probably different mechanisms contribute in chromatin assembly in different species. Another model to explain the specific central core nucleosome CENP-A deposition could depend on the presence of a functional kinetochore. A functional kinetochore interacts productively with spindle microtubules and segregate chromosomes. One possibility is that tension across the sister kinetochores produced by correct bi-orientation of the chromosome on the spindle, represents a mark for the deposition of CENP-A (Mellone & Allshire, 2003).

COMPONENTS OF CENTROMERIC HETEROCHROMATIN.

S. pombe centromeres are heterochromatic and genes placed within them become variably, but reversibly, repressed. This repressive effect at S. pombe centromeres resembles the classical phenomenon of position effect variegation imposed by Drosophila heterochromatin on nearby genes and it occurs also at the heterochromatic sites at the telomeres and the mating type loci. Silencing of genes placed in different parts of the centromere is very different: the outer regions are more strongly silenced than the central core. Insertion of the marker genes ura4+ and ade6+ within these sequences results in reversible repression of gene expression (Allshire et al., 1994). This system is very useful to screen for proteins involved in centromere structure and function: mutations that alleviate centromere induced gene silencing also interfere with chromosome segregation, providing a link between chromatin structure and centromere function (Allshire et al., 1995; Ekwall et al., 1995; Ekwall et al., 1996).

Using this system, mutations in clr4, rik1 (Ekwall and Ruusula, 1994; Ivanova et al., 1998) and swi6 genes (Lorentz et al., 1994) were found to affect transcriptional silencing of insertions across the outer repeats, but not across the central domain. In particular, the alleviation of silencing caused by mutations in clr4 and rik1 is more pronounced than in swi6 mutants (Allshire et al., 1995). clr4, rik1 and swi6 were originally identified as mutations that alleviate repression of two other heterochromatic loci in fission yeast, silent mating type loci mat2 and mat3 (Lorentz et al., 1992; Thon et Klar, 1992; Ekwall and Ruusula, 1994) and telomeres (Allshire et al., 1995).

Swi6

Swi6 protein contains a chromodomain, a motif first identified as a region of similarity between HP1 and Polycomb (Paro and Hogness, 1991), that is 46% identical to the chromodomain of Drosophila HP1 (Lorentz et al., 1994). Like HP1, Swi6 contains a chromo-shadow domain, which is responsible for the dimerization (Cowieson et al., 2000). Swi6 localizes at centromere, mating type loci and telomeres as shown by immunostaining and FISH (Ekwall et al., 1995) forming 2-5 discrete spots per haploid nucleus. Cells lacking Swi6 protein have increased frequency of lagging chromosomes in late mitotic cells and of minichromosome loss (Ekwall et al., 1995). Moreover its localization is dependent upon the presence of intact clr4 and rik1 gene products (Ekwall et al., 1996). Mutated clr4, rik1 and swi6 show cold sensitivity at 18° C and an increased
frequency of lagging chromosomes. They all are also sensitive to the microtubule poison TBZ and display genetic interaction with nda3, (α-tubulin).

Clr4

Clr4, as previously mentioned, is the fission yeast homologue of Drosophila Su(var)3-9, human SUV39H1 and murine Suv39h1. It contains a chromodomain and a SET domain, both essential for silencing (Ivanova et al., 1998). The SET domain of Clr4 displays histone H3 K9-specific methyltransferase activity in vitro (HMTase) (Nakayama et al., 2001) Methylated histone H3 K9 provides a binding site for Swi6 (Rea et al., 2000; Bannister et al., 2001) with a similar mechanism to that occurring at mammalian heterochromatin (Lachner et al., 2001). ChIP experiments, using a specific antibody against H3K9me2, showed the presence of this modification across the outer repeats and mating type loci and it is lost in clr4Δ cells (Nakayama et al., 2001).

Riki

Little is known about the Riki protein. Riki encodes a WD propeller repeat protein related to the DNA damage binding protein, DDB1 (Tuzon et al., 2004). DDB1 is ubiquitinated by Cullin 4A in response to the signalosome and mediates DNA repair in animals (Wittschieben and Wood, 2003) as well as gene repression in plants (Schroeder et al., 2002; Yanagawa et al., 2004). The WD propeller repeat domain is present in the large subunit of the cleavage and polyadenylation specific factors (CPSF) (Neuwald and Poleksic, 2000). Riki associates with centromeric repeats and interacts with Clr4 (Sadaie et al., 2004). Cells lacking Riki do not contain H3K9me2 at the centromeric outer repeats and at the mating type loci (Nakayama et al., 2001) indicating that Riki might act upstream of Clr4. In addition, a Riki associated cullin-dependent E3 ubiquitin ligase is essential for heterochromatin formation (Horn et al., 2005). Its precise role has not been defined yet.

Chp1 and Chp2

Chp1 is a protein that contains a chromo-shadow domain and is a component of fission yeast centromeric outer repeats. In its absence cells have chromosomes lagging on the spindle during anaphase and a high rate of chromosome loss (Doe et al., 1998). It is required for outer repeat transcriptional silencing but not to repress the silent mating type (Partridge et al., 2000; Thon and Verhein-Hansen, 2000; Doe et al., 1998). It binds H3K9me2 (Partridge et al., 2002) and it has been found recently as part of the RITS complex, together with Ago1 and an uncharacterized protein Tas3 (Verdel et al., 2004). In fact it is known that the RNAi pathway is involved in the establishment of silent chromatin at the outer repeats and at the mating type locus and siRNAs complementary to the outer repeats have been found in fission yeast cells (Volpe et al., 2003; Bartel & Reinhert, 2003; Hall et al., 2002).

Chp2 shares extensive sequence similarity with Swi6 and contains both a chromo- and a chromoshadow domain. chp2Δ cells have a weak alleviation of silencing at centromere, the mating type region and telomeres (Halverson et al., 2000; Thon and Verhein-Hansen, 2000).
Chp2 is required like Swi6 for the maintenance of H3K9 methylation and it is a structural component of heterochromatin (Sadaie et al., 2004).

**Abp1, Cbh1 and Cbh2**

Abp1, Cbh1 and Cbh2 are three fission yeast proteins with homology to the human centromere protein CENP-B that binds α-satellite DNA (Murakami et al., 1996; Lee et al., 1997; Irelan et al., 2001; Baum & Clarke, 2000). Alleviation of outer repeats silencing, reduction of H3K9 methylation and Swi6 binding characterized cells lacking combinations of these genes (Nakagawa et al., 2002). In addition, Abp1 is associated with the outer repeats. It has been proposed that they may act as site-specific nucleation factors for heterochromatin formation in the centromeric flanking domain (Nakagawa et al., 2002). Similarly, the CREB related transcription factor Atf1 represents a specific nucleation factor at the mating type loci (Jia et al., 2004). Hence Abp1 and Cbh1 may act upstream of the HDAC and the HMTase to generate the H3K9me epigenetic mark.

**Hip1 and Slm9**

In addition, proteins involved in chromatin assembly and nucleosome deposition affect centromere function: mutants in the *hip1* (HIRA-like involved in assembly of repressive chromatin) and *slm9* (similar to His2, a histone gene regulator in *S. cerevisiae*, Kanoh et al., 2003) genes partially alleviate silencing in the centromere outer repeats (Blackwell et al., 2004).

**Epel**

Epel is a jumonji-domain protein that when deleted, allows spreading of silencing beyond the normal heterochromatin barriers at the mating type locus. It affects silencing also at the centromeric outer repeats. Its role may be involved in influencing the spreading of silent chromatin at the heterochromatic loci in fission yeast (Ayoub et al., 2003). Recently, it has been proposed a role for Epel as a putative histone demethylase that could act by oxidative demethylation or a hydroxylase that catalizes a novel histone modification (Trewick et al., 2005).

**A role for heterochromatin in centromere function.**

It is known that both central core and outer repeat regions are required to form a functional centromere. However, only the genes encoding central core proteins are essential, while all the other outer repeat proteins are non-essential. This prompts us to think that the outer repeats must have a particular role possibly in directing properly central core components and ensuring the assembly of a functional kinetochore.

In fact, in fission yeast, an important link between heterochromatin and centromere cohesion has been demonstrated. Sister-chromatids cohesion is mediated by the cohesin complex, which is distributed along the entire chromosome with enrichment at the centromeres. At anaphase Scc1/Rad21 subunit of cohesin is cleaved and this allows sister-chromatids separation. In metazoa, arm cohesion is released early in mitosis and the sister-chromatids remain associated through the centromeric pool of cohesion (Cohen-Fix, 2001). It is not known what distinguishes arm cohesion from centromere cohesion. In fission yeast, the chromo-domain HP1 homolog...
protein, Swi6 is required for recruitment of Rad21 indicating that centromeric heterochromatin is important for centromere cohesion (Bernard et al. 2001; Nonaka et al., 2002). Cells lacking Swi6 show loss of centromeric cohesion but they maintain arm cohesion. The implication is that high frequency of lagging chromosomes is explained by the depletion of centromeric cohesion in mutants that are necessary for heterochromatin formation (Bernard et al., 2001). Swi6 has been shown to interact physically with another subunit of cohesin, Psc3 (Nonaka et al., 2002). It is possible that the outer repeats are required initially in the establishment of a functional centromere to specify where to deposit CENP-A for the assembly of the kinetochore. Then the essential role of the outer repeats may be over and they just function to function to recruit cohesions (Pidoux et al., 2004).

It is to consider also that the kinetochore is formed over the central core chromatin that is where microtubules contact the centromere; indeed microtubule-associated protein (MAP), Dis1 localized to this region by chip. However, another MAP, Alp14, is located at the imr and otr regions (Garcia et al., 2001; Nakaseko et al., 2001).

Despite centromere function in fission yeast is not only dependent on a particular DNA sequence, portions of the outer repeat can be placed in a euchromatic location and they induce ectopic silencing (Partridge et al., 2002). A 1.6 kb portion of the outer repeat inserted next to an ura4+ gene at a euchromatic site, becomes silenced and packaged into silent chromatin with H3K9me2, association of Swi6 and Chp1 and recruitment of Rad21-cohesin. A 1-kb fragment can drive ectopic silencing and this indicates that certain DNA sequences or some features of them like the ability to produce dsRNAs, can direct heterochromatin formation. The same observations have been made for a short sequence from the silent mating type locus (Ayoub et al., 2000).

**HISTONE MODIFICATION AT S. pombe CENTROMERES.**

Centromeres are generally underacetylated on histone H3 and H4 in S. pombe (Ekwall et al., 1997; Mellone et al., 2003). When S. pombe cells are treated with the histone deacetylase inhibitor Trichostatin A (TSA) centromeric heterochromatin becomes hyperacetylated and centromere function is defective (Ekwall et al., 1997). The underacetylated states of H3 on K9, K14, K18, K23, K27 and of H4 on K5, K8, K12, K16 at fission yeast centromeres have been demonstrated by ChIP analysis with specific antibodies for each of this acetylated K by Mellone et al., 2003.

Histone deacetylase plays a pivotal role in controlling correct silent chromatin formation in fission yeast. Three different HDACs, Clr6, Clr3 and Sir2 are required in heterochromatin formation at imr and otr regions. They all have different function in controlling the underacetylated silent chromatin state. Clr6 together with Pst1 (pombe sir three corepressor) forms a complex that is required to propagate underacetylated centromeres (Silverstein et al., 2003). Clr3 has been shown to affect the H3K9 methylation in the otr region (Nakayama et al., 2001). And recently
Sir2, has been shown to be required to keep the hypoacetylated state of the imr region more than the otr one (Shankaranarayana et al., 2003). Despite the fact that both H3 and H4 are underacetylated at centromeres, only mutations mimicking the acetylated state introduced at positions in H3 tails, in particular H3K14, and not of H4 influence and disrupt centromeric heterochromatin integrity and function (Mellone et al., 2003).

Swi6 localizes to centromeric heterochromatin in fission yeast, and this localization has been shown to require the histone methyltransferase C1r4 (Ekwall et al., 1996) that methylates H3K9 in vitro (Rea et al., 2000). H3K9me2 in the otr region is also dependent on the action of the C1r3 (Nakayama et al., 2001) and Sir2 (Shankaranarayana et al., 2003). Since C1r3 acts on H3K14Ac (Bjerling et al., 2002) it is possible that deacetylation of H3K14Ac by C1r3 and H3K9Ac by Sir2 are both required for the subsequent action of C1r4 on H3K9. In fact H3K14 mutated to both alanine (A) and arginine (R), causes alleviation of silencing at the centromere: K14 as well as K9 of histone H3 are required for heterochromatin formation (Mellone et al., 2003). In addition, cells mutated in the histone deacetylase C1r3 have slightly reduced methylated K9 at centromeric repeats, which decreases even more in cells mutated both in C1r3 and C1r4 (Nakayama et al., 2001).

Another covalent modification occurs at centromeric histones in S. pombe. Histone H3 is phosphorylated at serine 10 by the Aurora kinase Ark1 (Petersen et al., 2001). Ark1 interacts with the flanking centromere protein Bir1 (Leverson et al., 2002). Like the passenger proteins in other eukaryotes, the S. pombe Survivin homologue Bir1 and the associated Aurora kinase are only transiently bound to the centromere region and move away to the spindle midzone at anaphase (Morishita et al., 2001). bir1 mutants display lagging chromosomes and a cytokinesis defect (Rajagopalan & Balasubramanian, 2002). Ark1 is required for chromosome condensation and for the spindle checkpoint response to unattached kinetochores (Petersen & Hagan, 2003). H3S10A mutant cells show numerous mitotic defects like lagging chromosome on anaphase spindle, sensitivity to microtubule destabilizing drugs, increased minichromosome loss rates (Mellone et al., 2003).

The fact that mutants in histone deacetylation, methylation and phosphorylation, as well as point mutations of histone H3K14, H3K9 and H3S10 affect silencing in the outer centromeric repeats, is consistent with the idea that such modifications are part of a 'code' for silent and active chromatin (Jenuwein & Allis, 2001).

**RNA INTERFERENCE AND CENTROMERE FUNCTION IN S. pombe.**

RNA interference (RNAi) can act as a post-transcriptional process responsible for 'post-transcriptional gene silencing' (PTGS), in which small (21-23 nucleotide) inhibitory RNA molecules trigger degradation of homologous mRNA by Dicer (an RNAse III-like enzyme) and the RNA-induced silencing complex (RISC) (Fire et al., 1998; Zamore et al., 2000; Hammond et al., 2000; Bernstein et al., 2001). In addition, RNAi can act also at the transcriptional level (TGS)
and direct silent chromatin formation through DNA or histone modifications (Matzke et al., 2005; Grewal and Moazed, 2003; Lippman and Martienssen, 2004). The role of RNAi in directing silent chromatin formation is conserved in plants, *Drosophila*, mammals (Wassenegger, 2005). Studies in fission yeast have shown a role for RNAi in guiding heterochromatic silencing and histone modification at centromere and mating type loci (Volpe et al., 2002; Hall et al., 2002). Moreover as in *S. pombe*, centromeric transcripts have been detected from mouse and human satellite repeats (Lenhertz et al., 2003; Saffery et al., 2003) and RNA has a role in the higher order structure of pericentromeric heterochromatin (Maison et al., 2002). The connection between chromatin histone modifications and RNAi has been extended to more complex organisms like *Arabidopsis* (Ziberman et al., 2003; Melquist & Bender, 2003), *Drosophila* (Pal-Bhadra et al., 2004) and vertebrate cells (Fukagawa et al., 2004; Morris et al., 2004).

In *S. pombe* only one copy of each of the genes required for RNAi is present and this represents a very good system to genetically study RNAi (Wood et al., 2002). Dicer (Dcr1), the RISC component Argonaute (Ago1), the RNA-dependent polymerase Rdpl and several uncharacterized Csp gene products (centromere suppressor of position effect; Ekwall et al., 1999) are all required for heterochromatin formation and RNAi-mediated transcriptional silencing of centromeres (Volpe et al., 2002, 2003; Provost et al., 2002).

Transcript analysis revealed that the centromere otr repeats are transcribed from both strands but the transcripts can be detected only in the RNAi mutants: they are rapidly processed into small interfering RNAs (siRNAs) in wild type cells (Volpe et al., 2002). The reverse strand is continually transcribed whereas the forward strand is transcriptionally silenced. Indeed small interfering RNAs (siRNAs) from *S. pombe* have been cloned and found to be homologous to the centromeric repeat sequences (Reinhart & Bartel, 2002).

RNAi pathway mutants affect silencing at the outer repeats and not at the central core of centromere (Volpe et al., 2002). Silent chromatin at the outer repeats depends on H3K9me2 and Swi6 binding. Reporter genes inserted in the outer repeats are derepressed and lose H3K9me2 and Swi6 binding in the RNAi mutants, however, the centromeric repeats themselves just decreased few fold these heterochromatic marks (H3K9me2 and Swi6 binding) compared to wild (Volpe et al., 2002; Sadaie et al., 2004). This suggests that there must be an independent RNAi pathway to establish some silent chromatin nucleation (H3K9me2 and Swi6 binding).

In fact the RNAi pathway contributes to the establishment but not the maintenance of heterochromatin at the *mat2/3* mating type locus. This region contains the *cenH* DNA that shares 96% homology with a portion of the *dg/dh* centromeric repeats. At this genomic locus, another parallel pathway to the RNAi, initiates heterochromatin assembly: the two stress-response transcription factors, Atf1 and Pcr1 (Jia et al., 2004). Thus heterochromatin assembly can occur both in a RNAi-dependent and RNAi-independent manner. Something similar happens at the telomeres, where in RNAi mutants, the usual clustering of telomeres during meiosis is impaired (Hall et al., 2003). In addition some subtelomeric genes become more expressed in RNAi mutants (Mandell et al., 2005). However, like for the mating type locus, the hallmarks for
heterochromatin (H3K9me2, Swi6 binding) are not lost, suggesting a pathway RNAi-independent to initiate heterochromatin assembly at telomeres (Petrie et al., 2005).

In conclusion, heterochromatin at centromere, mating type and telomere regions of fission yeast is initiated by the RNAi but, only centromere heterochromatin depends on the RNAi pathway to be maintained.

From a recent study by the Grewal lab, the highly repetitive ribosomal RNA genes are also target of RNAi-directed H3K9 methylation (Cam et al., 2005).

The RITS complex initiates heterochromatin assembly through siRNAs.

The recently characterized RITS (RNA-induced initiation of transcriptional silencing) complex provides a direct link between the siRNA produced by Dicer and heterochromatin (Verdel et al., 2004). RITS contains both a known chromo-domain protein Chp1 (which binds the centromere outer repeats; Partridge et al., 2000), Ago1, as well as an uncharacterized protein, Tas3, and siRNAs that are complementary to centromeric DNA repeats where heterochromatin assembly is initiated (Verdel et al., 2004). RITS has been proposed to mediate targeting of specific regions of chromatin for heterochromatin formation: in dcr1Δ cells, siRNAs are absent, RITS does not localize to centromeric repeats and no silent chromatin forms (Verdel et al., 2004; Motamedi et al., 2004). In RNAi mutants H3K9me2 and Swi6 binding over the outer repeats are not completely lost like they are on a reporter genes: it is conceivable that RITS might promote the spreading from centromeric repeats to reporter genes integrated nearby (Sadaie et al., 2004).

However RITS has a role in heterochromatin initiation that is strictly linked to and dependent on H3K9me2, and thus Clr4. A possible model proposes that recognition of H3K9me2 by Chp1 might promote RNAi by recruiting Ago1 through the RITS complex. Failure to recruit the RNAi machinery in clr4Δ cells would then result in loss of siRNA (Martienssen et al., 2005).

Rdp1 and RDRC.

There are two possible ways to produce dsRNA that then are diced into siRNAs: centromeric transcripts are transcribed in sense and antisense RNA from both DNA strands or a complementary RNA is synthesized from a RNA template by an RNA-Directed RNA polymerase (RDRP).

Fission yeast has one RDRP, Rdp1 that is recruited to the centromere and by ChIP analysis binds to the outer repeats (Volpe et al., 2002). Recently biochemical data demonstrate that Rdp1 recruitment to centromere is Dcr1 and Ago1 dependent, thus RITS dependent (Motamedi et al., 2004). Moreover, Rdp1 is in a complex known as RDRC that contains two additional proteins, Hrr1, an RNA helicase and Cid12, a protein related to poly (A) polymerase. In the absence of Rdp1, the complex falls apart and Cid12 is found in association with the spliceosome instead (Motamedi et al., 2004).

It is not clear how the RDRC complex could interact and recruit the histone modification apparatus, thus Clr4. Data from the Grewal and Moazed laboratories, suggest the existence of a
Figure 1-5 Centromeric siRNAs direct heterochromatin formation. Centromeric repeats are transcribed by RNA pol II (Schramke et al., 2005; Kato et al., 2005) and dsRNAs are produced probably through the action of Rdpl. Dcr1 cleaves this long dsRNAs into 21-22 nt length siRNAs that target the RITS complex to centromeric nascent transcripts and H3K9 methylation over centromeric DNA (Sugiyama et al., 2005). RDRC is recruited, amplifies the signal through the production of more siRNAs to be incorporated in the RITS complex (Motamedi et al., 2004).
self-enforcing loop where RITS, tethered to chromatin through the binding of the chromo-domain protein Chop1 to H3K9me2, recruits RNAi factors such as Rdp1 to heterochromatic loci (Noma et al., 2004; Motamedi et al., 2004; Sugiyama et al., 2005). This enables the RNAi machinery to function in cis as a stable component of heterochromatin to process centromeric transcripts into siRNAs, allowing the further recruitment of heterochromatin and RNAi factors.

A study about the distribution of H3K4me2 and H3K9me2 modification and of RNAi components across the genome, revealed that RNAi machinery coats entire heterochromatic domains and this distribution depends on the presence of Clr4 (Cam et al., 2005). Rdp1 distribution closely mirrors the RITS and H3K9me2 profiles at most heterochromatic domains (silent mat interval, subtelomeric region and centromere) in wild type cells but it is abolished in clr4Δ cells. In contrast to RITS, Rdp1 is not substantially enriched at rDNA. Furthermore, profiling RITS-associated siRNAs, there is a marked correlation between localization of RNAi components and siRNA distribution. In addition to siRNAs mapping to centromeres, small RNAs corresponding to rDNA arrays, subtelomeric and silent mat region are observed (Cam et al., 2005). Although RDR enzymes are conserved in plants and C. elegans, they are absent in Drosophila and mammals where perhaps RNA pol II could use RNA as a template to produce dsRNA.

**The recruitment of Clr4 and its role in siRNAs production.**

Clr4 is the SET domain protein that catalyzes H3K9 methylation, the distinctive mark of heterochromatin in fission yeast. However, Clr4 plays also an important role in siRNAs production: RITS purified from clr4Δ cells does not contain siRNAs and northern analysis on total RNA from clr4Δ cells reveals absence of siRNAs. Possibly Clr4 regulates siRNAs production by promoting the association of RITS and RDRC with target chromatin (Motamedi et al., 2004; Verdel and Moazed, 2005). But how RNAi recruits Clr4 is still unclear.

Rik1 might have a role. rik1 mutants are defective in silencing at both mating type loci and at centromeres (Ekwall & Ruusala, 1994; Ekwall et al., 1996). It is possible that Rik1, whose C-terminus is 47% similar to the C-terminus of CPSF –A, is recruited as well as Cid12 and Rdrp to polyadenylation sites. In fact CPSF-A is a subunit of the Cleavage Polyadenylation factor that is thought to recognize the polyadenylation signal. In this context, considering its interaction with Clr4 (Sadaie et al., 2004), it may participate to initiate silent chromatin formation. Rik1 shares also 43% sequence similarity along its entire length with Ddb1, a protein involved in recognition of UV damaged DNA. In addition, three proteins have been discovered to interact with Rik1 and affect silent chromatin formation at centromere: an E3 ubiquitin ligase, Cullin4 (Cul4) (Horn et al., 2005), Dos1 and Dos2, two novel proteins that regulates heterochromatic RNAi and histone modification (Li et al., 2005). The Ddb1 protein associates with the Cul4 ubiquitin ligase (Groisman, 2003) just like Rik1. Maybe the ubiquitin ligase activity of the Ddb1-Cul4 is activated by a chromatin structure determined by damaged DNA as Rik1-Cul4 complexes are activated by recruitment to a specific chromatin structure such as siRNA/dsRNA. Very recently it has been
found that Clr4 and Swi6 are sumoylated and this modification is necessary for heterochromatin integrity (Shin et al., 2005).

ROLE OF RNA pol II IN CENTROMERIC TRANSCRIPTION.

What makes the first transcripts?

Recently three papers have shown that an essential component of the silencing process is the gene-expression machinery itself, in particular RNA pol II (Schramke et al., 2005; Kato et al., 2005; Djupedal et al., 2005). In this context the coupling of RNA-induced silencing with transcription becomes important for the correct targeting of the RITS complex and then of histone modifying enzymes. Very different mutations in RNA pol II disrupt the formation of heterochromatin: truncation of the RNA pol II largest subunit carboxy-terminal domain (CTD, normally required for coupling mRNA synthesis to mRNA processing (Maniatis & Reed, 2002) and a specific point mutation in the RNA pol II, Rpb2 subunit. These two mutants alleviate silencing but whereas the CTD truncation is still able to produce siRNAs, the Rpb2 mutant is blocked in processing them. A model to explain the coupling of RNA-induced silencing with transcription is that RITS goes along with RNA pol II elongation complex which produces the target mRNA. Through this interaction and the recognition of the specific transcript sequence, Clr4 then would be localized to the appropriate target gene and modify the histone as the gene is being transcribed (Buratowski & Moazed, 2005). Two other histone methyltransferases, Set1 and Set2 binding to elongating RNA pol II, modify transcribed regions of genes (Hampsey & Reinberg, 2003).

In plants this paradoxical involvement of an RNA silencing pathway in maintenance of transcriptional silencing has been solved by the existence of a distinct silencing polymerase, RNA pol IV in addition to the well-known DNA-dependent RNA polymerase I, II and III. (Herr et al., 2005; Onodera et al., 2005). In animal and fungi the Pol IV clade is absent but the silencing paradox may be solved if there are forms of RNA pol I-III with silencing-specific subunits that allow transcription of heterochromatin and, consequently, maintenance of the RNA-dependent silencing.

Another role of RNAi in fission yeast is to degrade mRNA: expression of a GFP hairpin RNA produces siRNAs complementary to GFP via the RNAi pathway without inducing chromatin modification (Sigova et al., 2004).

In animals and plants, endogenous PTGS is mediated by 20-22 nt small RNAs known as microRNAs (miRNAs), which are processed from imperfect hairpin precursor transcripts by Dicer and regulate target messages with complementary sequence via cleavage and translational arrest. These miRNAs are different from siRNAs that are derived from perfectly complementary dsRNA, typically from transposable elements. In Drosophila, miRNA precursors are exported into the cytoplasm by RanGTPase and exportin 5, where they are processed further and loaded into the RNA-induced silencing complex (RISC) (Lund et al., 2004; Yi et al., 2003). In S.pombe,
the exportin pathway impacts centromeric silencing (Kusano et al., 2004; Yoshida and Sazer, 2004) but no miRNA has so far been detected in fission yeast.

**csp Mutants: Centromere Suppressor Of Position Effect.**

The phenomenon of transcriptional repression at the centromere in S. pombe has been utilised to screen for novel centromere structural or regulatory components. To this end a strain containing ura4* and ade6* marker genes inserted on the left and right sites of centromere 1 respectively within the inner and outer repeats was constructed. This strain grows poorly on plates lacking supplement uracil and forms red colonies on indicator plates containing low supplementing adenine. This strain was ethylmethane-sulfonate (EMS) mutagenized (Ekwall et al., 1999). Two types of mutants that alleviate centromeric silencing were selected from the screen: temperature-sensitive (ts) mutants csp1 to 6 which formed white ade* colonies at 25°C but were inviable at 36°C and non temperature sensitive mutants csp7 to 13 which formed white ade* colonies. Screening against mutants that affect mat2/mat3 silencing, avoided the re-isolation of mutants such as clr4, rik1, swi6 that affect silencing at all tested heterochromatic loci (Allshire et al.,1995). All csp mutants were recessive with respect to centromere silencing defects and temperature sensitivity. Only the csp2 mutant alleviated silencing also on mating type loci and csp4 mutant alleviated telomeric silencing.

_csp9, csp10, csp11, csp12 and csp13_ have been shown to be involved in RNAi directed heterochromatin formation over the repeats and _csp9_ was found to be allelic to Argonaute1 (_ago1_), a component of RITS (Volpe et al., 2003; Verdel et al., 2004). _csp7_ and _csp10_ have been cloned by Sharon White in the lab and they are respectively allelic to _rdp1_ and _cid12_. The _csp1_ to 6 mutants are temperature sensitive and their characterization will be presented in this thesis. Some of them represent factors involved in general RNA metabolism and may provide a link between silencing chromatin components/RNA transcription and RNAi pathway.
AIMS OF THE PROJECT.

This project was focused on investigating the components of silent chromatin at fission yeast centromeres. At the beginning of this work some mutants, the csp non temperature sensitive mutants, found previously in a screen for alleviation of silencing of marker genes inserted in the outer repeats of centromere (Ekwall et al., 1999) were characterized and found to be involved in the RNAi pathway (Volpe et al., 2002). However some other mutants, the csp ts mutants found in the same screen, were not characterized.

To try to understand the role of this second class of mutants (csp 1-6), experiments of chromatin immunoprecipitation, centromeric siRNAs detection and centromeric transcripts accumulation were performed. In particular I focused on the csp3 gene that encodes for Rpb7, a subunit of RNA pol II. It has been discovered concomitantly with this work that mutations in two other subunits of RNA pol II, rpb2-m203 and rpb1-11 affect silencing at centromere (Kato et al., 2005; Schramke et al., 2005). Together in a collaborative work with the Ekwall laboratory we investigated the role of Rpb7 in centromeric transcript production and RNAi- directed silent chromatin formation. Also splicing factors, involved in general RNA metabolism, identified in the same screen as Rpb7, have been further characterized in respect to the RNAi pathway though their role is still unclear.

A mark for silent chromatin at centromere is H3K9 methylation. This modification is catalyzed in Schizosaccharomyces pombe by Clr4, a SET domain protein homolog to Suv3-9 in metazoa. There are 13 SET domain proteins in fission yeast and at the beginning of this study the target of only two of them (Setl/H3K4 and Clr4/H3K9) was known. During this study Set2 was shown to methylate H3K36 (Morris et al., 2005). In order to investigate the role and target of the other SET domain proteins, four were chosen and deleted: set3, set7, set9 and set10. The four SET domain chosen were the ones that did not have any homologue in Saccharomyces cerevisiae, H3K9 methylation being absent in budding yeast. The deleted strains were analysed for growth at different temperature and silencing phenotype at all the heterochromatic loci in fission yeast genome. In particular Set9, in collaboration with Steve Sanders in the Kouzarides laboratory, was shown to methylate H4K20 (Sanders et al., 2004). In fission yeast this modification is not required for silent chromatin formation as it is in mammalian cells and Drosophila melanogaster (Schotta et al., 2004), however it is necessary for genome integrity through the maintenance of the DNA damage checkpoint (Sanders et al., 2004).
2.1 FISSION YEAST GROWTH AND CLASSICAL GENETICS.

Growing *S. pombe*.

Haploid strains of *S. pombe* grow with the following generation times:

<table>
<thead>
<tr>
<th>Medium</th>
<th>Temperature °C</th>
<th>Generation time</th>
</tr>
</thead>
<tbody>
<tr>
<td>YE</td>
<td>25</td>
<td>3h</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>2h 40min</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>2h 10min</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>2h</td>
</tr>
<tr>
<td>minimal</td>
<td>25</td>
<td>4h</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>3h</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>2h 30min</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>2h 20min</td>
</tr>
</tbody>
</table>

For mutant strains the generation times may be longer. The time required for the population to double can be calculated more precisely using the following equation:

\[
T = \frac{\log (2^{\frac{y}{x}})}{\log (2)}
\]

Temperature sensitive strains were grown at 25°C, and 36°C was the restrictive temperature. The media used can also affect the temperature sensitivity of a mutant. Wild type strains were generally grown at 32°C.

**Liquid Cultures.**

For physiological experiments it is important that cultures are maintained in mid-exponential growth between $2 \times 10^6$ and $1 \times 10^7$ cells/ml.

To generate cultures in mid-exponential growth a fresh colony or overnight patch of a strain of checked phenotype was inoculated in 10 ml YES (or minimal medium if the strain carries a plasmid with selectable marker) and incubated the day before at the appropriate temperature until cells were in early stationary phase. This pre-culture was used to inoculate a larger culture, taking into consideration the generation times.

Flask size was selected according to the required volume of culture: 25 ml flasks for up to 10 ml
culture, 100 ml flasks for up to 50 ml, 200 ml for 100 ml cultures, 250ml for 125ml cultures and 500 ml for 250 ml cultures.

The media for growing S. pombe strains are given below. YES liquid or agar containing solid medium was used for vegetative growth.

Temperature/cold sensitivity.
Temperature sensitive (ts) and cold sensitive (cs) mutants were checked by replica plating onto YEP (YES + phloxin B) and incubated at the restrictive temperature. Phloxin B stains the dead cells.

Auxotrophy.
The auxotrophic markers most commonly used in S. pombe require adenine, arginine, histidine, leucine, lysine and uracil. We used 100 mg/L (4 ml of 10 mg/ml stock solution per 400ml medium). To test for auxotrophy the strain is grown up to single colonies on YES and then replica plated to minimal medium with and without the appropriate supplement. The plates are incubated for 1-2 days and then examined for growth under these conditions.

Serial dilution assay.
To assess the growth of various mutant strains on different media or at different temperatures, cells from a fresh plate were resuspended in 0.5 ml of liquid YES and counted with a coulter counter, followed by serial dilutions in YES in a sterile microtiter plate. 5 μl of cells were spotted at the following concentration per plate: 5x10^5, 5x10^3, 1x10^3, 200, 40, 8. Plates were then incubated at the desired temperature until colonies appeared.

Fission Yeast Classical Genetic Techniques.

Genetic Crosses.
Conjugation and sporulation cannot take place in S. pombe except under conditions of nutrient starvation. ME plates lacking nitrogen were generally used for genetic crosses. To cross two strains a loopful of freshly growing h^- and a loopful of freshly growing h^+ were mixed together on a ME plate. A loopful or two of sterile distilled water was then used to thoroughly mix the cells on the agar plate to an area of about 1 cm^2. The cross was left to dry and is then incubated at 25°C. Fully formed four spore asci could be seen after 2-3 days incubation.

Random spore analysis.
A three day old cross was checked for the presence of asci under the light microscope. Random spore analysis allows many more spores to be examined than in tetrad analysis and in this way recombination mapping and strain construction can be carried out. However, it is important that all the classes of spores are viable when studying recombination frequencies,
otherwise tetrad analysis becomes necessary. 1 ml of sterile distilled water was inoculated with a loopful of the cross, 400μl of a 1 in 100 dilution of glusulase was added and the mixture incubated overnight at 25°C or for at least 6 hours at 25°C. Glusulase is a crude snail gut enzyme that breaks down the ascus wall and kills vegetative cells. The spore number/ml is counted using a haemocytometer. Between 200-1000 spores/plate were plated out on YES Agar or selective medium and incubated at permissive temperature until colonies formed.

Tests for Allelism. Tetrad Dissection.
Recombination Frequency (looking for linkage).
Pair wise crosses were carried out between the mutants to be mapped; resulting asci were treated with glusulase as normal and approx. 200 spores plated out. It was helpful to have unlinked markers (histidine and arginine) present in at least one of the parental strains to ensure that colonies which result are the products of crossing, and not simply cells which have survived the glusulase treatment. The presence of wild type progeny in a cross between two temperature sensitive strains indicated that the two mutations under observation were in different genes, as recombination has taken place. In this case Mendelian segregation resulted in 25% of the spores being wild type, assuming the two genes are un-linked.

Construction of Double Mutants.
These can usually be made by free spore analysis and selecting by replica plating for suitable nutritional markers G418 for kanamycin, FOA for uracile or phloxine B for temperature sensitive.

Isolation of Diploid Strains.
Sporulating diploids can be isolated by crossing h⁻ and h⁺ haploids with complementary growth requirements, ade6-M210 and ade6-M216 mutations were used. ade6-M216 colonies are light pink and ade6-M210 dark pink on plates of YE medium or EMM containing 10 mg/L adenine because of accumulation of a red adenine precursor. On media containing adenine in excess of 10 mg/L the red colour is not observed. Diploid cells containing both mutations grow in the absence of adenine and the colonies formed are white due to intragenic complementation between the two ade6 alleles. Because the alleles are tightly linked and there is infrequent gene conversion, spores generated by meiosis are unlikely to be adenine prototrophs and will not form colonies on the selective media. Sporulating diploids are very unstable and will generate spores if they enter stationary phase from minimal medium. Non-sporulating diploids can be screened by replica plating onto malt extract and looking for non iodine positive colonies, which are not undergoing sporulation.
Protocol used for generating an h⁺/h⁻ sporulating diploid
1) cross an ade6-M210 strain to an ade6-M216 strain on ME-Glutamate. Incubated plates at 25 °C.
2) 8h, 12h and 24h after mixing the cells, take a loop-full and streak on minimal medium lacking adenine (PMG-ade). Incubate this plate at 25 °C.

3) Pick colonies as soon as they appear on the EMM-ade plate (typically after 2-4 days). If diploids are left too long on the EMM-ade plate they will undergo sporulation) and streak on YE (no supplements). Put the YE plate to grow at 25 °C for two days.

4) There are usually two types of colonies on the YE plate: larger, white colonies (these are diploids) and smaller, red colonies (these are haploid colonies ade- which grew from spores generated when the diploid reached stationary phase on the PMG ade plate and underwent sporulation, they are red because YE contains low levels of adenine and ade- cells will grow but accumulate a red metabolic precursor). Patch the diploids on YES and replica plate the YE plate on ME-glutamate. Check after 2-3 days that the diploids picked are sporulating.

**Mapping one gene**

The first step to map one gene is to assign it to a chromosome. There are several ways to do this. The way it was done here required chromosome loss experiments. A nonsporulating diploid with differently marked chromosomes was constructed. It was grown in liquid YES with 25 μg/ml thiabendazole (TBZ), a drug which promotes chromosome loss. Because *pombe* rapidly haploidises when one chromosome is lost, associated linkage of the mutation to a marked chromosome can be determined (See Figure 6-5 in chapter 6). After the chromosome assignment, mapping of the mutations against the markers on that chromosome can be performed using the swi5 technique. A swi5 mutant is not only defective in switching but has a general reduction in recombination frequency. That is, with a five fold reduction in recombination frequency, distant linkage that would otherwise be invisible can be seen. Thus with just a few markers in the background, the gene of interest is likely to show linkage to at least one. Practically, this means constructing strains with the mutation in a swi5 background, and then mapping against a swi5 strain with markers on all chromosomes.

**Media for the fission yeast *Schizosaccharomyces pombe***

All of the following media were autoclaved unless otherwise stated.

**Per litre**

PMG Agar in 1 L:
- 3.0 g Phthallic acid
- 2.2 g Di-sodium orthophosphate
- 3.75 g Glutamic acid
- 20.0 g D-Glucose anhydrous (Fisher Scientific)
- Vitamins 1000X 1.0 ml
- Minerals 10,000X 0.1 ml
- 50X Salts 20.0 ml
Agar 20.0 g (OXOID)

**PMG in 900ml:**
- 3.0 g Phthallic acid
- 2.2 g Di-sodium orthophosphate
- 20.0 g D-Glucose anhydrous (Fisher Scientific)
- 3.75 g Glutamic acid
- Vitamins 1000X 1.0 ml
- Minerals 10,000X 0.1 ml
- 50X Salts 20.0 ml

**YES (no ade) agar:**
- 5.0 g Yeast Extract (DIFCO)
- 30.0 g D-Glucose anhydrous (Fisher Scientific)
- 0.2 g Arginine (Sigma)
- 0.2 g Histidine (Sigma)
- 0.2 g Leucine (Sigma)
- 0.2 g Lysine (Sigma)
- 0.2 g Uracil (Sigma)
- 20.0 g Agar (OXOID)

**YES liquid:**
- 5.0 g Yeast Extract (DIFCO)
- 30.0 g D-Glucose anhydrous (Fisher Scientific)
- 0.2 g Adenine (Sigma)
- 0.2 g Arginine (Sigma)
- 0.2 g Histidine (Sigma)
- 0.2 g Leucine (Sigma)
- 0.2 g Lysine (Sigma)
- 0.2 g Uracil (Sigma)

**50X Salts:**
- 53.5 g Magnesium Chloride 6H2O
- 1.0 g Calcium Chloride 6H2O
- 50.0 g Potassium Chloride
- 2.0 g Di Sodium Sulphate

**1000X Vitamins (100ml):**
- 0.5 g Pantothenic acid
- 1 g Nicotinic acid
- 1 g Inositol
- 1 mg Biotin

(Filter sterilised)
10,000X Minerals: 
- Boric acid 5g
- MnSO₄ 4g
- ZnSO₄ 4g
- FeCl₂ 6H₂O 2g
- Molybdic acid 1.6g
- KI 1g
- CuSO₄ 5H₂O 0.4g
- Citric acid 10g

(Filter sterilised)

Supplement Stocks: 
- 50X adenine (Sigma) 5 g/l
- 100X arginine (Sigma) 10 g/l
- 100X histidine (Sigma) 10 g/l
- 100X leucine (Sigma) 10 g/l
- 100X uracil (Sigma) 10 g/l (dissolved by adding NaOH)

Malt Extract plates (ME): 
- Agar 20g/l (OXOID)
- Malt Extract 30 g/l (OXOID)
- adenine 250 mg/l (Sigma)
- arginine 250 mg/l (Sigma)
- histidine 250 mg/l (Sigma)
- leucine 250 mg/l (Sigma)
- uracil 250 mg/l (Sigma)

5-FOA plates: 
- PMG or YES-agar
- 1g/l 5-FOA (Melford laboratories)

(added to melted agar below 60°C)

TBZ plates: 
- YES-agar
- TBZ (stock 10mg/ml in DMSO) to 10mg/ml or 15mg/ml

(added to melted agar below 60°C)

All the solutions were made up to the final volume with added distilled H₂O.
2.2 MOLECULAR GENETICS.

S. pombe plasmids.

S. pombe plasmids consist of a bacterial origin of replication and selectable marker, a yeast selectable marker and an equivalent to an autonomous replication sequence (ars). Budding yeast markers used in S. pombe are the LEU2 and URA3 genes. Plasmids containing these markers complement the S. pombe mutations leu1^+ and ura4^+. In contrast to S. cerevisiae, in S. pombe a bacterial plasmid such as pBR322 carrying a marker gene such as LEU2 is able to replicate often to high copy number. However, the transformation frequency obtained when using such plasmids is very low. The addition of S. pombe autonomous replication sequences (ars) or the S. cerevisiae 2μ origin leads to high frequency of transformation and reduction in the copy number. In S. pombe high frequency of transformation and effective replication capacity are to some extent independent phenomena. Plasmid vectors based on 2μ (pDB248, YEp13) are mitotically unstable; their copy number is low, they are much more prone to rearrangements (tandem duplications or deletions) and they are more difficult to recover from fission yeast than plasmids carrying S. pombe ars1^+. Plasmids containing ars1^+ are also very unstable; their copy number is higher and they tend to produce polymers with various numbers of repeats units. pFL20 and pMB332 yield rather stable transformants both mitotically and meiotically, due to the presence of a stb (stable) element. This element is not an ars sequence nor it is a centromeric sequence. Plasmids containing this element still segregated asymmetrically ten times more frequently during mitosis than S. cerevisiae CEN plasmids.

Expression vectors.

Plasmids containing inducible promoters have been developed. These include various plasmids containing the thiamine repressible promoter, nmt1, as developed by Maundrell (1990). Several versions are available in which the promoter sequences have been mutated to different degrees to give lower levels of expression (Forsburg, 1993). pREP1 contains the wild type promoter. There is a significant background expression level with thiamine. The induced level is about 80X greater than the repressed level and about 6X greater than the level produced by the adh1 promoter. pREP41 has a 6X lower induced level and a 15X lower repressed level than the wild type promoter. pREP81 has an induced level about 80X lower than the wild type promoter (comparable to the repressed level of the wild type promoter), whilst in the absence of thiamine the level is reduced a further 250 fold.
Transformations.

Lithium acetate procedure.
1. Grow fission yeast cells in YES to 1x10^7 cells/ml.
2. Pellet 50ml of cells per transformation.
3. Wash cells in 50ml sterile water. Transfer to eppendorf in 1 ml water. Wash in 1ml of LiAcTE.
4. Resuspend in LiAc-TE at 2x10^9 cells/ml.
5. Mix 100 μl cells with 10 μl of DNA; mix gently.
6. Incubate at RT for 10 min.
7. Add 260μl of 40% PEG/LiAc-TE; mix gently.
8. Incubate 30-60 min at 30°C.
9. Add 43μl pre-warmed DMSO; mix gently.
11. Pellet and wash once with 1ml water.

Solutions: LiAc-TE: 0.1M lithium acetate, 10mM Tris pH 7.5, 1 mM EDTA
LiAc-TE-PEG: LiAc-TE plus 40% PEG4000

Electroporation.
1. Grow cells to a density of 1 x 10^7/ml in YES.
2. Harvest cells by spinning at 3000 rpm (J 2-21) for 5 minutes at 20°C. Wash once by resuspending in ice-cold water and harvesting; a second time by resuspending in ice-cold 1.2 M sorbitol and harvesting.
3. The final resuspension is in ice-cold 1.2M sorbitol at a density of 1 - 5 x 10^9 / ml.
4. 200 μl of the cell suspension are added to chilled cuvettes containing the DNA for transformation (100 ng ) and incubated on ice for 5 minutes.
5. Dry bottom of cuvette and pulse cells. The electroporator (Biorad) is set to 1.5kV, 200Ω, 25μF.
6. Cells and DNA are transferred to a pre-chilled cuvette and pulsed; 0.8 ml of ice-cold 1.2M sorbitol is then immediately added to the cuvette; the cell suspension is then returned to the Eppendorf and placed on ice.
7. Cells are plated as soon as possible onto selective medium. Transformants appear in 4 - 6 days at 30°C

Integration of a plasmid into the genome.
In S. pombe integration by homologous recombination is usually more frequent than non-homologous recombination; but for certain loci homologous recombination may only represent
about 5-10% of the integration events. On average about 0.1% of the transformants obtained after transformation with an ars plasmid will have an integrated copy of the plasmid at the homologous locus. The frequency of integration can be enhanced up to ten fold by linearizing the plasmid in the region of interest to facilitate the recombination event.

To isolate an integrant:
1. Transform a yeast strain with the plasmid of interest.
2. Isolate a transformant colony and grow up in 100 ml of YES medium (i.e. non selective conditions) for about 20 generations (re-inoculate 1ml of this culture into 100 ml of fresh YES medium 2-3 times).
3. Plate out about 1000 cells/plate onto selective medium and incubate until colonies form. These colonies should be stable due to integration of the plasmid into the genome. This can be tested by replica plating to YES medium twice and then back to selective medium.
4. The integration can be confirmed by Southern blotting.

**Gene disruption – SET domain proteins disruption (Bahler et al., 1998).**

The SET domain protein gene disruption was performed using PCR-based gene targetting plasmids containing kan module (ARC 784 ) The protocol used was the following:
1. Design primers: 100 mer oligos with 80 nt to gene of interest and 20 nt to kan module;
2. Do 2-5 independent PCR reactions using as template ARC 784. Pool reactions. Purify the DNA by phenol/chloroform extraction and concentrate by ethanol precipitation. Resuspend in 10µl of TE.
3. Follow lithium acetate method for transforming PCR products and a Leu2 plasmid. Grow transformed cell overnight at 25°C in selective PMG -LEU liquid media. Plate on PMG - LEU + kanamycin the day after.

**Histone Mutagenesis.**

Fission yeast has three pairs of genes encoding histones H3 and H4: h3.1-h4.1, h3.2-h4.2, h3.3-h4.3. Deletion and phenotypic analysis of each H3-H4 pairs have been previously performed (Mellone et al.,2003; Mellone and Allshire unpublished). These analysis established a procedure for retaining a single residual H3.2 or H4.2 gene so to create then strains with only the mutated histone H4 gene and the H3 partner. The K20 of the H4 within the h3.2-h4.2 gene locus was mutated to arginine following the procedure described in Mellone et al., 2003. A strain (Fy 4756) in which h4.2 has been replaced by ura4 was transformed with a PCR product for H4 where the codon for K20 (AAG) was substituted with that encoding for R (CGT). Following transformation, cells were spread on counterselective 5-FOA plates to select for those lacking ura4. FOA resistant colonies were checked by PCR for the presence of h4.2 and absence of ura4 using a pairs of primers just upstream and downstream h4.2 gene (data not shown).
Genomic preparation from these 5-FOAR transformants was done and PCR was used to amplify the H4.2 gene. The product was sequenced using specific 5' and 3' primers to detect the presence of K20R mutation. The mutated strains still retained wild type h4.1 and h4.3 genes and they were deleted by crosses. In the strain Fy 4640, h3.1-h4.1 gene pairs are deleted with the marker gene his3<sup>+</sup>, h3.3-h4.3 are deleted with the marker gene arg3<sup>+</sup>, in addition the remaining h3.2-h4.2 gene pairs tagged by insertion of ura4<sup>+</sup> nearby. This Fy 4640 strain was crossed to the newly generated cells bearing the h4.2K20R mutation and the spores plated on -HIS-ARG+5FOA. The colonies from the cross were tested by PCR for the presence of arg3 in place of h3.1/h4.1 loci and his3 in place of h3.3/h4.3. Sequencing with specific primers upstream and downstream h4.2 was performed in order to confirm that these progeny retained the H4K20R mutation.

Recovering plasmids from S. pombe.
The following procedure was used to recover the plasmids from S. pombe Shimoda library transformants:
1. Grow up 10 ml of cells under selective conditions (PMG-LEU) to 2 x 10<sup>7</sup> cells/ml.
2. Spin down the cells 3000 rpm 5 minutes.
3. Resuspend in 0.5 ml of SP1 buffer and add 1 mg/ml Zymolyase-100T. Transfer to an Eppendorf tube and incubate at 37°C for 1 hour.
4. Pellet the cells in an Eppendorf centrifuge for 30 seconds and resuspend in lysis buffer from Qiagen miniprep kit. Follow the protocol as for miniprep.
5. Transform 100 µl of DH5a competent cells with 20 µl of the miniprep of the recovered plasmids.
6. Plate on selective LB plates (usually LB + 30 µg/ml Ampicilline).

2.3 FISSION YEAST PHYSIOLOGY.
The cell cycle and cell growth.
S. pombe cells coordinate cell growth and the cell cycle. Exponentially growing wild type cells are born at a similar cell size and double their mass before entering the next round of cell division. This is the result of a size control coupling growth and the cell cycle. They grow by tip elongation, and therefore cell length is a measure of the cell cycle stage of a cell. To achieve this, exponentially growing wild type S. pombe cells use a size control in G2 which controls the size requirement for entry into mitosis. However, if this size control is eliminated by mutation or an environmental change, then a cryptic G1 size control comes into play. Rapidly growing wild type cells are born at a cell size that exceeds this G1 size control so it does not delay S-phase. However, small S. pombe cells spend a larger portion of the cell cycle in G1 and use this control
to coordinate growth and cell division. The existence of a size control, means that blocking cell growth leads to cell cycle arrest.

There are two accurate means of determining cell number, a haemocytometer or electronic cell counters.

The haemocytometer.

The haemocytometer is a specialised microscope slide on which 2 grids have been engraved, in a central region that is 0.1 mm lower than the rest of the slide. Each grid comprises 25 large squares, each containing 16 smaller squares of area 1/400 mm². This creates a region of known volume (0.1 mm³) when a special coverslip is placed over the central region 10 μl of culture were pipetted under the coverslip and cells counted in 5 gridsquares. Multiplying the total number of cells in the entire grid by 1x10⁴ gives the number of cells / ml.

The coulter counter.

For more accurate measurements of cell number the Coulter Counter can be used. It has the additional advantage that samples can be taken quickly during an experiment then processed at leisure. It also gives a measure of mean cell volume. 100μl of cell culture was mixed in 10ml FacsFlow solution, sonicated for 10 sec in a waterbath Sonicator (MSE Soniprep) and counted.

Microscopy.

Live analysis of GFP-Swi6 expressing cells.

Cells derived from mutants crossed to the strain containing the pREPnmt81X-GFP-Swi6 (FY2214) were grown to log phase in 2 ml of thiamine-free medium (PMG) supplemented with all amino acids. 1% low melting point agarose (Gibco) in culture medium was boiled and cooled to 37°C ready for use. Cells were pelleted, resuspended in small volume of culture medium and 4 μl put on a microscope glass slide. 6μl of agarose were added to the cells and mixed well. A glass coverslip was applied on top and pushed down firmly. Slides were observed at the fluorescence microscope and pictures taken on a Carl Zeiss MicroImaging, Inc. Axioplan2 IE fluorescence microscope. Metamorph software was used for image acquisition.

Immunostaining.

20-25ml of a cell culture was grown to a concentration of 5x10⁶ cells/ml. Cells were fixed by adding 3.8% paraformaldehyde dissolved in culture medium (a 10X stock was dissolved at 65°C and cooled-down to RT) and shaking the culture at RT or 18°C for 30 min. Cells were spun in a refrigerated benchtop centrifuge at 18°C, washed once with 15 ml PEM, transferred to a 1.5ml
tube and washed twice with PEMS. Cells were then incubated at 36°C for 90 min in PEMS containing 1 mg/ml zymolyase 100T (ICN) at a concentration of less than 10⁶ cells/ml. After this time cells were observed on a light microscope to check for digestion of the cell wall (cells become dark). After being washed with 1 ml of PEMS, cells were resuspended in 200 μl of PEMS containing 1% Triton-X100 and incubated on the bench for 5 min. Cells were then washed once with 500 μl of PEM, resuspended in 500 μl of PEMBAL and incubated on a rotating wheel for 60 min at RT. Aliquots of cells were taken to be incubated with the appropriate dilution of primary antibodies in 100 μl of PEMBAL overnight at 4°C on a rotating wheel.

The antibodies and the dilutions used in this thesis were:
- anti-GFP, polyclonal (rabbit) GENESIS 1/100
- anti-trimetil-H4K20, polyclonal (rabbit) ABCAM 1/100
- anti-Cnp1, polyclonal (sheep) crude serum, 1/500
- anti-Nop1, monoclonal (Tollervey), 1/100

After incubation with primary antibodies, cells were washed three times with 500 μl of PEMBAL incubating for 30 min with rotation at each wash. The required secondary antibodies (Molecular probes ALEXA anti-mouse, anti-rabbit or anti-sheep) conjugated with the desired fluorescent conjugate (Texas red or FITC) were added at the concentration of 1/1000 in 100 μl of PEMBAL. Tubes were wrapped in foil and incubated at RT with rotation for 4 h followed by one 30 min wash with 500 μl of PEMBAL. Followed a 5 min incubation in PEM-0.1% Sodium Azide containing 1 mg/ml DAPI (stock 500X stored at -20°C). Cells were finally spun and resuspended in 20 μl of PEM-0.1% Sodium Azide. 5 μl of cells were spread in a thin layer on a glass coverslip previously coated with 1 mg/ml Poly-L-lysine hydrobromide (Sigma) and allowed to dry. Coverslips were then placed face-down avoiding formation of air bubbles on a glass slide where 5 μl of Vectashield were placed. The coverslips were sealed with transparent nail polish and observed at a Carl Zeiss MicroImaging, Inc. Axioplan2 IE fluorescence microscope. Metamorph software was used for image acquisition.

The immunostaining with the monoclonal antibody against Nop1 was performed according to this protocol with the exception that the primary and secondary antibodies were diluted in 5% milk PBS (0.2% Tween) instead of PEMBAL. 5% milk PBS was used to pre-blocked the cells before incubation with the antibody.

PEM: 100 mM Pipes, pH 6.9; 1 mM EDTA; 1 mM Mg₂SO₄
PEMS: PEM containing 1.2 M Sorbitol
PEMBAL: PEM containing 1% BSA (Sigma), 0.1% Na Azide; 100 mM Lysine hydrochloride (BDH).
2.4 *S. pombe* MOLECULAR BIOLOGY.

DNA preparation and analysis.

Small scale *S. pombe* DNA preparations.

5 ml of stationary phase cell culture was pelleted at maximum speed in a benchtop centrifuge. The pellet was resuspended in 250 ml SP1 containing 0.4 mg/ml Zymolyase 100T (ICN) and incubated for 60 min at 37°C. Spheroplasted cells were pelleted at 8,000 rpm in an eppendorf centrifuge for 15 sec. The pellet was resuspended in 0.5 ml TE, 50 ml of 10% SDS was added, followed by vortexing and addition of 165 µl 5M potassium (K) acetate. Samples were stored on ice for 30 min. The supernatant was added to 0.75 ml isopropanol, placed on dry ice for 5 min and centrifuged for 10 min. The pellet was resuspended in 0.3 ml TE and RNase added to 10 mg/ml. After 1 h 30min at 37°C, the sample was extracted with phenol/chloroform and precipitated with Ethanol. Genomic DNA was resuspended in 20 µl TE.

SP1: 1.2M Sorbitol, 50 mM Sodium Citrate; 50 mM Sodium Phosphate; 40 mM EDTA. pH to 5.6.

Rapid *S. pombe* genomic DNA preparation for PCR (Ling et al., 1995).

A small amount of cells was picked with a sterilised cocktail stick from a fresh patch of cells and placed in a sterile microfuge tube containing 15µl of SPZ buffer. After mixing the suspension was incubated at RT for 10-15 min and 1 µl of it was used as DNA template for a 20-50 µl PCR reaction.

SPZ buffer: 1.2 Sorbitol; 100 mM Sodium phosphate, pH 7.4; 2.5 mg/ml Zymolyase 100-T. Stored at −20°C.

Total *S. pombe* RNA preparation.

Cells were grown at 32°C in YES medium to a density of approximately 1x10⁷ cells/ml. The cultures were pelleted by centrifugation, washed in TE and transferred to microfuge tubes prior to being resuspended in 300 µl RNA extraction buffer. 300 µl of glass beads (Sigma) were added followed by 300 µl phenol/chloroform pH 4.7 (Sigma). The microfuge tubes were shaken at high speed on a multi-head vortexer for 30 minutes at 4°C to lyse the cells, followed by centrifugation at 10,000 rpm for 5 minutes and removal of the supernatant. The supernatant was extracted twice with phenol chloroform, and then once with chloroform. The RNA obtained was precipitated with 3 volumes cold 100% ETOH and centrifuged at 10,000 rpm for 15 minutes at 4°C. The pellet was air-dried and resuspended in 25 µl dH₂O. The concentration of each sample of RNA was determined by measuring the optical density of a 1/500 dilution in a spectrophotometer set at a wavelength of 260 nm. The samples were then diluted to a concentration of 1mg/ml with dH₂O.
RNA Extraction Buffer: 50 mM Tris-HCl pH 7.5, 10 mM EDTA, 100 mM NaCl, 1% SDS.

RT-PCR.

For RT-PCR 1 μg of each sample of RNA was aliquoted into a microfuge tube. 1 μl DNAse buffer (INVITROGEN), 1μl DNAse and dH₂O were added to a final volume of 10 μl and the mix was incubated at 25°C for 1 hour. After this time 1 μl 25 mM EDTA was added and the reactions incubated at 65 °C for 10 minutes to denature the DNAse. 1 μl of the reaction was used as template for a PCR reaction to check that all of the DNA had been digested. Once it had been established that there was no DNA remaining as a contaminant in the sample, 1 μg oligo dT₁₇ or CenRev or CenFor primers and dH₂O were added to a final volume of 25 μl. The samples were incubated for 10 minutes at 70 °C, then placed on ice. The samples were centrifuged briefly at 4°C and returned to ice before adding 8 μl 5X Reverse Transcriptase+MgCl₂ buffer (Roche), 4 μl DTT(25 mM) and 2 μl 2.5 mM dNTPs. The samples were mixed then split into two 19ml aliquots and incubated at 42 °C for 5 minutes. 1 μl reverse transcriptase (Superscript GIBCO) was added to one of each pair of tubes only. The incubation at 42 °C was continued for another 50 minutes then the samples were incubated at 70 °C for 15 minutes to stop the reaction and the samples returned to ice. 1ml of each sample was used as template in a 20 μl PCR reaction to determine the amount of transcript.

SMALL RNA PREPARATION.
- 40 ml culture at ~1 x 10⁷ / ml.
- Pellet.
- Resuspend in 500 μl TE, transfer to microfuge tube. Pellet.
- Resuspend in 500 μl extraction buffer,
  Add 500 μl glass beads (use PCR tube to measure beads) 500 μl phenol
- Lyse cells on multi-head vortexer for 15-30 min high speed at 4°C.
- Spin 5 min. Transfer s/n to fresh tube.
- Extract with phenol/chloroform.
- Extract with chloroform.
- Remove large rRNA, mRNA and genomic DNA by precipitation with 10% PEG (mol wt 8000)/0.5 M NaCl. Incubate on ice for 30 minutes
- The low molecular weight RNA including the siRNA is precipitated from the PEG/NaCl supernatant by the addition of 3 volumes of ethanol/-20/ overnight.

Save the PEG precipitated material (pellet) and dissolve it in 30 ul of 50% deionised formamide to compare it with the subsequent low molecular weight fraction. It can also be used for mRNA analysis, PCR etc.
- Wash pellets containing siRNA in 95% ethanol (cold) and spin again. Air dry and dissolve in 30 ul 50% deionised formamide. Formamide protects against chance ribonuclease activity and also helps loading onto the gel.

The efficiency of enrichment can be checked by running the two fractions (i.e. PEG-insoluble high molecular weight nucleic acid versus PEG-soluble/EtOH-insoluble low molecular weight nucleic acid) side by side on a standard, non-denaturing agarose minigel and stain with ethidium bromide. Most or all of the DNA, rRNA and mRNA will have precipitated and be in the high mol. wt. fraction. The siRNA will be in the low mol.wt. fraction which should migrate as a bright "blob" near the 100bp DNA marker (it is mainly 5S, 5.8S rRNA and tRNA).

- Store at —80°C.

**Extraction buffer**

- 50 mM Tris-HCl pH 7.5
- 10 mM EDTA
- 100 mM NaCl
- 1% SDS

DEPC-treated dH₂O can be used for solutions, but this is not necessary, as long as fresh autoclaved dH₂O is used and gloves are worn.

**Denaturing PAGE.**

Electrophoresis on 17.5% polyacrylamide-7M urea-gel using mini-protein gels from Hoefer.

**Preparation of the Gel(s):**

Preparation for 30 ml (for 2 Gels):
- 12.6 g urea (Gibco-BRL, gives 7M)
- 13.1 ml 40% acrylamide:bis-acrylamide 19:1 (Severn Biotech Ltd.)
- 1.5 ml 10X TBE
- heat (37°C waterbath) until the urea is completely dissolved
- leave to cool down
- Pour gel, for 1 gel use 15 ml and add:
  - 240 µl 10% ammonium persulfate
  - 11µl TEMED
- mix well, pour gel
- allow to polymerize
- set up with 0.5X TBE running buffer

Prior to loading, the gel is pre-run for 30 minutes at 80V and just before loading samples, any urea that has leached out is flushed from the wells with running buffer (0.5xTBE).

**Preparation of the RNA Samples:**

Generally between 30 and 40 µg of siRNAs are loaded:
- add 4x sample buffer to 10 μl of RNA
- denature 5 min 95°C,
- put on ice until loaded

4x Sample buffer
5 mM EDTA
0.03% Bromophenol Blue
50% glycerol
50 mM Tris pH 7.7
Mix well and load.
For the ladder a DNA oligo which corresponds to 1 μl of a 10 μM dilution of the primer + H2O+
100% formamide + loading buffer. Heat-denature as for the RNA samples.

Electrophoresis:
- set limits to 80V, 400mA, 10Watt
- run until bromophenol blue reaches lower edge of gel

Stain:
- stain 10 min in 0.5X TBE, 1μg/ml ethidium bromide
- wash 20 min in 0.5X TBE

Blotting and Fixing.
Transfer by electroblotting
Equipment: Semi-dry blotter
Membrane: Hybond N

Setup:
(in the cold room, from bottom-anode to top-cathode)
- 3 sheets Whatman soaked with 0.5X TBE
- membrane soaked with 0.5X TBE
- gel, equilibrated to 0.5X TBE
- 3 sheets Whatman 3MM, soaked with 0.5X TBE

Transfer:
- transfer RNA to the membrane at 12V for 1h

Hybridization.
- Equilibrate the membrane on Whatmann 3MM soaked with SSCx2 for 10-20 minutes;
- UV cross-link the RNA with a Startalinker (2400 microJoules);
- Let the membrane dry for 5 minute on air;
- Pre-hybridize with with 25 ml of Perfect-Hyb Plus buffer (SIGMA) at 42°C for 1 hour;
- Add the desalted probe and hybridize at 42°C overnight.

Probe preparation
To detect a small RNA population, the random priming reaction was used: this generates many small fragments as a result of Klenow activity. These fragments are ideal to hybridize to 21 nt long RNAs.
- 25 ng of template DNA were added to H₂O to reach a volume of 11 μl;
- the DNA was denatured at 95°C for 10 minutes and chilled quickly in ice;
- 4 μl of High Prime was added to the denatured DNA;
- 5 μl of 50 μCi [α³²P]dCTP were added.
The reaction was incubated 1 hour at 37°C, the unincorporated nucleotides were removed with a G25 sephadex column, the probe was heat denatured and added to the hybridization buffer.

Washes
The membrane was washed twice with 25 ml of 2x SSC, 2% SDS for 20 minutes at 50°C and exposed to phosphorimager for at least 4 hours.

Protein techniques.

Chromatin immunoprecipitations (ChIP).

This protocol was adapted from Ekwall and Partridge, 1998.
50 ml of exponentially growing cells (5x10⁶ cells/ml) were shifted from the growing temperature (25°C or 32°C) to 18°C for 30min. The same number of cells were utilized in all samples. Cells were fixed for 30min at 18°C with 3% paraformaldehyde dissolved in culture medium. Fixation was stopped by adding 2.5 M glycine (20X) to cultures for 5 min at RT with agitation. Cells were then washed twice with 20ml of ice-cold PBS, washed once in 1ml of PBS. Cells were resuspended at 10⁹cells/ml in PEMS added with 0.4 mg/ml 100T Zymoliase, incubated at 37°C for 20 min. After checking at the microscope the percentage of digestion cells were spun and resuspended in 500 μl of ice-cold lysis buffer containing protease inhibitors cocktail (100X Sigma) and 2mM PMSF. Lysates were sonicated to shear the chromatin for 20 sec for three times at maximum amplitude on ice. This results in shearing the chromatin to approximately 500-1000bp. The lysate was spun at maximum speed in a microfuge for 5 min at 4°C to eliminate debris. The supernatant was transferred to a fresh tube and spun again for 15 min. The chromatin obtained was pre-cleared by adding 1/6 of the volume of beads (Roche) protein G agarose (for anti-HA and anti-Pol II ChIP) or protein A agarose (for, α-H4K20me3, α-H3K9me2
and α-Swi6 ChIP) in a 1:1 v/v suspension in lysis buffer and incubating with gentle rocking for 1-2 h at 4°C. Beads were spun at 8000 rpm for 3 minutes at 4°C and the supernatant was transferred to a fresh tube. 1/10 of this pre-cleared lysate was frozen and constituted the total input DNA sample. The appropriate amount of antibodies was then added to the remaining lysate at the following concentrations:

- α-H3K9me2 (T. Jenuwein) and α-H4K20me3 (Abcam): 10μl in 300μl lysate (1:30).
- α-H3K4me2 (Upstate): 8 μl in 200 μl (1:25).
- α-Swi6: 10μl in 400μl lysate (1:40).
- α-HA (monoclonal B. Earnshaw): 10μl in 300 μl lysate (1:30).
- α-Pol II (Covance): 6μl in 300μl (1:50).

Tubes were incubated at 4°C with gentle rocking for 1 hr. A volume of beads resuspended 1:1 in lysis buffer equivalent to 1/6 of the volume of lysate was added and tubes were incubated overnight at 4°C with gentle rocking. Beads were spun at 8000 rpm, washed for 5 min at RT with gentle rocking with 1ml of each of the following buffers: Lysis buffer, Lysis buffer 0.5M salt, Wash buffer, TE pH8. After the washes 250 μl of TES were added to the beads and tubes were incubated at 65°C overnight. TES was added also to the total input DNA samples (T) to obtain the same final volume. This treatment is made to reverse the crosslinking. After cooling down samples, 250 μl of TE and 25 μl of 10 mg/ml Proteinase K (Roche) were added and tubes were incubated at 37°C for 2 hr. Samples were phenol+chloroform extracted to eliminate proteins and the DNA samples were transferred into 2 ml tubes. To the DNA samples were added 1/10 of the volume of 3 M NaOAc; 2.5 volumes of ice cold ethanol 1.5 μl of 10 mg/ml Glycogen. Samples were mixed thoroughly by vortexing and incubated in dry ice for 1 hr. DNA was recovered by centrifuging the samples at 4°C for 30 min at maximum speed. The pellet was dried under a fume hood. ChIP DNA (IP) was resuspended in 30 μl and total input DNA (T) in 300 μl of TE. 2 μl of DNA were used for 20 μl PCR reactions containing 3 mM MgCl₂ with primer pair 41/42 for *ura4* PCRs and primers otrA/otrB, tmA/tmB, imrAl/imrB and fpb5'/fbp3' for multiplex PCR. The 'ura4' program was used for all the PCR of ChIP DNA described in this thesis: 94°C 4min; (94°C 30sec; 55°C 30sec; 72°C 1 min) 30 times; 72°C 5 min.

Lysis buffer: 50 mM Hepes-KOH, pH7.5; 140 mM NaCl; 1 mM EDTA; 1% (v/v) triton X-100; 0.1% (w/v) sodium deoxycholate.

Lysis buffer 0.5M NaCl: Lysis buffer containing 500 mM NaCl.

Wash buffer: 10 mM Tris-HCl, pH8; 0.25 M LiCl; 0.5% NP-40; 0.5% (w/v) sodium deoxycholate; 1 mM EDTA.

TE: 10 mM Tris-HCl pH8; 1 mM EDTA.

TES: 50 mM Tris-HCl pH8; 10 mM EDTA; 1% SDS.

When performed ChIP for tap-tagged proteins (Chp1, in Chapter 4), the same procedure was
used except that, while the lysate was precleared with protein G agarose, 40 μl of IgG sepharose beads (for each sample)(Amersham), were blocked with 0.1mg/ml BSA, 1 μg of 100 bp ladder (Biolabs) and lysis buffer. The lysate precleared was then incubated overnight at 4°C with the 40 μl IgG sepharose blocked beads.

**Total protein extraction from S. pombe cells.**

10 ml of cells were grown to 5x10^6 cells/ml in YES or PMG selective medium. The same number of cells was used for each strain. Cells were spun in a benchtop centrifuge at maximum speed for 2 min. Pellet was resuspended in 1 ml dH₂O, transferred to microfuge tube, and centrifuged for further 30 sec followed by removal of the supernatant. Cell pellet was resuspended to obtain 5x10^7 cells in 100μl of 2XSB containing PMSF (50μl of 100 mM stock per ml of SB) and transferred to a screw-cap tube. An equal amount of glass beads (Sigma) was added and cells were lysed in a bead beater for 2 min at maximum speed, followed by boiling for 5-10 min. The extract was transferred to a fresh microfuge tube using a duck-bill tip to avoid glass beads. The extract was then spun and transferred to a fresh tube avoiding cell debris. 10-20 μl were loaded in SDS PAGE.

2X SB (Sample buffer) for 10 ml: 2.5 ml 0.5M Tris-HCl, pH 6.8; 2m1 Glycerol; 4 ml 10% SDS; 0.5 ml 0.1% (w/v) Bromophenol blue; 0.5 μl 2-mercaptoethanol; dH₂O. Stored at -20°C.

**Rapid S. pombe histone preparations (adapted from Ekwall et al., 1997).**

Cultures were grown in YES approximately to a concentration of 5x10^6 cells/ml. For each preparation 1x10^8 cells (20 ml of culture) were washed in 10 ml of ice cold NIB buffer and resuspended in 500 μl of NIB buffer with protease inhibitors and PMSF. 500 μl of glass beads (acid washed, 400-600 nm SIGMA) were added. Cells were lysate using the beadbeater for two minutes, keeping samples on ice all the time, good percentage (>70%) of lysis was checked at the microscope. The lysate was recovered by punctuating the bottom of the beadbeater tube and placing it on top of a new 1.5 ml tube. Both tubes were then placed in a “snap top” 15ml tube and centrifugated for 1 min at max speed in a benchtop centrifuge at 4°C. The lysate was spun in a fresh 1.5 ml tube for 10 min in microfuge at 13000 rpm at 4°C. The supernatant was discarded. The pellet was resuspended in 0.5 ml of 0.4 M sulfuric acid (Aldrich) and incubated for 1 hr on ice to extract basic proteins. The extract was then centrifuged at max speed for 5 min at 4°C and the supernatant was collected to a fresh microfuge tube. The acid extraction is repeated once, without further incubation on ice. The pooled supernatants (1 ml total) were precipitated overnight in glass Corex tubes at -20°C with 12 volumes of ice-cold acetone. The precipitate was collected by centrifugation at 7,000 rpm for 15 min at 4°C (Sorvall SS-34). The pellet was air dried the pellet under a fume hood and resuspended in 4 M urea.
The protein concentration was determined using Biorad protein assay kit at 595 nm following the manufacturer’s instructions. This method yielded 50-100 μg of protein. 3-5 up to 10 μg were used to visualise histone bands in Coomassie-stained gels and in Western blots. 2XSB was added, samples were boiled and electrophoresed on a 16% acrylamide (Prosieve) gel; pre-stained protein marker (Benchmark or Prosieve) were used to run the gel long enough (10KDa band almost to the bottom of the gel) to separate well the histone bands.

NIB buffer: 0.25 M Sucrose, 60 mM KCl, 15 mM NaCl, 5 mM MgCl₂, 1 mM CaCl₂, 15 mM Pipes pH 6.8, 0.8% Triton X-100.

2X SB (Sample buffer) for 10ml: 2.5 ml 0.5 M Tris-HCl, pH6.8; 2 ml Glycerol; 4 ml 10% SDS; 0.5 ml 0.1% (w/v) Bromophenol blue; 0.5 ml 2-mercaptoethanol; dH₂O. Stored at -20°C

SDS-PAGE (Laemmli, 1970).

Proteins were separated on 1 mm thick discontinuous SDS-PAGE (Sodium dodecyl sulphate-polyacrylamide gel electrophoresis) with the Hoefer minigel apparatus. The Prosieve 50 gel solution (FMC) was used to improve the separation of small molecular weight proteins, such as histones and histone variants. Two concentrations of resolving gel (14% and 16%) were used which should separate optimally proteins within the size range of 5-50 KDa.

Resolving gel (for 10ml):
14%: 4.5 ml dH₂O, 2.8 ml Prosieve 50; 2.5 ml 1.5M Tris-HCl, pH8.8; 0.1 ml 10% SDS; 0.1 ml 10% APS; 4 ml TEMED.
16%: 4.1 ml dH₂O, 3.2 ml Prosieve 50; 2.5 ml 1.5M Tris-HCl, pH8.8; 0.1 ml 10% SDS; 0.1 ml 10% APS; 4 μl TEMED.

Stacking gel (5%) for 5ml: 3.75 ml dH₂O, 0.5 ml Prosieve 50; 0.65 ml 1.5 M Tris-HCl, pH6.8; 50 μl 10% SDS; 50 μl 10% APS; 5 μl TEMED.

Resolving gel was poured first, followed by stacking gel. Gels were run in 1X Tris/Tricine/SDS buffer at 130V (constant voltage) for approximately 30min.

10X Tris/Tricine/SDS (1 litre): 121 g Tris Base; 179 g Tricine; 100 ml 10% SDS, dH₂O.

Western blot analysis to detect histone H3 and H4.

Proteins were transferred on Protran nitrocellulose (Schleicher & Schuell) using a Hoefer semi-dry electroblotter. The membrane was floated on dH₂O then soaked on blotting buffer, then placed on top of 6 pieces of 3MM paper of the gel size, also previously soaked in blotting buffer. The SDS gel was placed on top of the membrane followed by 6 more pieces of 3MM paper soaked in blotting buffer. As each layer was added, bubbles were rolled out using a plastic pipette. Transfer was done at the constant amperage of 65 mA for 1hr. The membrane was washed in dH₂O, followed by staining with Ponceau solution (Sigma) to verify good protein transfer. The membrane was rinsed in PBS, followed by incubation in blocking buffer for 30
minutes at RT with agitation. After a brief wash with PBS, the membrane was placed in a sealed plastic bag and incubated with the primary antibody of interest in PBS-0.2% Tween overnight at 4°C with agitation. The membrane was washed twice with water and incubated with secondary antibody HRP conjugated anti-rabbit 1:2500 in PBS-0.2% Tween for 1 hrs at RT with agitation. The blot was washed twice with water, followed by one wash with PBS-0.05% Tween for 5min, three with PBS for 5 min and by one wash with water for 5 min. The histones were revealed using the the Enhanced Chemi-Luminescence kit (Amersham) following the manufacturer's instructions. The blot was exposed to kodak film for 10min-1hr.

Blotting buffer (for 100ml): 10 ml 10X Tris/Tricine/SDS; 70 ml dH₂O; 20 ml MeOH.
Blocking buffer: 3% Marvel dried non-fat milk; 0.2% Tween in PBS.
Dilution: 1:500 for anti-rabbit α--H3C A. Verrault and α-H4(UPSTATE),
1:300 for anti-rabbit α-H4K20me1,α-H4K20me3 (Abcam).
1:250 for anti-mouse α-H4K20me2 (Abcam).

2.5 BACTERIAL METHODS.

Bacterial cells DH5α were used for all the cloning performed in this thesis. Cells were grown at 37°C in LB medium, solid or liquid, supplemented with 30 μg/ml ampicillin for plasmid selection.

Bacterial media.
LB (per litre): 10 g Bacto-peptone; 5 g Yeast extract; 10 g NaCl. Autoclaved.

Preparation of competent cells.
A single colony of DH5α cells was grown overnight in 5 ml LB. This was diluted into 100 ml LB and grown for about 2 hrs until the OD₅₅₀ reached 0.48. Cells were chilled on ice then pelleted at maximum speed for 5 min in an eppendorf centrifuge at 4°C. The supernatant was discarded and the pellet resuspended in a minimal volume of LB. 40 ml ice-cold TFB-I was added and the cells were gently mixed with a pipette. Cells were incubated on ice for 15-45 min, then centrifuged for 10 min at maximum speed at 4°C. The pellet was gently resuspended in 8 ml ice-cold TFB-II and incubated on ice for further 30 min. Cells were aliquoted into pre-chilled tubes and frozen on dry ice. Competent cells were stored at -70°C until use.
TFB-I: 30 mM K acetate, 0.1 M RbCl₂; 10 mM CaCl₂ 2H₂O; 50 mM MnCl₂; 15% glycerol; pH to 5.8 with acetic acid. Filter sterilised.
TFB-II: 10 mM PIPES; 75 mM CaCl₂ 2H₂O; 10 mM RbCl₂; 15% glycerol; pH to 6.5 with KOH. Filter sterilised.
Transformation of competent cells.
An aliquot of frozen competent bacteria was thawed on ice. 100 μl of cells was added to a 10 μl ligation mixture or to 50 ng plasmid DNA and incubated on ice for 30 min. Cells were heat-shocked for 1 min at 42°C followed by 2 min on ice. 1 ml of LB was added and cells were incubated at 37°C for 30-45min. Cells were spread on LB-agar + 30 μg/ml Ampicillin plates.

Plasmid minipreps.
2-3 ml of LB+Amp cultures of bacteria were grown overnight. The plasmids were isolated using a miniprep kit (Qiagen) following the instructions provided by the manufacturer.

2.6 GENERAL SOLUTIONS

PBS per litre: 10 g NaCl
0.25 g KCl
1.43 g Na₂HPO₄
0.25 g KH₂PO₄
Autoclaved.

TE: 1 mM EDTA
10 mM Tris-HCL pH8
Autoclaved.

20xTBE per litre: 216.0 g Tris Base
110.0 g Boric Acid
0.5 M EDTA (pH8.0) 80.0ml
CHAPTER 3

A ROLE FOR csp ts MUTANTS IN SILENT CHROMATIN FORMATION

INTRODUCTION

Correct chromosome segregation is in part ensured by proper centromere assembly (Miyazaki & Orr-Weaver, 1994; Karpen and Allshire, 1997). Active centromeres in many organisms consist of arrays of repetitive DNA and much of this centromeric repeat DNA is assembled into heterochromatin (Richard & Dawe, 1998; Sullivan et al., 2001). In the fission yeast Schizosaccharomyces pombe, centromeres are composed of outer repeat sequences (dg and dh) that flank the central kinetochore domain (Takahashi et al., 1992; Steiner et al., 1993). As in metazoans, centromere regions in fission yeast are heterochromatic since marker genes placed within the outer repeats are subjected to position effects and rendered transcriptionally silent. Mutations that disrupt the formation of this silent chromatin allow the expression of these markers and concomitantly interfere with centromere function.

The outer dg-dh repeats regions at fission yeast centromeres are packaged in nucleosomes on which the N-terminal tails of histone H3 and H4 are underacetylated (Ekwall et al., 1997). Moreover lysine 9 of histone H3 is specifically methylated by Clr4 (homolog of metazoan Suv3-9) allowing binding of Swi6 (the fission yeast counterpart of heterochromatin protein 1) and Chp1 (another chromodomain protein). Binding of Swi6 and Chp1 is mediated via their chromodomains and promotes the assembly of silent chromatin (Partridge et al., 2000; Bannister et al., 2001; Partridge et al., 2002) (Figure3-1).

Centromeric repeats are transcribed and RNA interference is required for the establishment and the maintenance of silent chromatin over the dg-dh repeats at the outer regions of centromeres (Volpe et al., 2002; Reinhart and Bartel, 2002). Components of the RNAi machinery, the Ribonuclease III (RNAselli)-like enzyme Dicer1 (Dcr1)(Fire et al., 1998; Hannon, 2002; Zamore, 2002), the PIWI/family protein that targets cognate mRNA for degradation, Argonaute1 (Ago1)(Hannon, 2002; Hammond et al., 2001; Hutvagner et al., 2002) and the RNA-dependent RNA polymerase (Rdp1) are essential for the processing of the non-coding antiparallel RNAs derived from the upper and lower strands of the centromeric repeats (Volpe et al., 2002; Reinhart and Bartel, 2002; Volpe et al., 2003). The siRNAs produced are then assembled in the RNAi effector complex, RITS (RNA-induced initiation of transcriptional gene silencing) to target specific chromosome regions by sequence-specific interactions and contribute to the formation of silent chromatin by promoting methylation of H3K9 over these repeats and binding of Swi6.
Two distinctive chromatin domains can be found in fission yeast centromere: the central domain where the histone H3-like protein CENP-A binds and the flanking repeats where silent chromatin assemble. Heterochromatin is required to recruit cohesin (Bernard et al., 2001) that mediates tight physical cohesion between sister centromeres. The outer repeats domain incorporates the dg/dh elements and a small part of the \( imr \) elements. Clusters of tRNA genes are indicated by the black thick bars in the \( imr \) elements. siRNAs complementary to the \( dh/dg \) repeats and primers used in ChIP and RT-PCR experiments are indicated.
Components of the RITS complex are Ago1, the chromo-domain protein Chp1 and a novel protein Tas3 (Verdel et al., 2004). In this chapter I will focus on a class of mutants that specifically affect silencing at centromeres: the csp ts mutants (centromere suppressor of position effect). 12 csp mutants were found in a genetic screen for trans-acting factors which are required for centromeric silencing of marker genes located within the centromere dg-dh repeat (Ekwall et al., 1999). Of these csp9, csp10, csp11, csp12 and csp13 have been shown to be involved in RNAi-directed heterochromatin formation over the repeats. csp7 and csp9 were found to be allelic respectively to rdp1 (S. White and R. Allshire personal communication) and to ago1, a component of RITS (Volpe et al., 2003; Verdel et al., 2004). The csp1 to 6 mutants are temperature sensitive (ts) and display defective silencing of centromeric marker genes at 25°C but are inviable at 36°C. These ts mutants are variable in this phenotype but in general show alleviation of silencing at the permissive temperature, high rate of chromosome loss, sensitivity to a microtubule destabilizing drugs (thiabendazole, TBZ) and chromosome segregation defects (Ekwall et al., 1999).

This is an explorative chapter that aims to further analyse the contribution of the ts mutants (csp1 tp csp6) in silent chromatin formation and assess their possible role in RNAi mediated transcriptional silencing. All the experiments presented have been reproduced more than twice. However the quantifications of the ChIP performed refers specifically to the pictures presented (Figure 3-3, 3-5).

RESULTS

Analysis of silencing phenotypes.

3.1 Alleviation of silencing.

In wild type fission yeast cells the ade6+ marker gene is silenced when inserted at the outer repeats of centromere 1 (cen1otrR(Sphl):ade6+) (Allshire et al., 1995). Silencing of this ade6+ gene, results in red colonies on low adenine supplemented plates (1/10 adenine). This phenotype is caused by the accumulation of a substrate of adenine metabolism that when exposed to oxygen, confers red colour to colonies grown on limiting adenine plates. clr4Δ cells alleviate repression, allowing the expression of ade6+ and the formation of white colonies on the 1/10 adenine plates (Figure 3-2A). The csp ts mutants result in a higher incidence of paler pink coloured colonies than wild type cells. Similar effects were observed in silencing assay with the ura4+ gene inserted in the outer repeats (cen1otrR(Sphl):ura4+) (Ekwall et al., 1999). Northern analysis to quantify the expression of a marker gene inserted at the outer repeats (cen1otrR(Sphl):ura4+) compared to an euchromatic minigene (ura4DS/E), previously showed that silencing of cen1otrR(Sphl):ura4+ was most defective in csp3 (0.77) and csp6 (0.75) relative to wild type (0.2) cells; csp4 (0.48) and csp5 (0.48) had an intermediate effect whereas csp1 (0.33) and csp2 (0.26) cells had a relatively weak silencing phenotype. csp9, that is an allele of ago1, almost completely abolishes silencing (0.92) (Ekwall et al., 1999; Volpe et al., 2002).
Figure 3-2 Monitoring silencing defects on outer repeats of csp ts mutants.

A) Colony colour assay was used to analyse the degree of defect in silencing for csp ts mutants. Cells with the ade6+ marker gene inserted at the outer repeats of centromere1 (cen10r(SphI):ade6+) were plated on low adenine supplemented YES plates (YES 1/10 ade) and grown at permissive temperature (25°C). While the ade6+ gene is repressed in wild type cells, indicated by red colour, clr4Δ cells lose this transcriptional repression and form white colonies. The mutants csp1, csp2, csp3, csp4, csp5 and csp6 all affect silencing of the ade6 gene inserted in centromeric outer repeats as indicated by pink colony colour. None of the csp ts mutants display silencing defects as strong as clr4Δ.

B) Northern analysis of RNAs from strains with a fully expressed ura4+ integrated in the genome (R.Int.), of wild type and csp mutant strains with ura4+ inserted within the otr1R repeats of centromere1. All the strains contain the ura4DS/E minigene at the endogenous ura4 locus (Ekwall et al., 1999). The cen1Rura4+/ura4DS/E ratio is indicated below each lane.
Recently, Sharon White in the lab cloned csp7 and csp10 and found that they are RNAi components, respectively rdpl and cid12. Thus the csp ts mutants exhibit a variable effect on alleviation of silencing of marker genes inserted within centromere1 at the permissive temperature and this effect is not as strong as for the csp non-ts mutants.

3.2 H3K9me2 and Swi6 binding over the centromeric repeats.

The Clr4 histone methylase is required to methylate histone H3 on lysine 9 (Rea et al., 2000; Nakayama et al., 2001) allowing binding of Swi6 (Partridge et al., 2000; Bannister et al., 2001; Partridge et al., 2002). Swi6 is in turn required for the recruitment of cohesin to mediate cohesion between sister-centromeres (Bernard et al., 2001; Nonaka et al., 2002). To assess whether these markers for heterochromatin were intact in the csp ts mutants, chromatin immunoprecipitation experiments (ChIP) were performed with anti-H3K9me2, and anti-Swi6 antibodies. The immunoprecipitated DNA was amplified using two pairs of primers complementary to centromeric outer repeats (otr) or the fbp1 euchromatic gene, which serves as a negative control. At 25°C csp1 and csp2 cells still retain H3K9me2 and Swi6 binding on the outer repeat chromatin (otrA/otrB), whereas csp3, csp4 mutants have slightly decreased H3K9me2 and Swi6 binding. The csp6 mutant is distinct since cells completely lose H3K9me2 modification and Swi6 binding over the centromeric repeats similar to cells lacking Clr4 (Figure 3-3).

Swi6 is concentrated in all the heterochromatic loci in fission yeast and it localizes in several punctuate foci in the nucleus (two to six spots) (Ekwall et al., 1995; Ekwall et al., 1997). Live analysis of wild type and csp mutant cells expressing Swi6-GFP (C-terminal tagged) was performed at the permissive temperature (25°C). In these cells swi6 is overexpressed because it is under the attenuated pREP81 nmt promoter (Pidoux et al., 2000). Wild type and csp1, csp2, csp3, csp4, csp5 and csp6 cells exhibit two to three foci in most cells. From previous analysis it is known that those foci represent clustered centromeres and telomeres (Ekwall et al.; 1995). The Swi6-GFP signal in csp1 to csp5 was similar to wild type, however the GFP foci in nuclei was consistently fainter in csp6 cells (Figure 3-4).

The discrepancy between Swi6-GFP localization and ChIP with α-Swi6, suggesting normal localization but defective association, may be related to the fact that Swi6-GFP is overexpressed allowing some residual association.

3.3 Rad21 recruitment over the centromeric repeats.

It has previously been shown that one function of Swi6 at centromeres is to ensure that a high density of cohesions (Rad21) and tight physical cohesion are maintained at centromeres prior to anaphase (Bernard et al., 2001). Swi6 deleted cells lose chromosomes at elevated rate due to loss of cohesin. csp1 to 6 had been shown to display defective chromosome segregation; to
Figure 3-3 Silent chromatin chromatin is retained over the centromeric repeats of csp ts mutants except in a csp6 mutant. Chromatin immunoprecipitation was performed on csp ts mutants at the permissive temperature (25°C) with antibodies specific for H3K9me2 and Swi6. The immunoprecipitated DNA was amplified in a multiplex PCR with pairs of primers for outer repeats (otrA and otrB) and a euchromatic gene, fbp1. csp1 and csp2 retain similar levels of H3K9me2 and Swi6 association to wild type cells. csp3 and csp4 mutants have reduced levels of these markers on the outer repeats. H3K9me2 is completely abolished in csp6 cells like in clr4 deleted cells and the level of Swi6 bound at the repeats is comparable to the one for swi6 deleted cells. Quantitation of bands was performed using the Eastman Kodak Co. EDAS 290 system and 1D Image Analysis software (n.d. = not possible to determine). This experiment has been repeated more than twice.
Figure 3-4 Swi6-GFP is still localized in the csp ts mutants.
Images of csp ts cells expressing Swi6-GFP from nmt81 promoter. Cells were grown in minimal media at permissive temperature (25°C). Wild type cells were used as a control. Only csp6 cells present a weaker signal.
determine whether the csp ts mutants also affect recruitment of cohesin, chromatin immunoprecipitation was performed on wild type, csp1 to 6 and swi6Δ cells expressing Rad21 tagged with HA at the endogenous locus and expressed from its own promoter. The presence of Rad21 over the outer repeats was quantified in a multiplex PCR using specific primer pairs to centromeric outer repeats otr or the euchromatic locus fbp1. The relative enrichment observed in wild type versus the swi6Δ control was 4 fold, confirming that Swi6 does affect the association of Rad21-HA with centromeric repeats (Figure 3-5). All the csp ts mutants affected the association of Rad21-HA with centromere repeats to some extent (2 to 4 fold). csp6 is a s defective as swi6Δ with respect to loss of Rad21-HA. The data suggest that even in csp1 and csp2 Rad21-HA recruitment to the outer repeats is not efficient despite the fact that H3K9me2 and Swi6 appear to be present on centromere repeats at same wild type levels in csp1 and csp2 cells and at intermediate levels in csp3 and csp4 cells.

In summary, the csp ts mutants have different levels of effects on the alleviation of silencing and function in silent chromatin formation. Clearly csp6 has the most severe phenotype since it loses H3K9me2 and Swi6 binding and Rad21-HA recruitment.

3.4 Centromeric repeat transcript analyses in csp ts mutants.

Small RNAs (between 21 and 24 nt) homologous to the outer centromeric repeats (dg-dh) have been detected in wild type cells (Reinhart & Bartel, 2002) and transcripts from the bottom dg-dh strand have been shown to be produced in wild type cells (Volpe et al., 2002). These transcripts are normally processed in cells by RNAi so that their steady state levels are very low, whereas they accumulate to high levels in dcr1Δ and other RNAi mutants. Top strand dg-dh repeat transcripts thought to be repressed by Swi6 heterochromatin in wild type cells, are produced in swi6 and RNAi mutants due to loss of processing of the reverse transcripts by RNAi and a failure in heterochromatin assembly as explained in Figure 3-6B (Volpe et al., 2002). To determine if the production of top and bottom strand transcripts from the dg-dh repeats is defective in the csp ts mutants, strand specific RT-PCR was performed on RNA samples extracted from csp ts cells grown at the permissive temperature. As illustrated in Figure 3-6A, first strand cDNA synthesis reactions were primed either by the forward or the reverse primers complementary to centromeric transcripts (cen for or cen rev). In addition, the control actin1 reverse (act1 rev) primer was added simultaneously. Subsequently the centromeric cen for and cen rev were used with actin 1 primers (act1) in a multiplex PCR to assess the relative amount of cen reverse or forward transcripts. These analyses indicate that the csp1, csp2, csp3, csp4 and csp5 mutants accumulate centromeric transcripts to some extent but this is a relatively weak effect compared with dcr1Δ cells. These mutants mainly affect accumulation of the top strand rather than the bottom (Figure 3-6C). It is this top strand that in wild type cells is silenced by Swi6 when heterochromatin is established while, the bottom strand is continuously transcribed but processed by RNAi (Volpe et al., 2002) (Figure 3-6B). The accumulation of the top strand again suggests that these mutants are defective in silent chromatin formation. However the
Figure 3-5 The cohesin subunit Rad21 is recruited less efficiently to centromeric repeats in csp ts mutants.
Chromatin immunoprecipitation was performed on csp ts cells expressing HA-tagged Rad21 (Rad21-HA), using anti-HA monoclonal antibody. A 2-4 fold reduction in the level of cohesin associated with centromere repeats was obtained. The level of reduction of cohesin at centromeres in csp6 cells is similar to swi6Δ cells. Primers amplify the dg repeats were used (otrA/otrB) an a euchromatic gene, fpb1 were used. This experiment have been repeated more than twice.
RNA extraction from cells grown at $10^7$ cells/ml

DNAse treatment

First cDNA synthesis using strand-specific primers instead of only oligodT

Multiplex PCR using two pairs of primers; $cen$ for- $cen$ rev / $act1$ for-$act1$ rev

Figure 3-6A RT-PCR procedure to detect centromeric transcripts.
Figure 3-6B  Model for establishment of silent chromatin at centromeres.
Centromeric repeats are transcribed. These transcripts are processed by the RNAi components, then incorporated in the RITS complex that recruits the methyltransferase Clr4. Clr4 dimethylates H3K9 that is subsequently bound by the chromodomain Swi6 and silent chromatin is established. The signal is continuously amplified by Rdpl that is part of the RDRC complex together with Cid12 and Hrr1. In RNAi mutants the starting transcripts are not processed, they accumulate and silent chromatin formation is defective.
Figure 3-6C Analysis of centromeric transcripts in \textit{csp} ts mutants. Strand specific RT-PCR analysis was performed in the presence (+RT) or absence (-RT) of reverse transcriptase in \textit{csp} ts, wild type and \textit{dcr1}\textDelta \textit{cells}. Apart from \textit{csp6} cells, all \textit{csp} ts mutants show a slightly increased in the amount of top and bottom centromeric transcripts. However, none of the mutants show as much an increase as that observed in \textit{dcr1}\textDelta \textit{cells}. The bottom strand is amplified in the first cDNA synthesis by \textit{cen} for primer while the top one is amplified by \textit{cen} rev primer (see Figure 6A).
accumulation of a small amount of the bottom strand could suggest a role for those csp gene products in directing chromatin modification via the RNAi pathway.

3.5 Production of siRNAs from centromeric repeats.
In fission yeast, Dicer (Dcrl), the RISC component Argonaute (Ago1), the RNA-dependent polymerase 1 (Rdpl) and the csp mutants, csp7-13 (non-ts), have been shown to be required for efficient silent chromatin formation over the centromeric repeats (Ekwall et al., 1999; Volpe et al., 2002, 2003). In these mutants, long unprocessed transcripts from both strands of the centromere dg and dh sequences accumulate. The long transcripts are normally short-lived since the RNAi pathway processes them into small siRNAs (21-24 nt) which are homologous to the dg-dh outer centromeric repeats. These siRNAs are loaded into the RITS complex (RNA induced initiation of transcriptional silencing) which contains Ago1, the chromo-domain protein Chpl, and an uncharacterized protein, Tas3. Cells lacking Dicer activity (DcrlA) fail to make siRNA and thus they are not detected in RITS (Verde et al., 2004). Ago, Chpl and Tas3 have been shown to associate with centromeric repeats by ChIP in a wild type but not in a dcr1Δ strain (Verdel et al., 2004). Thus the RITS complex appears to direct the formation of heterochromatin over sequences that are complementary to the loaded siRNA. To further investigate the role of the csp genes in processing centromeric transcripts and in centromeric siRNAs production, northern analysis of small RNAs prepared from the csp 1, csp2, csp3, csp4, csp5 and csp6 cells was performed.

Approximately 40 μg of small RNAs from extracts of all csp ts mutants grown at 25°C were loaded on a 17.5% polyacrylamide 7M urea gel. Ethidium bromide stain of the gel prior to transfer demonstrated that equivalent amount of RNAs were loaded. The resulting gel was probed with a region from dh repeat that was amplified using cen for and siRNAH primers, was used (Figure 3-1). As expected siRNAs were readily detected in wild type but not in dcr1Δ cells. csp2 cells retain cen siRNAs suggesting that the RNAi pathway is intact. In contrast cen siRNAs are not detectable in csp1 and csp6 cells; csp3, csp4 and csp5 may retain a low level of siRNAs (Figure 3-7).

DISCUSSION
The data collected on the csp ts mutants are summarized in the Table1. Some of these genes encode essential proteins and thus they may influence heterochromatin formation indirectly by affecting other pathways such as general RNA metabolism. In the next two chapters I will focus my analysis on csp3, csp4, csp5 and csp6 in order to understand how they affect silent chromatin formation. Below I summarize what it is known about each mutant.

csp1
This mutant does not have a strong silencing defect. It maintains the two features of silent chromatin at the centromere: H3K9me2 and Swi6 binding. However it recruits less cohesin subunit Rad21 than wild type cells. This phenotype may be due to a problem of Rad21

63
Figure 3-7 Analysis of siRNAs from centromere repeats.
Small RNAs from dcr1Δ, csp1, csp3, csp2, csp4, csp5, csp6 and wild type cells were resuspended in 50% formamide and separated on a 17.5% urea-denaturing polyacrylamide gel. RNAs were transferred by electroblotting to a nylon filter that was hybridized with a 32P DNA probe, complementary to centromeric dh repeats, generated with Random Primed DNA Labeling Kit (Roche). Phosphorimage was exposed for minimum 3 hours. A DNA oligonucleotide homologous to centromeric repeats, of 22 nt length, was used as size marker.
recruitment being less efficient despite Swi6 presence. The major defects are related to the RNAi pathway: it has a slight defect in processing centromeric transcripts, in particular the bottom strand. This strand as shown in the Figure 3-6B, in a wild type cell, it is continuously produced and processed by the RNAi pathway. This phenotype is in accordance with the fact that cen siRNAs are not detectable by northern analysis. All these data agree with a possible role of Csp1 more in Post-Transcriptional Gene Silencing (PTGS) or siRNAs production than in the Transcriptional Gene Silencing (TGS) or Swi6-based chromatin silencing. However, being an essential gene, it could be that it is involved in general RNA metabolism and this may affect indirectly the RNAi pathway. We do not know what csp1 encodes for and I tried once its cloning by complementation but without success.

csp2
This mutant exhibits the same phenotypes as wild type cells for all the features relative to silent chromatin examined (H3K9me2, Swi6 localization, centromeric transcripts processing and siRNAs production). The silencing defect is really weak as seen from the level of transcription of the marker gene ura4 inserted between the outer repeats (Ekwall et al., 1999). The only defect found is related to the levels of the cohesin subunit Rad21 bound at the outer repeats. They are 3 fold less than in wild type cells. Like for csp1 this phenotype could be due to inefficient recruitment of Rad21 despite the presence of Swi6 and suggest a possible complementary pathway for the delivery and/or stabilization of cohesin at centromeres.

csp3
This mutant shows a defect in silent chromatin formation, decrease in H3K9me2 and Swi6 binding over the outer repeats. These defects are less pronounced than in a clr4Δ mutant. Data related to centromeric transcripts processing and cen siRNAs production suggest that the assembly of silent chromatin is not so efficient possibly due to less cen siRNAs which targets Clr4 to methylate H3K9. csp3 cells accumulate the TOP strand transcript more than the BOTTOM one even if not at the same extent as dcr1Δ cells. This suggests that the RNAi pathway is intact and what is defective, it can be an upstream process, like for example, the initial transcription that then determines deficiency in silent chromatin formation. Moreover the decrease in Rad21 bound to the outer repeats is in accordance with less presence of Swi6. csp3 has been cloned by Karl Ekwall and it is an allele of rbp7. Rpb7 is a conserved subunit of RNA pol II. Its role in silent chromatin formation and its further characterization will be the topic of the next chapter.

csp4
This mutant has an intermediate phenotype relative to silencing defects: it has decreased levels of H3K9me2 over the centromeric repeats and less Swi6 is associated at the centromere. The cohesin subunit Rad21 is less present probably as an effect of less Swi6 bound over the repeats. These defects are less pronounced than in csp3 cells. csp4 cells partially accumulate both the top and the bottom centromeric transcripts. Northern analysis of cen siRNAs reveals a defect in their production: possibly in csp4 cells the RNAi pathway is compromised and this
influences proper silent chromatin formation. *csp4* has been cloned by Karl Ekwall and it is an allele of *cwf10*, a gene encoding for a splicing factor ortholog to Snu114p in *S. cerevisiae*. Snu114p is involved in unwinding U4/U6 RNA during activation of the spliceosome in budding yeast. Csp4 could affect the RNAi pathway directly or indirectly by being involved in general RNA metabolism. Further characterization of this mutant will be the topic of one of the next chapters (Chapter 5).

**csp5**

This mutant by northern analysis, can produce *cen* siRNAs but in lower amount than wild type cells. As for *csp3* cells the defective silencing phenotype could derive from the initial step in the production of the centromeric RNAs or in the subsequent one, in their processing. In fact *csp5* is an allele of *prp39*, a U1 associated protein involved in pre-mRNA splicing and again as for *csp4*, it could affect directly or indirectly the RNAi pathway. Its cloning will be described in the same chapter as for *csp4* (Chapter 5).

**csp6**

This mutant has the most dramatic phenotype relative to silencing defects, H3K9me2 and Swi6 localization. It behaves like *clr4Δ* cells losing completely all the markers of silent chromatin. From these data it seem to be defective in some components of silent chromatin more than in the RNAi pathway. However, despite the fact that it does not accumulate any transcripts (top or bottom), it does not also produce any *cen* siRNAs by northern analysis. The only possible explanation for these two results is that in *csp6* cells *cen* transcripts are not made. In this scenario *csp6* cells are compromised in the first step for the RNAi pathway to act and trigger silent chromatin formation. It may be that in this context the RNAi pathway is perfectly functional and *csp6* cells are deprived of some components necessary for specific centromere transcription. Many attempts were made to try to clone this gene but the only suppressor recovered was *hsp70*, a gene encoding for a chaperone protein. For its interesting phenotype relative to silencing and also to cell cycle defects, found by Karl Ekwall in the initial screen (Ekwall et al., 1999), the efforts to clone it and its initial mapping are presented in Chapter 6.

From this initial analysis we can divide the functions of the *csp* ts mutants based on their action at three levels: the transcriptional gene silencing (TGS/chromatin, DNA modifier), the post-transcriptional gene silencing (PTGS/ RNA degradation), and transcription of the centromeric transcripts (Figure 3-8). This distinction is mostly based on the results obtained from the northern analysis and the ability of each mutant to produce *cen* siRNAs. Csp1 and Csp4 probably function in the RNA degradation while Csp3, Csp5 function upstream, at the transcriptional level. Csp6 is a particular mutant whose function can be placed very upstream of everything else: it loses all the heterochromatic features, it cannot produce *cen* siRNAs but also it does not accumulate centromeric transcripts. This suggests a problem even with their production.
In conclusion the csp ts genes may encode for factors involved in general RNA metabolism and RNAi pathway. The fact that they are essential genes and the mutants that so far have been found to affect silencing at the outer repeats are non essential genes suggest two main models.

1) Their defect in silencing is an indirect result of a general compromised RNA metabolism function or

2) certain general RNA metabolism factors have specific and direct function or interaction with the RNAi components.

In Chapter 4, relative to csp3 mutants, data will be presented to validate the second model and explain its direct role. In Chapter 5, relative to csp4 and csp5 data will be presented to try to understand how splicing factors might interfere with silent chromatin formation.

<table>
<thead>
<tr>
<th></th>
<th>Silencing defect</th>
<th>H3K9me2 localization</th>
<th>Swi6 localization</th>
<th>cen transcripts</th>
<th>siRNAs production</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt</td>
<td>-</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
<td>++</td>
</tr>
<tr>
<td>csp1</td>
<td>-</td>
<td>+++</td>
<td>+++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>csp2</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>csp3</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>csp4</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>csp5</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>csp6</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>cIr4Δ</td>
<td>+++</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>dcr1Δ</td>
<td>+++</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>csp9</td>
<td>+++</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1 Summary of the phenotypes of the csp ts mutants.

Silencing defects are relative to the expression of a marker gene inserted at the Sphi site of the outer repeats. None of the csp ts mutants have such strong derepression as cIr4Δ cells or dcr1Δ cells. The strongest defect is apparent present in csp3 and csp6 cells.

H3K9me2 over the centromeric repeats is completely lost in csp6 cells as well as in cIr4Δ. csp3 cells like dcr1Δ and csp9 cells lose partially this methylation. csp4 and csp5 cells have some reduction of H3K9me2 while csp1 and csp2 cells exhibit the same levels of methylation as wild type cells. Swi6 localization is completely in accordance with the H3K9me2 results: it is lost in csp6 cells and cIr4Δ cells. Its presence is partially decrease in csp3 like in dcr1Δ and csp9 cells. csp4 and csp5 cells have reduced levels of Swi6 over the centromeric repeats while in csp1 and csp2, Swi6 is present as in wild type cells.

Centromeric transcripts partially accumulates in all the csp ts mutants except for csp6 cells where they do not accumulate as in wild type cells. The amount that accumulates in the rest of the csp ts is not as much as in cIr4Δ, dcr1Δ and csp9 cells.

siRNAs production is defective in csp1, csp4 and csp6 cells like in cIr4Δ, dcr1Δ and csp9. csp3 and csp5 cells are able to produce some siRNAs but less than wild type or csp2 cells.
Figure 3-8 Chromatin and RNAi: possible role of csp gene products. The csp ts mutants may be involved in:

1) centromere transcription
2) siRNAs production
3) events that are concomitant with or downstream of the transcription process and the RNAi pathway
CHAPTER 4
Rpb7, A SUBUNIT OF RNA POL II THAT IS INVOLVED IN CENTROMERIC siRNAs PRODUCTION

INTRODUCTION

In the RNA interference (RNAi) pathway, long double stranded RNA molecules (dsRNAs) are processed into small interfering RNAs (siRNAs) of 21-24 nt length to trigger the destruction of homologous target mRNAs. The RNAi pathway can also induce DNA methylation and/or H3K9 methylation at homologous genomic loci that result in transcriptional silencing in Arabidopsis thaliana (Baulcombe, 2004; Lippman et al., 2003; Jones et al., 2001; Mette et al., 2000), S. pombe (Volpe et al., 2002; Lippman & Martienssen, 2004), Tetrahymena (Mochizuchi et al., 2002; Mochizuchi et al., 2004), Drosophila melanogaster (Pal-Bhadra et al., 2004) and mammalian cells (Kawasaki et al., 2004; Morris et al., 2004). Dicer, an RNAse III-containing enzyme, Argonaute, a PAZ/PIWI-domain containing protein which is a component of the RISC complex (RNA- induced silencing complex), and RNA-dependent RNA polymerase (Rdrp), not present in Drosophila and mammalian cells, are all necessary components of this pathway. They are involved in a variety of gene-silencing phenomena at the transcriptional level (TGS), post-transcriptional (PTGS), or translational level, depending on the organism.

In metazoa and plants there are frequently several Dicers (2 in Drosophila, 4 in Arabidopsis) and many different Argonaute proteins (5 in Drosophila, 10 in Arabidopsis), making it difficult to determine their functions. However in S. pombe only one gene encodes for Dicer, Argonaute, Rdrp allowing comprehensive and informative series of studies on RNAi-mediated TGS. Endogenous RNAs of 21-24 nt length (siRNAs) corresponding to the outer centromeric repeats (otr) have been examined (Reinhart & Bartel, 2002). It has been shown that these centromeric siRNAs associate with an RNA-induced initiation of transcriptional gene silencing (RITS) complex (Verdel et al., 2004) and this appears to induce H3K9 methylation, followed by recruitment of Swi6 (an HP1 homolog) and cohesin to centromeres and mating type loci (Bernard et al., 2001; Nonaka et al., 2001; Hall et al., 2002; Volpe et al., 2002, 2003). Mutations within this pathway affect centromere function and manifest themselves as defects in sister-chromatid cohesion, chromosome segregation and ultimately in aneuploidy (Volpe et al., 2003). Moreover fission yeast has centromeres with structure that resembles that of metazoa where repetitive sequences (pericentromeric heterochromatin) surround a central domain where the
kinetochores are assembled. Genes inserted within the repeats are assembled into transcriptionally silent chromatin which has properties of heterochromatin (Allshire et al., 1995). Recently it has been found that Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing, suggesting that a similar mechanism to induce heterochromatin in mammalian cells could exist (Kanellopoulou et al., 2005).

How the siRNAs complementary to the cen repeats gain access to DNA and recruit DNA and chromatin modifying activities to assemble silent chromatin is still not clear. In plants this paradoxical involvement of an RNA silencing pathway in maintenance of transcriptional silencing for some repeats has been solved by the existence of a distinct silencing polymerase, RNA pol IV in addition to the well-known DNA-dependent RNA polymerase I, II and III (Herr et al., 2005; Onodera et al., 2005). The mechanism by which it works it is not completely clear. A possible model suggests that heterochromatic regions might not normally be transcribed from DNA. Rather, pol IV and RdRp could act in tandem to propagate mitotically transmissible dsRNAs from which siRNAs are generated to help maintain heterochromatic silencing (Vaughn & Martienssen, 2005). In animals and fungi, the Pol IV homologue is absent but the silencing paradox may be solved if there are forms of RNA pol I-III that associate with silencing-specific subunits and allow transcription of heterochromatic DNA and, consequently, maintenance of the RNA-dependent silencing. Alternatively the presence of siRNAs may alter the rate of RNA polymerase allowing it to act as a platform for the assembly of silent chromatin.

In this chapter I will focus on csp3. As mentioned in the previous chapter, csp3 was isolated by Karl Ekwall in a genetic screen for trans-acting mutants that alleviate silencing at the outer repeats of centromere 1 (Ekwall et al., 1999). It is one of the six temperature sensitive genes isolated together with the other five non-temperature sensitive genes (csp9 to csp13) that have been found to be involved in RNAi directed heterochromatin formation over the repeats (Volpe et al., 2003). csp3 has been cloned by Karl Ekwall and found to be an allele of rpb7, that encodes a very highly conserved and essential subunit of RNA pol II (Sakurai & Ishihama, 1997; Mitsuzawa et al., 2003). I will present data to support the idea that the function of Rpb7, is to initiate the transcription of the centromeric repeats that become targets of the RNAi pathway to direct silent chromatin formation. This work was performed in collaboration with the Ekwall lab and some their data will also be presented.

RESULTS

4.1 Cloning, structural and functional characterization of rbp7-G150D mutation.

The csp3/rpb7 gene was cloned by complementation of the csp3 temperature sensitive phenotype by Karl Ekwall. Sequencing of the csp3/rpb7 allele revealed a point mutation that substitutes a glycine with an aspartic acid at position 150 (G150D) (Karl Ekwall). The mutation lies within an RNA binding motif (Todone et al., 2001) that is conserved in S. cerevisiae and human Rpb7.
The N-terminal domain of Rpb7 is composed of 4 \( \beta \) sheets (A1, A2, A3, A4) (Figure 4-1A). In the secondary structure they wrap around a helix (K2). The C-terminal half of the protein folds into five-stranded antiparallel \( \beta \) barrel (B1, B2, B3, B4, B5) with an OB-fold topology, known as S1 motif. The \( rpb7^{-} \) G150D mutation is in the S1 motif. This motif has been found in proteins that bind single-stranded RNA \textit{in vitro} (Subramanian, 1983). The unusual feature of the S1 motif in Rpb7 is that the B3 and B4 strands are separated by a long insertion where the short helix (K3) (always present) is followed by a three stranded antiparallel \( \beta \) sheet (C1-C3) (Todone et al., 2001). From structural data, the S1 motif of Rpb7 could interact with the nascent RNA transcripts (Todone et al, 2001). This may be important during the initial step of the transcription process as previous studies indicate a role for Rbp7 in initiation of transcription (Edwards et al., 1991).

RNA pol II is a large multi-subunit complex that is responsible for the synthesis of all eukaryotic mRNAs. The best studied RNA pol II is the yeast complex, consisting of 12 subunits (Rpb1-12) (Sakurai et al., 1997; Williams et al.,1996). The tertiary structure and the RNA pol II context is illustrated in Figure 4-1B. The crystal structures of the 12 subunits of yeast pol II have recently been resolved (Greenleaf, 2003; Bushnell & Kornberg, 2003), revealing two distinct structures: a ten-subunit core, which constitutes the bulk of pol II structure that includes the catalytic active site; a two-subunit complex comprising Rpb4 and Rpb7. The additional constituent of pol II, the C-terminal domain (CTD) of Rpb1, is not seen in the crystal structure because it is unstructured. The CTD is very important during the different stages of the transcription cycle, by acting as a platform for the different proteins that affect both transcription and processing of the emerging transcript. The Rpb4/Rpb7 structure is located near the transcript–exit groove and adjacent to the CTD linker region. Rpb4/Rpb7 complex can dissociate from pol II in a reversible manner and this has been used to show \textit{in vitro} that it is required for promoter-directed initiation of transcription (Edwards et al., 1991; Orlicky et al., 2001).

Purification of the 12 RNA pol II subunits carried out in the Ekwall lab, was performed in wild type and \( rpb7-G150D \) cells grown at the permissive temperature as previously described (Spahr, 2003). No difference was found in the composition of wild type and mutant cells by silver staining analysis. The result was confirmed by mass spectroscopy (Spahr & Gustafsson data not shown). Basal transcription in an \textit{in-vitro} assay at permissive temperature was slightly reduced in \( rpb7-G150D \) cells compared to wild type (Spahr & Gustafsson data not shown). However, the affinity of RNA pol II for binding ssRNA \textit{in vitro} was higher in \( rpb7-G150D \) than in wild type cells both at permissive and restrictive temperature (Spahr & Gustafsson data not shown). From this first functional analysis we can conclude that the subunit composition of RNA pol II in \( rpb7-G150D \) cells is not defective at permissive temperature (25°C).

### 4.2 Silencing defect analysis.

In the previous chapter, in the analysis of \( csp \) ts phenotypes with respect to silencing, \( csp3/rpb7-G150D \) cells showed quite strong alleviation of silencing of marker genes (\textit{ade6} \( ^{+} \) and
Figure 4-1 Rpb7 from budding yeast, fission yeast and human.
A) Alignment of Rpb7 from the three species was done using CLUSTAW. The conserved amino acids, present in all the three species are indicated by red stars. The β-sheet (A), β-barrel (B), antiparallel β-sheet (C) and α-helix (K) structures, predicted by Todone et al, 2001 from the ortholog amino acid sequence from archea Methanococcus jannaschii E subunit (RpoE), are drawn at the top of each alignment.

B) Structure of RNA pol II and location of all its subunits. A ribbon representation of the pol II structure of the top and back view is shown. The two encircled regions indicate the core and the Rpb4/7 heterodimer. The subunit 1 and 2 form opposite sites of a central cleft where the template DNA accommodates and is located the active center. The cleft is in between a protein wall and a mobile clamp, the mobility of which depends on Rpb4/7. The N-terminal half of Rpb7, which contains the RNP domain (a conserved fold involved in RNA-binding) is dark blue, the C-terminal half of Rpb7, which contains the oligo-binding fold is light-blue (Choder, 2004)
ura4\textsuperscript{*} inserted at the outer repeats compared to the other csp ts mutants. However, rpb7-G150D cells partially retained features of silent chromatin over the outer repeats of centromere 1 (H3K9me2, Swi6) (see Table 1 Chapter 1). Defective Chp1 or any other protein involved in the RNAi pathway alleviate silencing and abolish H3K9me2 on the ura4\textsuperscript{*} reporter gene inserted in centromeric outer repeats (otr) (Partridge et al., 2002; Volpe et al., 2002; Verdel et al., 2004; Motamedi et al., 2004) but the methylation level on centromeric dg-dh repeats just reduces 3 fold (Volpe et al., 2002; Sadaie et al., 2004). This could suggest that RNAi is required to establish but not to maintain all H3K9me2.

To investigate in more detail whether rpb7-G150D cells behave like chp1 and RITS mutants, chromatin prepared from a strain in which ura4\textsuperscript{*} gene was inserted into centromeric otr repeats (otr1(SphI):ura4\textsuperscript{*}) was immunoprecipitated with anti-H3K9me2 and anti-Swi6 antibodies. Subsequently, semiquantitative PCR was used to determine whether the cen1-ura4\textsuperscript{*} vs the euchromatic minigene ura4DS/E was enriched in IP's. rpb7-G150D cells lose H3K9me2 to the same extent as clr4\Delta cells from otr1:ura4\textsuperscript{*} (Figure 4-2A). In contrast, on otr/dg repeat sequences, H3K9me2 level is reduced in rpb7-G150D cells but the levels are clearly higher than in clr4\Delta cells (Figure 4-2B). Similar results were obtained when chromatin was immunoprecipitated with an antibody against Swi6 (Figure 4-3). rpb7-G150D cells lose Swi6 binding (Figure 4-3A) whilst the levels of Swi6 bound to the otr/dg locus is just decreased in rpb7-G150D cells relative to wild type (Figure 4-3B). These data demonstrate that H3me2K9 and Swi6 association with outer repeats are maintained, at least in part, in rpb7-G150D mutant, while H3K9me2 and Swi6 association with otr1(SphI):ura4\textsuperscript{*} apparently depends on rpb7\textsuperscript{*}. The differential reduction of H3K9me2 and Swi6 levels observed on otr/dg sequences versus otr1(SphI):ura4\textsuperscript{*} in rpb7-G150D cells suggests that Rpb7 is involved (like Chp1 and other RNAi components) in spreading H3K9 methylation and Swi6 binding on non-centromeric DNA (ura4\textsuperscript{*} gene) in the centromeric repeats.

4.3 Rpb7 functions upstream of Dcr1.

siRNAs homologous to centromeric dg-dh repeats have been cloned in wild type cells (Reinhart & Bartel, 2002) and transcripts from the top and bottom strands of dg-dh are detected (Volpe et al., 2002). These transcripts are rapidly processed in wild type cells, whereas they accumulate to high level in dcr1\Delta or RITS mutant cells. Top strand dg-dh transcript is repressed by recruitment of Swi6 to centromeric repeats in wild type cells. This top strand transcript is detected in dcr1\Delta, RITS and clr4\Delta mutants probably as a secondary effect of loss of processing of the bottom strand and consequently less heterochromatin formation. Consistent with this, Swi6 mutants predominantly accumulate top strand transcript. To determine if silencing and processing of top and bottom dg transcripts were affected in rpb7-G150D mutants, strand specific RT-PCR was performed (as described in the previous chapter) on wild type, rpb7-G150D, rpb7-G150D dcr1\Delta and dcr1\Delta cells. As shown in Figure 4-4A, both bottom and top
Figure 4-2  

Chromatin Immunoprecipitation was performed with anti-H3K9me2 antibody. The DNA immunoprecipitated was amplified using specific primers for *ura4* gene. These primers amplify the euchromatic *uraDS/E* minigene, used as control and the *ura4* gene inserted in the centromeric *dg-dh* repeats (*otr1*(SphI):*ura4*+) in a competitive PCR reaction as shown in the diagram (A). Primers specific for *dg-dh* repeats (*otrA/otrB*) and the euchromatic gene *fbp1* were used (B). The loss of H3K9me2 is more pronounced on a marker gene inserted in the *dg-dh* repeats than on the *dg-dh* repeats themselves. The numbers indicate the mean and standard deviations of three experiments.
Figure 4-3 *rbp7-G150D* cells lose Swi6 binding over a marker gene inserted in the *dg-dh* repeats.

Chromatin Immunoprecipitation was performed with an antibody specific for Swi6. The DNA immunoprecipitated was amplified using specific primers for *ura4* gene. These amplify the euchromatic *uraDS/E* minigene, used as control and the *ura4* gene inserted at the heterochromatic *dg-dh* repeats (*otr1(SphI):ura4*) in a competitive PCR reaction (A). Primers specific for *dg-dh* repeats (*otrA/otrB*) and the euchromatic gene *fbp1* were used (B). Swi6 is lost from a marker gene inserted at the *dg-dh* repeats but only reduced in levels over the *dg-dh* repeats themselves. The numbers indicate the mean and standard deviations of three experiments.
strand dg-dh transcripts were detectable in dcr1Δ cells and undetectable in wild type cells as expected (Volpe et al., 2002). rpb7-G150D cells accumulate some top and bottom transcript but to a lesser extent than seen in dcr1Δ cells. The rpb7-G150Ddcr1Δ mutant accumulates less top strand transcript than dcr1Δ cells but much further less (half the amount) bottom strand than dcr1Δ cells. This result suggests that rpb7-G150D cells are mainly defective in the production of the bottom dg-dh transcripts. Similar results were found by Northern analysis of the centromere transcripts performed by Ingela Djupedal (Ekwall lab). These data suggest that Rpb7 acts upstream of Dcr1. However, rpb7-G150Ddcr1Δ cells also show a higher level of accumulation of both dg-dh top and bottom transcripts compared to wild type. It is possible that this represent a secondary effect due to less silent chromatin formation. These data suggest that Rpb7 is required for the efficient generation of bottom centromeric transcripts in a heterochromatic environment, and since it is a subunit of RNA pol II, it also suggests that RNA pol II is responsible for the synthesis of the dg-dh transcripts.

4.4 RNA pol II binds dg-dh repeats.

To test directly if RNA pol II associates with centromeric dg-dh repeats, an antibody that recognizes the C-terminal domain of the large catalytic subunit Rpb1, was used for chromatin immunoprecipitation in wild type, rpb7-G150D, dcr1Δ and rpb7-G150Ddcr1Δ cells grown at permissive temperature (25°C). The immunoprecipitated DNA was amplified using pair of primers complementary to dg-dh repeats (cen for/cen rev) and to a region encompassing the Ser-tRNA gene and Met-tRNA gene. Since tRNAs genes are transcribed by RNA pol III it is expected that Rpb1/Pol II should be absent from this region and therefore it should provide a negative control. A second PCR using primers specific for an actively transcribed gene, act1 was performed to validate the ChIP (Figure 4-5). These analyses demonstrate that RNA pol II is associated with dg-dh repeats. The levels of RNA pol II present at the dg-dh repeats in dcr1Δ and rpb7-G150Ddcr1Δ cells are the same as observed in wild type and rpb7-G150D cells. This suggests that Rpb7 is not required for RNA pol II recruitment to these centromere repeats and RNA pol II-centromere association is Dcr1-independent.

A hallmark of RNA pol II transcripts is 3’ end polyadenylation. 3’-RACE (Rapid Amplification of cDNA Ends) mapping of centromeric dg-dh transcripts was performed on RNA extracted from dcr1Δ cells and the resulting sequence aligned with genomic DNA sequences: the bottom transcripts are indeed polyadenylated (Ingela Djupedal data not shown, Ekwall lab). The association of RNA pol II with the dg-dh repeats and the fact that there are polyadenylated species complementary to the dg-dh repeats support the idea that centromeric transcripts are synthesized by RNA pol II.

Transcriptional run-on (TRO) experiments on wild type cells revealed that the top strand does not accumulate, presumably because it is repressed via Swi6. In contrast, the bottom strand is frequently transcribed but rapidly turned over by RNAi (Volpe et al., 2002).
Figure 4-4 Rpb7 acts upstream of Dcr1 in the production of the dg-dh bottom strand.

Strand specific RT-PCR analysis of dg-dh repeat transcripts was performed on wild type, rpb7-G150D, rpb7-G150DdcrlΔ and dcrlΔ cells grown at 25°C. Top and bottom dg-dh transcripts were analysed as indicated in the accompanying diagram of cenl dg-dh repeats. Actin serves to show that same amount of RNA template was present in each sample after the first strand cDNA synthesis. The primers used in this reaction are indicated at the top of figure 4-4B.
Many factors can influence RNA levels, for example defective splicing and 3'end processing can result in unscheduled turnover. Alternatively, defects in RNA production may reduce the amount of ongoing transcription. To test whether the altered *dg-dh* transcript levels in *rpb7-G150D* cells was due to a change in the amount of ongoing synthesis, transcription run on analysis at 25°C in *rpb7-G150D* cells, of top and bottom transcripts was also performed by the Ekwall lab. This analysis showed a reproducible 2-fold reduction in bottom transcript synthesis and a 1.5 fold increase in top transcript synthesis in *rpb7-G150D* cells versus wild type. This is consistent with the interpretation for the RT-PCR and northern data and support the idea that the amount of on going transcription to produce the bottom *dg-dh* transcript is reduced in *rpb7-G150D* cells. This could due to defective initiation or impaired elongation.

In *dcr1Δ* cells the processing of centromeric transcripts to siRNAs is defective so that centromeric transcripts accumulate and *dg-dh* siRNAs are no longer detected in RITS (Volpe et al., 2002; Verdel et al., 2003). Northern analysis of *dg-dh* siRNAs was performed on RNA prepared from wild type, *rpb7-G150D, dcr1Δ* and *rpb7-G150Ddcr1Δ* RNA cells grown at 25°C (Figure 4-6). As expected wild type cells produce *dg-dh* siRNAs, whereas *dcr1Δ* cells do not (Verdel et al., 2004; previous chapter). A very faint signal is visible in *rpb7-G150D* cells, indicating that the levels of *dg-dh* siRNA are dramatically reduced. No *dg-dh* siRNAs could be detected in *rpb7-G150Ddcr1Δ* double mutant. Equal amounts of small RNAs (~40 µg) were loaded for each sample and this is indicated by ethidium bromide staining, following transfer. These data again supports the idea that Rpb7 acts upstream of Dcr1 in the RNAi pathway and is involved in producing the substrate RNAs on which Dicer cuts.

4.5 Rpb7 is not defective in general transcription at permissive temperature.

TRO analysis of *rpb7-G150D* cells revealed a defect in the amount of ongoing transcription across the *dg-dh* repeats. This defect might derive from a problem in initiation or transcriptional elongation. In addition, defects in termination could lead to unscheduled turnover of the defective transcripts resulting in reduced levels of steady state transcripts. To examine the possibility that *rpb7-G150D* cells are defective in transcription elongation or transcription termination, the following experiments were performed.

6-Azauracil inhibits transcriptional elongation by depleting the intracellular pool of GTP and UTP (Exinger et al., 1992). In *S. cerevisiae* mutants defective in elongation such as Rpb1 and Rpb2, the two largest subunit of RNA pol II display increased sensitivity to 6-Azauracil (6AU) (Archambault et al., 1992; Lennon et al., 1998). *rpb7-1* cells were tested and growth was unaffected on plates containing 100 µg/ml 6AU at 25°C. This indicated that *rpb7-G150D* has no severe general transcription elongation defect at the permissive temperature (Djupedal Ingela data not shown).

Defects in the transcription termination step were investigated using previously described tester strains (Aranda & Proudfoot, 2001). These strains have a *ura4* gene engineered to contain an
ChIP with α-CTD

Figure 4-5 RNA pol II is associated with the dg-dh repeats.

Chromatin Immunoprecipitation was performed using a specific antibody for C-terminal domain of RNA pol II (Covance) on wild type, *rpb7-G150D, dcr1Δ* and *rpb7-G150Ddcr1Δ* cells grown at 25°C. The DNA immunoprecipitated was amplified with specific primers for *dg-dh* repeats (*cen for/ cen rev*) and a region between Ser-tRNA and Met-tRNA (*tRNA*) gene as a control. RNA pol II binding at the centromeric repeats is not Rpb7 or Dcr1-dependent at 25°C (A). The DNA immunoprecipitated was also amplified with primers specific for an euchromatic gene, *act1* as a control (B).
Figure 4-6 *rpb7-G150D* cells have dramatically reduced *dg-dh* siRNAs. Phosphorimage show northern analysis of PEG precipitated small RNA extracted at 25°C from wild type, *rpb7-G150D*, *dcr1Δ* and *rpb7-G150D dcr1Δ* double mutant cells. The northern was probed with *dg-dh* centromeric probe prepared by random priming of the fragment amplified by *cen for* and *dhH* siRNAs as indicated in the flanking diagram of centromere 1. The ethidium bromide stained gel after the transfer is shown to demonstrate that overall loading was similar.
efficient transcription termination module inserted in an artificial intron within the ura4 gene (Figure 4-7A). In the Ter<sup>wt</sup> strain this transcriptional terminator causes greater than 99% of ura4 transcripts to prematurely terminate so that cells do not produce full length ura4 transcript and are thus unable to grow on media lacking uracil. Defects in transcriptional termination due to a trans-acting mutation or a cis-acting mutation in the terminator (Ter<sup>M5</sup>), allow transcriptional read through, expression of full length ura4 transcript and growth on plates lacking uracil (Aranda and Proudfoot, 2001) (diagram Figure 4-7B). These ura4<sup>+</sup> constructs inserted at the nmt1 locus, were combined with the rpb7- G150D mutation. In addition, the strains contain the ura4-DS/E allele as an internal control. Six Ter<sup>+</sup> rpb7-G150D isolates were tested for growth on media lacking uracil and the presence of ura4 transcripts assessed by RT-PCR in two of these (Figure 4-7C). The Ter<sup>wt</sup> strain does not grow on plates lacking uracil and transcript does not proceed beyond the Ter<sup>+</sup> terminator in the ura4 gene. In contrast, the control strain Ter<sup>M5</sup>, grows on plates lacking uracil and read through is shown by production of ura4 transcripts by RT-PCR. This is consistent with the observation that 75% of transcripts read through the M5 terminator (Aranda & Proudfoot, 2001). Ter<sup>+</sup>rpb7-G150D cl.2 and Ter<sup>+</sup>rpb7-G150D cl.6 behave as wild type cells in that they do not grow on media lacking uracil and do not express full length ura4 transcripts. Thus the rpb7- G150D mutation generally does not affect transcriptional termination in this assay. It seems that termination of centromeric transcripts is also normal as no downstream products had been found by RT-PCR. To determine if the rpb7-G150D mutation generally affects transcription and to investigate if other genes in the genome are affected by this mutation at 25°C and 36°C, the Ekwall lab performed expression profiling by microarray analyses on RNA from rpb7-G150D cells (Xue et al., 2004). The data generated did not reveal any defect in gene regulation at 25°C. In contrast, the majority of genes were down regulated (>2 fold) after one hour at 36°C. Expression of genes with known functions in centromere silencing was not significantly altered at 25°C. Thus rpb7- G150D cells are generally defective for transcription at the restrictive temperature (36°C) but display no general defect at 25°C. Taken together these data enforce the idea that rpb7-G150D cells display a specific defect in synthesis of dg-dh transcripts at 25°C.

4.6 Rpb7 specifically promotes transcription of the bottom dg-dh strand.

Previous studies indicate a role for Rpb7 in initiation of transcription. Together with Rpb4, Rpb7 forms a dissociable substructure of RNA pol II (Choder, 2004) and interacts strongly in protein-DNA photo-crosslinking experiments with promoter DNA (Chen et al., 2004). Moreover RNA pol II lacking Rpb4/7 is inactive in promoter-directed initiation of transcription and addition of Rpb4/7 restored activity in yeast (Edwards et al., 1991). To test if transcriptional initiation from a putative promoter in the centromere was defective in rpb7-G150D cells, the 5' end of the dg-dh transcripts were mapped by 5'-RACE (Ingela Djupedal data/ Ekwall lab). Carolina Bonilla (Ekwall lab) cloned the putative top and bottom transcript promoters in the dg-dh region into a
Figure 4-7A rpb7-G150D cells are proficient in general transcription termination.

*rpbl-1* cells were tested for general transcription termination defects on engineering *ura4* construct as indicated in the diagram at the top (Aranda & Proudfoot, 2001). Growth on plates lacking uracil or with FOA was tested on wild type (*Terwt*), *TerM5* and six different *Ter*rpb7-G150D clones (1-6). *TerM5* cells have a mutation in the terminator allowing read through of the terminator signal and growth on plates lacking uracil; but no growth on FOA plates. Like *Ter* cells, *Ter*rpb7-G150D cells are not defective in transcription termination and all the six *Ter*rpb7-G150D clones do not grow on plates lacking uracil but do grow on FOA plates.
Two primers were used separately in the first cDNA synthesis: either \textit{uraS} or \textit{uraL}. In a wild type \textit{ura4*}/\textit{uraDS/E} strain, in a \textit{Ter*wt}/\textit{ura4-DS/E} strain and in a \textit{TerM5}/\textit{ura4-DS/E} strain, \textit{uraS} primes a band of 300 bp. In a wild type \textit{ura4*}/\textit{ura4-DS/E} strain, \textit{uraL} primes both the \textit{ura4*} gene (660 bp) and the \textit{ura4-DS/E} (400 bp); in a \textit{Ter*wt} strain, \textit{uraL} primes just the \textit{ura4-DS/E} minigene (400 bp) because transcription of \textit{ura4*} stops at the \textit{Ter*} terminator; in a \textit{TerM5}/\textit{ura4-DS/E} strain, \textit{uraL} primes both the \textit{ura4-DS/E} (400 bp) and the \textit{ura4*} gene (660 bp) because transcription can read through the \textit{TerM5} terminator. The light blue box in the \textit{ura4*} gene, represents the region that is absent in the \textit{ura4DS/E} minigene.
RT-PCR analysis was performed on Ter*wt, TerM5 and two Ter*bp7-G150D clones (2 and 6). First cDNA synthesis step was primed with either uraS (located 5' to the Ter*M5 signal) or uraL (located 3' to the Ter*M5 signal) primers as indicated in the diagram. Subsequent PCR was done using ura5' and either uraS or uraL primers. TerM5 cells have a mutation in the terminator signal and they read through it. rpb7-G150D cells do not read through. All the strains contain the *ura4DS/E* minigene allele as an internal control.

**Figure 4-7C rpb7-G150D cells are proficient in general transcription termination.**
LacZ reporter vector (Forsburg, 1993). The plasmids were introduced into wild type, rpb7-G150D, rpb1-11 and dcr1Δ cells and β-galactosidase assays were performed: the top promoter activity is equal in rpb7-G150D and wild type cells whilst the bottom promoter activity is four times reduced in rpb7-G150D cells compared to wild type cells at 25°C. Importantly, stability of the LacZ transcript was not affected when nmt1-promoter-LacZ construct was introduced in rpb7-G150D cells as the amount of LacZ transcript was comparable to that in wild type cells (Figure 4-8). The effect is specific for rpb7-G150D cells since in rpb7-G150D cells the transcription is comparable to the one in wild type cells, considering the standard deviations. This suggests that initiation of transcription from the bottom promoter is specifically defective in rpb7-G150D cells at 25°C.

4.7 Localization of RITS complex.
The RNA-induced initiation of transcriptional gene silencing (RITS) complex contains Dicer-generated siRNAs and is required for heterochromatic silencing. RITS components, Ago1, Tas3 and Chp1, associate to the dg-dh repeats by ChIP (Verdel et al., 2004; Noma et al., 2004). The specific defect in the production of the dg-dh transcripts in rpb7-G150D cells may interfere with the recruitment of the RITS complex to the centromeric repeats and subsequently this may affect silent chromatin formation. To test this hypothesis, ChIP analysis was performed in rbp7-G150D cells expressing Chp1TAP tagged in wild type cells and untagged cells (Figure 4-9). The DNA immunoprecipitated with IgG-sepharose beads, was amplified with pairs of primers specific for the dg-dh repeats (cenfor/cenrev) and fbp1 genes as a control. Chp1 localizes to the dg-dh repeats in rbp7-G150D cells as in wild type cells. The RITS complex is normally recruited at the dg-dh repeats in rpb7-1 cells. This suggests that despite "few" dg-dh siRNAs are produced in rpb7-1 cells, they are sufficient to ensure the RITS complex localization. However rpb7-G150D cells have reduced, but not abolished H3me2K9 and partially defective silencing. It could be that this residual H3me2K9 is sufficient to allow RNAi independent binding of RITS.

DISCUSSION
Rpb7 functions in making centromere transcripts.
The data presented in this chapter suggest that RNA pol II is responsible for the generation of fission yeast centromere repeat transcripts that are processed by RNAi into homologous siRNAs. In particular a subunit of RNA pol II, Rpb7, is required for initiation of transcription of centromeric repeats and its mutation leads to loss of these transcripts and siRNAs. rpb7-G150D mutation was identified in the same screen as mutants that affect silencing at the outer repeats and are involved in RNAi (csp9-csp13, Volpe et al., 2003). In particular Sharon White in the lab, sequenced csp7 and csp9 and found that they are respectively an allele of rdrp1 and ago1. In contrast to RNAi components which are not essential, rpb7-G150D is a
Figure 4-8 Rpb7 is specifically required to promote initiation of centromere repeat transcription.

A) The transcription start site was revealed with 5' RACE analysis of centromeric transcripts originating in the outer centromeric repeats of a dcr1Δ strain. Centromeric dg-dh promoter fragments were cloned upstream of the LacZ reporter gene. RevL RevM RevS are different truncated versions of the reverse (Rev) promoter starting at ^ and ending at the transcription start site (green). The position of the putative TATA box is indicated in red.

B) Analysis of dg-dh promoter gene expression using forward (For) and reverse (Rev) lacZ fusions in wild type, rpb7-G150D and rpb1-11 cells. β-galactosidase activity (Miller Units) was assayed in cells grown at 25°C. The nmt1 promoter (pREP41X-lacZ) was used as positive control. The mean activity of five independent transformed colonies is shown as % of wild type. The error bars represent the standard deviation.
Figure 4-9 The RITS complex is localized at the dg-dh repeats in rpb7-G150D cells.
Chromatin Immunoprecipitation was performed on wild type Chp1TAP tagged, rpb7-G150DChp1TAP tagged and untagged cells. The DNA immunoprecipitated was then amplified using pair of primers for dg-dh repeats (cen for/cen rev) and fbp1 as a control. Fold enrichment values are indicated underneath each lane. They were calculated as a ratio between the value obtained for the IP over the one obtained for the Total DNA. The dg-dh values were normalized to the fbp1 value both from the Total and the IP DNA. The quantification was performed using the EASTMAN Kodak co. EDAS 290 System and 1D Image Analysis software.
temperature sensitive allele of an essential gene that encodes for a subunit of RNA pol II and its function is upstream of the RNAi pathway.

Several observations indicate that RNA pol II is required for the production of centromeric transcripts:

1) ChIP analyses indicates that the large subunit is associated with outer repeats,
2) centromere transcripts are polyadenylated,
3) centromere transcripts do not accumulate to the same extent in rpb7-G150Ddcr1Δ compared to dcr1Δ cells,
4) northern analyses reveals that the levels of dg-dh siRNAs are dramatically reduced in rpb7-G150D cells,
5) expression of a LacZ reporter gene from the putative centromeric reverse promoter is dramatically reduced in rpb7-G150D relative to wild type cells.

Together these data indicate that RNA pol II is required to produce the centromeric transcripts that are substrate for Dicer and the RNAi pathway to form centromeric siRNAs that then direct silent chromatin formation. In fact the Rpb7- G150D mutation results in less transcription, thus less dsRNA substrates and then lower levels of siRNA and defective silencing.

To further validate this, attempts were made to detect Rpb7 on centromeric repeats with α-ScRpb7 (Santa Cruz Biotechnology) without success. Unfortunately we were unable to obtain published reagents to detect S. pombe Rpb7 and its partner Rpb4 (Ishimama/University of Tokyo).

The C-terminal domain (CTD) of the large subunit of RNA pol II (Rpb1), contains multiple conserved YSPTSPS heptad repeats, the phosphorylation state of which regulates the binding of various mRNA processing factors, thus coupling mRNA processing to transcription (Orphanides et al., 2002; Reed et al., 2003 Proudfoot, 2004). Deletion of 17 heptad repeats (rpb1-11) results in defective centromeric silencing but rpb1-11 does not affect the production of centromere transcripts as Rpb7 (Schramke et al., 2005). The phenotype of rpb1-11 is distinct from rpb7-G150D described here in that it is not temperature sensitive, does not accumulate centromere transcripts by RT-PCR, and does produce dg-dh siRNAs. However, despite centromeric transcripts and siRNA being intact, rpb1-11 presents a reduction in H3K9 methylation and Swi6 binding over a marker gene (ura4) inserted at the dg-dh repeats. This suggests that the CTD of Rpb1 acts downstream to allow the conversion of an RNAi signal into chromatin modification. The transcription of a LacZ gene driven by the putative centromere promoter is not affected in rpb1-11 as it is in rpb7-G150D. This further suggests that specifically and only Rpb7 subunit of RNA pol II acts upstream the RNAi pathway, in particular in the production of pre-siRNAs.

The fact that Rpb7, Rpb1 and Rpb2 subunits (Kato et al., 2005) of RNA pol II, affect silencing at the centromere but have different phenotypes relating to siRNAs production, H3K9 methylation and Swi6 localization, suggests that the RNA pol II holoenzyme complex acts upstream to produce centromeric transcripts but is also required downstream to allow targetting of RNAi and
thus chromatin modifying enzymes to homologous chromatin. Interestingly Ago1 has also been found to Co-IP with Rpb1/ pol II (Schramke et al., 2005). Together the analyses of rpb7- G150D, rpb1-11 and rpb2-m203 confirm that RNA pol II is required in silent chromatin formation.

**rpb7-1 cells are defective in spreading silent chromatin.**
Like the RNAi mutants and chp1Δ cells, rpb7-G150D cells exhibit only partial loss of H3K9me2 and Swi6 over the centromeric dg-dh sequence but complete loss from a non centromeric DNA ura4 marker inserted within centromeric repeat DNA. This suggests that some centromeric siRNAs are still produced in rpb7-G150D cells allowing the nucleation of silent chromatin on outer repeats but that this silencing is less robust and unable to spread into neighbouring marker genes.

**Rpb7 has a specific defect in the production of the bottom strand transcripts.**
Analyses of forward and reverse centromeric promoter constructs and transcriptional run on experiments suggest that rpb7-G150D mainly affect the production of the reverse centromeric transcript.

The order of the events that lead to the establishment of silent chromatin at the centromere is not known but the following model seems reasonable (Figure 4-10). Forward and reverse transcripts provide a substrate which is processed into a first generation of siRNAs (primary siRNAs). These primary siRNAs are incorporated into Ago1 (RITS) which targets nascent transcripts that remain associated with elongating RNA pol II. Rdpl (RDRC) then starts to produce more dsRNA and secondary siRNAs, creating a self-enforcing loop (Motamedi et al., 2004; Sugiyama et al., 2005). Such a process is expected to be less efficient in rpb7-1 cells where already the primary dg-dh siRNAs are reduced. In this context it could be interesting to assess the phenotypes of double mutants such as rbp7-G150DrdplΔ or chp1Δ rbp7-G150D cells. It may be that the increased ssRNA-binding property of rbp7-G150D inhibits initiation of transcription by stalling the transition from Pol II initiation to elongation complex (Djupedal et al., 2005).

Rpb7 is highly conserved between yeast and human and cen transcripts are also produced in vertebrate (Fukagawa et al., 2004). It seems likely that RNA pol II subunits such as Rpb7 play a conserved role in heterochromatin formation in animal cells. In fact while in plants RNA pol IV is responsible for the transcription of 5S gene clusters and ATSn1 retroelements and it is involved with the RNAi pathway in maintaing silent these repeats through DNA methylation-dependent heterochromatin formation (Onodera et al., 2005; Herr et al. 2005), in yeast and mammalian cells, this specific RNA polymerase does not exist. This again support the idea that subunits of RNA pol II evolved a very fined system and regulation for the paradoxical transcription of silent genomic loci.
Figure 4-10 Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing.
Overlapping transcripts are produced from dg-dh centromeric repeats. *rpb7-G150D* cells have a specific defect in the production of the bottom strand. Less siRNAs are produced but they are still enough to direct the recruitment of the RITS complex over the centromeric repeats. Somehow the H3K9 methyltransferase Clr4, is recruited but its activity is reduced. This causes less H3K9me2, less Swi6 binding and inefficient silent chromatin formation.
CHAPTER 5

SPLICING MUTANTS AFFECT SILENT CHROMATIN FORMATION AT FISSION YEAST CENTROMERE

INTRODUCTION

In fission yeast, two complexes involved in RNAi, RITS and RDRC, physically interact with each other and associate with noncoding centromeric RNAs (Verdel et al., 2004; Motamedi et al., 2004). The RNAi effector complex, RITS which contains Ago1, the chromo-domain protein Chp1, and Tas3, is required to mediate RNAi-mediated heterochromatin assembly (Verdel et al., 2004). RITS also contains siRNAs that match the sequence of repetitive DNA at the outer centromeric repeats where heterochromatin assembly occurs (Reinhart & Bartel, 2002; Verdel et al., 2004). In dcr1Δ cells, RITS is not loaded with siRNAs and does not associate with heterochromatic DNA regions. This, together with recent analyses, supports the idea that RITS uses Dicer-produced siRNAs to target specific chromosome regions by a recognition mechanism that involves RNA/RNA base-pairing interaction.

In addition to siRNAs, the association of RITS with chromatin requires Clr4, the H3K9 methyltransferase. H3K9me2 stabilizes the association of RITS with chromatin by creating a binding site for the chromo-domain of Chp1 (Partridge et al., 2002). RNAi-mediated silent chromatin formation also requires an RNA-directed RNA polymerase (Rdp1). Rdp1 is part of the RDRC complex (RNA-directed RNA polymerase complex), which also contains two other highly conserved proteins: Hrr1, a putative RNA helicase, and Cid12, a protein that belongs to the polyA polymerase/2'-5' oligoadenylate synthetase family of enzymes. This complex has RNA template-dependent RNA polymerase activity (Motamedi et al., 2004) and both its activity and associated subunits are required for association of siRNAs with RITS and for centromeric gene silencing. RDRC interacts physically with RITS and this interaction requires Dcr1 and Clr4 (Motamedi et al., 2004). In particular, Cid12 together with Cid13, are members of a highly conserved family of proteins that includes the budding yeast Trf4 and Trf5. These proteins in budding yeast have been implicated in non-coding RNA polyadenylation: they are part of a poly(A) polymerase complex involved in RNA quality control (Read et al., 2002, Saltoh et al., 2002; Vanacova et al., 2005; La Cava et al., 2005). Affinity purification of fission yeast Cid12-TAP tagged and associated proteins from a rdp1Δ strain identified many subunits of the spliceosome (29) (Motamedi et al., 2004), although the significance of this is unclear.
In a genetic screen, performed by Karl Ekwall, 12 mutations falling into 11 complementation groups, required for the integrity of silent chromatin over the outer repeats at fission yeast centromeres were identified (Ekwall et al., 1999). Of them csp9, csp10, csp11, csp12 and csp13 have been shown to be involved in RNAi directed heterochromatin formation over the repeats and csp9 was found to be allelic to Argonaute1 (ago1), a component of RITS (Volpe et al., 2003; Verdel et al., 2004). The csp1 to 6 mutants are temperature sensitive (ts) and display defective silencing of centromeric marker genes at 25°C but are inviable at 36°C. csp4' gene was identified by complementation with a genomic library by Karl Ekwall. csp4-29 cells are temperature sensitive lethal in a gene designated cwf10* (SPBC215.12: csp4-29 is now called cwf10-1) (Karl Ekwall). The Cwf10 protein associates with a large pre-RNA splicing complex containing at least 26 proteins. Cwf10 is homologous to human ribosomal translocase and the ortholog of S. cerevisiae Snu114p (the only GTPase involved in splicing).

The spliceosome is a collection of snRNAs and proteins recruited to nascent transcripts (pre-mRNAs) to carry out intron excision. Within the assembled spliceosome, intron excision occurs in two chemical steps (Moore et al., 1993): first, 5' splice site cleavage and lariat formation, then 3' splice site cleavage and exon ligation. Following exon ligation, complete spliceosome disassembly would then free up its components for de novo assembly of other spliceosomes. During the splicing cycle, the RNA and protein components of the spliceosome undergo numerous rearrangements, which must be highly coordinated in order to ensure fidelity of the process (Staley & Guthrie, 1998). Most of these rearrangements appear to be energy dependent, and are correlated with the activity of individual ATPases of the DExD/H-box family. Eight known DExD/H-box proteins are required for the splicing cycle, and mutations in these proteins inhibit the ATP-dependent steps of splicing (Stanley & Guthrie, 1998). Additionally, splicing requires one GTP-ase, Snu114p, which is an essential protein in S.cerevisiae (Fabrizio et al., 1997). Snu114p is homologous to the ribosomal translocase elongation factor G (EF-G in prokaryotes/EF-2 in eukaryotes) leading to the hypothesis that Snu114p may similarly use the energy of GTP hydrolysis to drive rearrangements of the spliceosome (Fabrizio et al., 1997).

Snu114p is packaged with other proteins and the U5 snRNA to form U5 snRNP (small ribonucleoprotein particle). As represented in Figure 5-1 (top panel), prior to formation of the spliceosome, U5 snRNP interacts with the U4/U6 di-snRNP, in which U4 and U6 snRNAs are extensively base paired, thus forming U4/U6.U5 tri-snRNP (Burge et al., 1998). This tri-snRNP is then recruited to the pre-spliceosome formed by U1 snRNA and U2 snRNA and the sequence to splice. Rearrangements that occur during the early stages of spliceosome activation are regulated by several components of the U5 snRNP (Brow 2002). The Prp28 ATPase is required to unwind the U1/5’ splice site duplex, possibly by destabilizing protein components of U1 snRNP (Chen et al., 2001; Staley & Guthrie, 1998) and the Brr2 ATPase is required to unwind the U4/U6 duplex (Kim & Rossi, 1999; Laggerbauer et al., 1998;
Figure 5-1 ScSnu114p interactions in the pre-spliceosome.
Top panel shows the splicing cycle in *S. cerevisiae* as described in the text. Bottom panel shows the model for Snu114p activity during spliceosome activation by Brenner & Guthrie, 2005. Snu114p is bound to GTP when tri-snRNP first interacts with the spliceosome. Proper interaction between tri-snRNP and the spliceosome involves GTP hydrolysis causing a conformational rearrangement of Snu114p. This alters the interaction between the IV domain of Snu114p and Prp8p, changing the conformation of Prp8p. Subsequently Prp28p and Brr2p ATPase activity is triggered and U1 and U4 released.
Prp8, a large well-conserved U5 snRNP protein (220K in mammalian cells, Cwf6 in S. pombe) regulates the ATPase activity of Prp28 and Brr2 until spliceosome formation has occurred (Collins & Guthrie, 2000; Kuhn & Brow, 2000; Kuhn et al., 1999; Graiger & Beggs, 2005). In addition to the putative U4/U6 helicase Brr2p, five splicing factors in S. cerevisiae have already been implicated in the release of U4 snRNA during spliceosome activation. These are: the U4/U6 snRNP protein Prp4p (Ayadi et al., 1997), the non-snRNP protein Prp19p (Tarn et al., 1993), the tri-snRNP protein Prp38p and the U5 snRNP proteins Prp8p and Snu114p (Xie et al., 1998; Kuhn et al., 1999; Fabrizio et al., 1997; Bartels et al., 2002). A strong physical interaction between Prp8p and Snu114p may also play a regulatory role during spliceosome activation: the human orthologs, 15-220kD/Prp8p and U5-16kD/Snu114p, remain associated as a heterodimer after treatment of U5 snRNP with high concentration of chaotropic salts that disrupts the complex (Achsel et al., 1998). By two hybrid analysis, also the U1 proteins Prp39p and Prp40p, interact with the neighboring regions of Prp8p (Abovich & Rosbash, 1997; Dix et al., 1998; Grainger & Beggs, 2005; Van Nues & Beggs, 2001). In a recent paper by Brenner & Guthrie, 2005 as shown in the bottom panel of Figure 5-1, it has been proposed that GTP hydrolysis by Snu114p could be triggered by interactions with U1 snRNP.

In this chapter I will focus on csp4/cwf10, the ortholog of Snu114p. I will also describe the identification of the gene mutated in csp5, which is another splicing factor, prp39, ortholog of S. cerevisiae prp39 and I will discuss a possible explanation for their involvement in silent chromatin formation.

RESULTS

5.1 Sequencing of csp4/cwf10 and structural analysis.
Csp4/Cwf10 is the ortholog of Snu114p. It is known that Snu114p is a GTPase with homology to the ribosomal translocase EF-G/EF2 (Fabrizio et al., 1997). Over a stretch of 904 amino acids Snu114p is 27% identical and 47% similar to EF-2. Cwf10 (983 aa) is 35% identical and 53% similar over a stretch of 1050 amino acids to Snu114p (1008 aa). In order to verify whether Cwf10 has the same structural domains as EF-G/EF-2, Clustal W alignment was performed (Thompson et al., 1994). As shown in figure 5-2 the N-terminal domain of Cwf10 has a stretch of hydrophobic amino acids that is evolutionary conserved with Snu114p. Aside from the N-terminal acidic domain (120 amino acids) that is not present in EF-2, Cwf10 and Snu114p have the same domain structure as EF-2, including the G-domain. S. cerevisiae EF-2 structure has been resolved by X-ray crystallography. It has five structural domains (Jorgensen et al., 2003) and by the alignment as shown in Figure 5-2, Snu114p and Cwf10 contain most of the conserved amino acids in the G-domain which is
subdivided into the p-Loop and the Guanine-ring binding sequence, indicated by the blue and red boxes; as well as in Domain II, Domain III, Domain IV and Domain V.

I sequenced cwf10 gene in csp4-29 cells and wild type cells and found that the mutation in csp4-29 is after the hydrophobic N-terminal domain, precisely at the 323th amino acid. It causes a missense mutation that converts a cysteine to a tyrosine (C323Y), as shown in Figure 5-3. All the members of the GTPase superfamily share a similar architecture of the G domain and contain conserved motifs (G1-G5) that interact with GTP/GDP (Bourne et al., 1991; Sprang, 1997).

Despite the fact that this amino acid is not conserved between Cwf10, Snu114p and EF-2, it is located in the stretch of amino acids that are part of the conserved G-domain of EF-2 and it may be important for the overall structure and function of Cwf10 specific protein. Interestingly the fifth amino acid after this cysteine is another cysteine, that is also not conserved in Snu114p or EF-2: together these two cysteines may form a disulphur bridge.

The importance of G-domain has been studied in Snu114p where it seems to be involved in binding and hydrolysis of GTP. Several amino acid substitutions in the P loop of the GTP-binding domain of Snu114p are lethal, suggesting that GTP binding and hydrolysis are important for its function (Bartels et al., 2002). This step is in fact required during the unwinding of U4 from U6 in the pre-mRNA splicing process. Other mutations in the G-domain of Snu114p in S. cerevisiae produce a temperature sensitive phenotype and splicing defect at restrictive temperature (Bartels et al., 2003). A similar function may be conserved in Csp4/Cwf10 since csp4-39/cwf10-1 mutation lies in the G-domain.

5.2 Cloning of csp5 and identification of csp5-39 mutation.

The csp5 mutant was identified in the same way as csp4/cwf10-1. Mutants which display strong defects in silencing over the centromeric repeats of centromere 1, such as swi6, clr4 and rik1, are sensitive to microtubule destabilizing drugs and exhibit elevated chromosome loss rates (Allshire et al, 1995; Ekwall et al., 1996). csp5-39 cells grow slowly even at permissive temperature, and have the highest rate of minichromosome loss among the temperature sensitive class mutants (140 fold higher than wild type cells). csp5-39 cells are also supersenssitive to Thiabendazole (TBZ), a drug that destabilizes microtubules. However the silencing defect in csp5-39 cells is intermediate and comparable to that of csp4-29 cells and is clearly less than that seen in mutants lacking silencing such as clr4Δ (Ekwall et al., 1999).

In order to try to understand the role of this mutation in silent chromatin formation, the Shimoda genomic library (Tanaka et al., 2000; Nakamura et al., 2001) was used to identify by complementation the genomic clones. All the transformants were plated on minimal media lacking leucine and with phloxin (0.02%v/v). After five days of growth at 25°C, the plates were moved at 36°C for three days. Phloxin is a vital dye taken up by dead cells (they have a darker pink colour). No phloxin pale colonies were detected on the plates where csp5-39 cells
Figure 5-2 Alignment of ScE6-2, SpCwf10 and ScSnu114p.
Alignment using Clustal w program (Thompson et al., 1994). The conserved GKS residues in the P-loop are indicated by the blue box, the conserved NKXDR motif in the guanine-ring binding is indicated by the red box. These residues are all conserved in fission yeast Cwf10 and budding yeast Snu114p. In orange is the residue mutated in Cwf10. The alignment was shaded using BOXSHADE (version 3.2; Hofmann & Baron).
Figure 5-3 *csp4-29* is an allele of *cwfl0*, encoding for ScSnu114p homologue.

Sequencing analysis was performed on *cwfl0* gene from wild type and *csp4-57* cells in both orientations. *csp4-29* has a missense mutation in the *cwfl0* gene. This mutation converts the triplet coding for cysteine, C (323) in tyrosine Y. In the picture two chromatograms are represented from the sequence analysis using Sequencher 4.1 program.
were transformed with the empty plasmid whilst 13 pink colonies out of a total of 50000 transformants plated, were recovered on plates lacking leucine and supplemented with phloxine at 36°C. Plasmids were successful rescued into E.coli from 10 of the 13 pink ts+ colonies, and miniprep DNA analyzed by restriction pattern analysis using Sall and NotI, two restriction enzymes that cut once in the multiple cloning site of the pAL-KS. As shown in Figure 5-5A from the digestion pattern, the 10 clones fell into 5 groups (A, B, C, 2 and 5). These plasmids were transformed back in csp5-39 cells and silencing at the outer repeats (otr1R:ade6") was monitored.

Serial dilution assays demonstrate that all 10 plasmids, not only rescued the temperature sensitivity of the csp5-39 mutant but also the silencing defect (Figure 5-5B). csp5-39 cells bearing the ade6" marker gene at the outer repeat of centromere 1 (otr1R:ade6") transformed with plasmid A or B or C or 5 exhibit the same red colour as wild type cells transformed with pAL-KS empty plasmid. On the contrary, csp5-39 cells transformed with the empty pAL-KS plasmid remained temperature sensitive and were silencing defective (white).

Partial DNA sequencing of the insert in these five clones using universal 5’ and 3’ primers annealing with the multiple cloning site from the plasmid was performed. In all the five plasmids recovered as shown in Table 1 and the diagram below (Figure 5-6), part of the cosmid SPBC4B4 (19475-25609) was common, and the three genes present in all the rescued plasmid were SPBC4B4.08c (19647-21242), SPBC4B4.09c (21876-24049) and SPBC4B4.10c (24277-25265). These sequences are on chromosome II and they encode respectively for a hexose transporter, Ght2; a splicing factor, Prp39; and an apoptosis specific protein homolog, Apg5. We reasoned that the mutation, being a temperature sensitive one, should be in an essential gene. The published literature indicates that ght2 and apg5 are not essential genes whereas prp39 is required for cell viability.

Sequencing was performed on the prp39 ORF amplified by PCR from wild type and the csp5-39 mutant. A missense mutation was found which converts Trp 550 to a stop codon (TRP550STOP)(Figure 5-7). As shown in the alignment in Figure 5-8, Prp39 is conserved in human (it shares over a stretch of 410 aa, 32% identity and 52% similarity with S. pombe Prp39) and S. cerevisiae (it shares over a stretch of 528 aa, 23% identity and 43% similarity with S. pombe Prp39).

Prp39 contains also repetitive motifs, the HAT helix, that is implicated in RNA processing (Preker & Keller, 1998). The repetitive pattern is characterized by 3 aromatic amino acid with conserved spacing. The first residue, usually trp (W), is the most conserved. These repeats are present in human cleavage stimulation factor Ctf-77, in Drosophila melanogaster homologue su(f), suppressor of forked and in S. cerevisiae homologue Rna14p as shown in Figure 5-9. Its function is usually linked with RNA processing.

As shown in the diagram of Figure 5-10, plasmid C was digested with the restriction enzyme BgIII to be linearized in the prp39 gene. Subsequently this linearized plasmid was transformed in csp5-39 cells to facilitate its genomic integration at the endogenous csp5-
Figure 5-4 Strategy to clone *csp5*.

- **csp5**-39° cells
- Shimoda high copy genomic library
- TRANSFORMATION BY ELECTROPORATION
- PMG-LEU + phloxine plates 25°C for 5 days
- MOVE to 37°C for 3 days
- Recovered 13 clones
- Recovered plasmid after *E.coli* transformation: 10/13.

Group these 10 clones by digestion pattern in 5 groups
Sequence them.
Figure 5-5 csp5-39 temperature and silencing defects are rescued by five plasmids.

(A) Plasmids that rescued the temperature sensitivity of csp5-39 cells and empty pAL-KS plasmid (last lane) were digested and grouped according to their restriction pattern in five categories (1,6,8 = group A; 3,4 = group B, 7,9,10 = group C; 2,5). (B) These plasmids were retransformed in csp5-39 cells (otr1R:ade6") that then were tested for silencing defect by a spotting assay on plates with low adenine concentration. As control wild type cells and csp5-39 cells transformed with the empty plasmid pAL-KS were plated too.
Figure 5-6 All the five plasmids contain prp39+ and ght2+ genes. Sequencing of all the five inserts in the pAL-KS plasmids in both directions, 5' and 3' was performed and by BLAST search the common region was mapped to SPBC4B4 cosmid, between the 19475 and 25809 nucleotides. This region comprises three genes encoding for Ght2 (19647-21242), Prp39 (21826-24049) and Apg5 (24277-25265).
Figure 5-7 *csp5-39* is an allele of *prp39*.
Sequencing analysis was performed on *prp39* gene from wild type and *csp5-39* cells in both orientations. *csp5-39* has a missense mutation that converts the triplet 550 encoding for a tryptophan (W) into a stop codon. In the picture two chromatograms from wild type *prp39* cells and *csp5-39/prp39-1* cells are represented from the sequence analysis using Sequencher 4.1 program.
Figure 5-8 Alignment of Prp39 from *S. pombe*, *human*, and *S. cerevisiae*.

Clustal W alignment was performed on Prp39. The mutation in csp539/prp39-1 lies on the 550th tryptophan that is circled in red. SpPrp39 has 32% identity and 52% similarity with human Prp39 over a stretch of 410 amino acids. SpPrp39 has 23% identity and 43% similarity with ScPrp39 over a stretch of 528 amino acids.
Figure 5-9 Prp39 contains a HAT motif in the N-terminus.
The HAT helix is a repetitive motif implicated in RNA processing (Preker & Keller, 1998). It is characterized by 3 aromatic amino acids with conserved spacing. The first residue, usually W, is the most conserved. In green are marked 4 of these motives. Alignment based on Clustal W, was done with human cleavage stimulation factor Ctf-77, Drosophila melanogaster homologue su(f), suppressor of forked, S. cerevisiae homologue Rna14p. Only the first N-terminus with this motif is reported.
39/prp39-1 allele. This strain has ade6+ at the outer repeats and was monitored for temperature sensitivity and silencing: most leu+ transformants obtained were red (silencing restored) and grew at 36°C (ts+) although a few white ts transformants were obtained that probably derived from rearrangement events destroying the wild type plasmid bearing prp39 gene.

To verify that csp5-39 is complementated by prp39+, the red temperature resistant integrant (1) and the white temperature sensitive strains (2) were crossed to csp5-39 cells and to isogenic wild type cells. From all the leu+ colonies, 50% of the progeny coming from 1 crossed to csp5-39 cells, were both temperature sensitive and silencing defective. The remaining 50% were wild type. In addition none of the progeny resulting from the cross between 1 and wild type cells, was temperature sensitive and silencing defective. Both sequencing and genetics confirmed that csp5-39 is indeed an allele of prp39.

5.3 Other splicing mutants affect centromere silencing.

As described in Chapter 1, csp4-29/cwf10-1 and csp5/prp39 cells were identified in a genetic screen for factors required for the integrity of silent chromatin over the outer repeats (Ekwall et al., 1999). Other splicing mutants (prp2, prp3, prp5, prp8 and prp12) have been found to alleviate silencing at centromeres to some extent (Elizabeth Bayne and Vera Schramke unpublished results). Like cwf10-1 and prp39, they show intermediate defects in silencing and some also exhibit variegation. When assaying silencing of the ade6 gene inserted in centromeric repeats, prp2 and prp3 have a weak defect, while prp5, prp8 and prp12 have stronger phenotypes.

ChIP analyses on these strains with antibodies specific for H3K9me2, Swi6, the heterochromatin markers, and H3K4me2 were performed. Over the dg-dh repeats, like wild type cells, prp2, prp3, prp5, prp8, prp12, cwf10 and prp39 all retain H3K9me2 and Swi6 association (Figure 5-11A); whilst over the ura4 marker gene inserted in these repeats, prp8 prp12 and prp39 cells, exhibit a drop of Swi6 binding and an increase in H3K4me2, a mark for actively transcribed genes. This drop is not as dramatic as in swi6A and clr4A cells respectively (Figure 5-11B). These differences between heterochromatin formation over the outer repeats themselves and a maker gene inserted between these repeats may be due to a problem in heterochromatin spreading.

Heterochromatin spreading is defective in swi6A cells but it can derive also from inefficiency of the RNAi pathway (Sadaie et al., 2004). In order to investigate whether the formation of silent chromatin formation via-RNAi pathway was functional, northern analysis was performed together with Liz Bayne, on centromeric siRNAs and centromeric transcripts from these mutants and the results are summarized in Table 1 (Figure 5-12): prp2 and prp3 cells produce some siRNAs but not at the same level as wild type cells and do not accumulate centromeric transcripts; cwf10, prp8 cells produce very low amounts of siRNAs and do not accumulate centromeric transcripts; while prp5, prp12, prp39 not only produce less siRNAs
Figure 5-10 Strategy to integrate the plasmid linearized at prp39 locus. Linearized plasmid on prp39+ gene was transformed to be integrated in the prp39-1 genome locus. If recombination occurs, a wild type copy of the gene is formed and the temperature sensitivity (ts) and silencing defect (ade6*) are rescued.
but also accumulate some centromeric transcripts. This prompted us to think that different splicing factors have different functions: some might act upstream of Dcr1 in the production of pre-siRNAs (cwf10, prp8) and others might be involved also in the processing (prp5 and prp12, prp39). This is reminiscent of the fact that different subunits of RNA pol II also act at different stages in silent chromatin formation (Rpb7, Rpb2 and Rpb1). Splicing factors, like all the other mRNA processing factors interact with the CTD domain of Rpb1 during transcription. The rpbl-11 mutant has some phenotypic resemblance to some of the splicing mutants in that it variegates, it produces some siRNAs but less than wild type cells, it does not accumulate centromeric transcripts and it has decreased H3K9me2 and Swi6 binding on a marker gene inserted at the outer repeats (Schramke et al., 2005).

Silencing defects could be due to failure to splice RNA from genes encoding components of the RNAi machinery. Therefore, in order to determine whether there is a correlation between splicing activity and silencing defect, RT-PCR analysis was performed at permissive temperature and restrictive temperatures to check the accumulation of unspliced mRNA species of several genes: cdc2, nda3, tbp1 in prp2, prp3, prp5, prp12 and cwf10-1 mutants. Specific primers complementary to the sequences of each of these genes located at an intron junction were used in a RT-PCR. As a control, genomic DNA was used to determine the size of the unspliced form.

As shown in Figure 5-13, apart from tdp1 which has a very sensitive intron, and the unspliced form accumulates even in wild type cells, none of the mutants exhibits splicing defect at permissive temperature. At restrictive temperature, there is accumulation of the unspliced form for all the genes examined in prp2, prp3, prp5, prp12 and cwf10-1 cells but not in wild type cells. ago1 gene has an intron: RT-PCR was performed by Liz Bayne on this gene and it seems to have a quite sensitive intron even in wild type cells, at permissive temperature. Transformation with a high copy number plasmid bearing the ago1 gene has to be performed to assess whether the silencing defect in these splicing mutants can be rescued and therefore it is just an indirect phenotype.

**DISCUSSION**

This chapter plus other data from Chapter 3 and from work done in our laboratory, implicate a link between splicing and centromeric silencing.

Mutation in Cwf10 and Prp39, two characterized splicing factors, known to be involved in pre-spliceosome formation (McDonald et al., 1999), were isolated in a screen for mutants that alleviate silencing over the outer repeats of centromeres (Ekwall et al., 1999) and they have been further characterized in this chapter.

Cwf10 is an essential protein whose splicing defect becomes evident only at restrictive temperature. At permissive temperature when cwf10-1 cells exhibit the silencing defect, none of the silent chromatin features are severely compromised. Only dg-dh siRNA
Chromatin immunoprecipitation was performed at the permissive temperature (25°C) with antibodies specific for H3K9me2 and Swi6. The immunoprecipitated DNA was amplified in a multiplex PCR with pairs of primers for outer repeats (cen for/cen rev) and an euchromatic gene, fbp1. The partial decreased signals in both H3K9me2 and Swi6 ChIP on csp5 cells is not reproducible.
**Swi6 ChIP at 25°C**

<table>
<thead>
<tr>
<th></th>
<th>wt</th>
<th>prp2</th>
<th>prp3</th>
<th>prp8</th>
<th>prp12</th>
<th>cwf10</th>
<th>prp39</th>
<th>swi6Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>iP</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>Ratio iP/T</td>
<td>8</td>
<td>7.2</td>
<td>10</td>
<td>4.1</td>
<td>4.6</td>
<td>11</td>
<td>2.7</td>
<td>0.8</td>
</tr>
</tbody>
</table>

**H3K4me2 ChIP at 25°C**

<table>
<thead>
<tr>
<th></th>
<th>wt</th>
<th>prp2</th>
<th>prp3</th>
<th>prp5</th>
<th>prp8</th>
<th>prp12</th>
<th>cwf10</th>
<th>prp39</th>
<th>clr4Δ</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>iP</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>Ratio iP/T</td>
<td>0.19</td>
<td>0.26</td>
<td>0.32</td>
<td>0.37</td>
<td>0.47</td>
<td>0.50</td>
<td>0.05</td>
<td>0.25</td>
<td>0.5</td>
</tr>
</tbody>
</table>

**Figure 5-11B Splicing mutants affect silent chromatin formation over a marker gene inserted at the outer repeats.**

Chromatin Immunoprecipitation was performed at 25°C with Swi6 and H3K4me2 antibodies. The DNA immunoprecipitated was amplified using specific primers for ura4 gene. These primers amplify the euchromatic uraDS/E minigene, used as control and the ura4* gene inserted in the centromeric dg-dh repeats (otr1(SphI):ura4*) in a competitive PCR reaction. Quantification is indicated below and is the result of the ratio between the value obtained from the IP over the one from the total of each sample.
Figure 5-12 Table 1. Summary of the features relative to RNAi in splicing mutants.
Splicing mutants behave differently in regard to siRNAs production and centromeric transcript accumulation. All produce less siRNAs than wild type cells. *cwf10* and *prp8* cells have the strongest effect on siRNA production and they do not accumulate cen transcripts. *prp3*, *prp5*, *prp12* and *prp39* produce more siRNAs than *cwf10* and *prp8* but accumulate some cen transcripts.
RT-PCR was performed on RNA preparations from prp2, prp3, prp5, prp12 and csp4 cells grown either at 25°C or 36°C. Specific primers across the intron junction of cdc2, nda3 and tbp1 genes as indicated in the scheme at the bottom, were used. Genomic DNA was used as template in a control PCR for the unspliced forms of each gene.

Figure 5-13 prp2, prp3, prp5, prp12 and csp4/cwf10 cells are not defective in splicing at 25°C.
production appears to be drastically reduced. This supports the idea that *cwf10-1* cells might have problems in general siRNAs production, and thus silencing. The same could be true for Prp39. In Chapter 3 Figure 3-12, by northern analysis low levels of *dg-dh* siRNAs were detected in *csp5-39/prp39-1* cells as in *csp4-29/cwf10-1* cells. It is still not clear what is the function of these splicing factors in silent chromatin formation. However the fact that alleles of some particular other splicing factors like *prp5, prp8* and *prp12* affect silencing at centromere at permissive temperature when they do not exhibit an apparent splicing defect, support the idea that somehow splicing factors might be involved in silent chromatin formation independently from their specific splicing activity. How can splicing factors impinge upon RNAi directed silencing?

There are two possible explanations for the involvement of such factors in silent chromatin formation at centromeres in fission yeast. These two concern a direct role for splicing factors, in that they would interact directly at the centromere, with the CTD domain of pol II or with the RDRC complex (Figure 5-14). These two models suggest an interaction between centromere sequences and transcription machinery or RNAi components/chromatin components at the centromere.

**A direct role for splicing factors in silencing.**

A possible direct role for splicing factors in silent chromatin formation over the centromeric repeats is suggested by the fact that RNAi appears to act cotranscriptionally to mediate silent chromatin formation and Cid12 associates with splicing factors in *rdrp1A* cells (Motamed et al., 2004).

In Chapter 4, a subunit of pol II, Rpb7, has been found to act upstream of the RNAi pathway in making the centromeric transcripts. Two other mutations, in Rpb2 and Rpb1 subunits of Pol II, have been found to affect the transmission of the RNAi signal to establish silent chromatin (Goto et al., 2005; Schramke et al., 2005). Moreover Ago1 has been shown to interact with the CTD domain of pol II (Schramke et al., 2005). The implication is that RNAi-directed silent chromatin assembly is linked to ongoing transcription.

The CTD domain is known to be important for the interaction with all the general mRNA processing factors involved in polyadenylation, capping and splicing (Orphanides & Reinberg, 2002; Proudfoot, 2004; Reed, 2003). All these processes are linked and dependent on each other and occur co-transcriptionally in budding yeast.

The interaction between the CTD of RNA pol II and Ago1, a RNAi component, suggests that RNAi may also act co-transcriptionally. Although this has not been shown in *S. pombe*, evidence in *S. cerevisiae* suggest that splicing factors associate with chromatin in an intron dependent manner while transcription proceeds (Kotovic et al., 2003). A possible explanation for splicing factors being directly involved in silent chromatin formation may derive from their interaction with the CTD domain (Figure 5-14).
However no sequence similar to an exon-intron junction has been found at the centromere that could more strongly justify the presence of splicing factors during centromere transcription, has been found.

A direct role of splicing factors in silent chromatin formation may be implicated by the fact that 29 splicing factors have been pulled down by Cid12 in the RDRC complex in rdp1Δ cells. It is known that RDRC interacts with RITS whose components bind centromeric sequences (Motamedi et al., 2004; Verdel et al., 2004). Furthermore pol II localized to the centromeric repeats and coimmunoprecipitates with Ago1, a component of the RITS complex (Schramke et al., 2005). It may be that splicing components interact with Cid12/RDRC at the centromere. ChIP experiments to assess the association of Cid12 to the centromeric repeats in splicing mutants have not been carried out yet but they could be useful to clarify whether Cid12 localization is lost. Another useful experiment to elucidate what is happening in these splicing mutants could be RIP (RNA immunoprecipitation) of Cid12.

Do Cwfl10 and Prp39 contribute in making dg-dh siRNAs?

cwf10/csp4 and prp39/csp5 are defective in silencing marker genes (ade6* or ura4* genes) inserted in the centromeric repeats but they do not completely lose H3K9me2 and Swi6 localization over the centromeric repeats. The most serious defect they present, is that the production of dg-dh siRNAs is much less than in wild type cells. This suggests that they may be involved directly in siRNAs production.

It is not known whether RITS is always associated with centromeric transcripts or if it dissociates after slicing. Presumably, if there is dissociation, this process may involve unwinding of the dsRNAs. As previously mentioned, Cwf10/Csp4 is the ortholog of the U4/U6 unwinding protein Snull4p. It has a specific GTPase domain that is fundamental for this unwinding activity. cwf10-1 is mutated in this domain. There are no data to indicate a direct involvement of Cwf10 in the RNAi pathway, for example in unwinding dsRNAs or ds/siRNAs and no biochemistry has been done. It is also not known so far what it is the activity that promotes the unwinding of dsRNA once they are processed by Dicer and enter the RISC complex. If Cwf10 had such an activity, this would confirm a direct action of RNAi machinery and splicing components (like Cwf10) at the centromere. ChIP and RIP experiments to assess whether splicing components are bound to centromere sequences/transcripts could be carried out to better elucidate what it is happening at the centromere.

Prp39 has a HAT domain that is common in RNA processing factors such as human polyadenylation factor Ctf-77, Drosophila melanogaster su(f) and S. cerevisiae Rna14p. Functional information concerning HAT-repeat-containing proteins is currently limited. Where known, HAT proteins are components of macromolecular complexes that are required for mRNA processing. Short amphipathic α-helical repeats are a recurring motif in HAT domains. HAT-containing proteins may provide the structural framework of multiprotein
complexes such as U1snRNP and Cstf. Alternatively they may play a more active role in the rearrangement of protein-protein or protein-RNA interactions (Preker & Keller, 1998). The presence of such a motif could be important in some step of the RNAi pathway during the degradation of dsRNAs.

**An indirect role for splicing factors in silencing.**

Recently a new mechanism has been discovered which acts to clear aberrant RNAs. It is known that a regulated pathway for nuclear pre-mRNA turnover exists in budding yeast (Bousquet-Antonelli et al., 2000). This pathway comprises two degradation mechanisms: a major 3'->5' pathway requiring Rrp41, Mtr3p, Rrp44p and Rrp6p and most of the entire exosome and a minor 5'->3' pathway involving Rat1p. Recent observations (LaCava et al., 2005; Vanacova et al., 2005), suggest the existence of a new *S. cerevisiae* poly(A)polymerase complex involved in RNA quality control. This complex, TRAMP, consists of Trf4p, the catalytic subunit of a new poly(A)polymerase complex that contains two potential RNA-binding subunits and a RNA helicase, Mtr4p. The polyadenylation activity of the Trf4 complex stimulates the subsequent exosome-mediated degradation of hypomethylated initiator tRNA\(^{Met}\), snoRNAs and non-coding RNAs.

Cid1 and Cid13 are two *S. pombe* poly(A)polymerases, that are members of the Trf4/5 family of proteins first described in budding yeast (Sadoff et al., 1995) discovered to have RNA-dependent poly(A)polymerase activity in vitro (Read et al., 2002; Saitoh et al., 2002). Trf4p is in a complex with the helicase Mtr4p. This is reminiscent of the association of Cid12 with the helicase, Hrr1 in RDRC. It has not been shown whether Cid12 has poly(A)polymerase activity but it has a NTD (nucleotransferase domain) common to Cid family members, poly(A) polymerases and 2'-5' -oligoadenylate synthetases. One possibility is that Cid12 performs an additional role in clearing unspliced RNAs and this may explain its association with spliceosome in *rdrp1Δ* cells where the RDRC complex falls apart. Another possibility is that in splicing mutants Cid12 is sequestered to such unspliced RNAs to trigger their degradation and cannot form an active RDRC.

In fission yeast these RNA quality control mechanisms have not been shown to exist but orthologs to the budding yeast Trf4p and Mtr4p do exist and might have the same activity but act in a different context: it might be that in *S. pombe* some components of these RNA quality control mechanism interact with RNAi components but the same process can not occur in budding yeast which lacks the RNAi pathway.

A connection between aberrant mRNA production and transcriptional gene silencing of the homologous genes mammalian, has been observed.

In mammalian cells transcriptional silencing happens on nonsense codon-containing immunoglobulin minigenes (Buhler et al., 2005). This mechanism called nonsense-mediated transcriptional gene silencing (NMTGS), represents a nonsense surveillance mechanism by which truncation of a gene's ORF induces transcriptional silencing through chromatin
remodelling. Together with the RNAi pathway, the exosome takes part in the clearance of aberrant transcripts at the post-transcriptional level (Orban & Izaurralde, 2005). The idea is that splicing mutants may be involved indirectly in silent chromatin formation: the presence of unspliced mRNA might "deviate" some components in the cells that are usually targeted to clear aberrant mRNA but are also part of the RNAi pathway. It is in fact unknown how the mRNA surveillance and RNAi machinery interact with general RNA decay enzymes and recruit them specifically to their targeted mRNAs. More studies need to be done, combining RNAi mutants and exosome mutants that are the main enzymes involved in the final step of RNA degradation.
Figure 5.14 Role for splicing factors at fission yeast centromeres.
Splicing factors might interact at the centromere, with the CTD domain of pol II (1) (Schramke et al., 2005) or with the RDRC complex (2) (Motamedi et al., 2004) and therefore "stabilize" components for silent chromatin formation.
INTRODUCTION

Centromeres of fission yeast resemble those of metazoa. The kinetochore is surrounded by inverted repeats that are heterochromatic: underacetylated on H3 and H4, methylated on H3K9. This latter modification is bound by chromo-domain proteins such as Chp1 and Swi6 (HP1 homolog) that ensures the recruitment of cohesin, indispensable for correct chromosome segregation (Bernard et al., 2001; Nonaka et al., 2002). Factors that are directly involved in silent chromatin formation, such as the H3K9 methyltransferase,Clr4 (Suv 3-9 homolog), are necessary for silencing of marker genes inserted at the outer repeats. In the absence of Clr4, cells completely lose the markers of heterochromatin: H3K9me2 and Swi6 localization (Ekwall et al., 1996). As a consequence they exhibit chromosome segregation defects. Clr4 is also necessary for the RNAi pathway to be functional since production of centromeric siRNAs is lost (Verdel et al., 2004; Sigova et al., 2004).

Among the csp ts mutants that alleviate silencing at the outer repeats, csp6 cells have the most interesting phenotype, as described in Chapter 1: like clr4Δ cells, they completely lose the markers of heterochromatin: H3K9me2 and Swi6 localization over the centromeric repeats is lost and the RNAi pathway is not completely functional. Moreover, they present a cell cycle arrest phenotype (Ekwall et al., 1999). Cultures of csp6 cells are unable to divide at the restrictive temperature and viability decreases to 60% after 6 hours. csp6 cells undergo normal spindle elongation and mitosis at permissive temperature, but mitosis fails to be completed at the restrictive temperature. After a 10 hour shift to restrictive temperature: 18% of csp6 cells accumulate with condensed chromosomes and a V-shaped spindle. Frequently only one SPB is detectable at the base of the V-shaped spindle by an antibody against Sad1, a spindle pole body component (Hagan & Yanagida, 1995). After 12 hours csp6 cells start dying with the disappearance of V-shaped spindles and the appearance of cells that undergo a lethal cytokinesis culminating in the “cut” phenotype (cell utterly torn/ cells divide aberrantly giving rise to nucleated and anucleated progeny or DNA that appears to be bisected by the septum). It is thought that csp6 cells arrest during prometaphase perhaps as a result of defective centromere function (Ekwall et al., 1999).
To identify the gene affected by the csp6 mutation, attempts were made to identify the affected gene by complementation with a high copy genomic library and by preliminary genetic mapping. Those analyses are presented in this chapter.

RESULTS

6.1 Attempts to csp6 using a genomic library.

The first attempt to clone csp6* by complementation of the temperature sensitive phenotype, was performed using the Shimoda genomic library (Tanaka et al., 2000; Nakamura et al., 2001). It contains Sau3AI fragments of S. pombe genomic DNA cloned in the multicopy pAL-KS plasmid bearing ScLEU2+ marker gene.

I amplified the library from E.coli. Two strains:
csp6-75 (h csp6-75 ade6-210 leu1-32 ura4DS/E imr1L(Ncol):ura4* otr1R(Sphl):ade6*) and csp6-95 (h — csp6-95 ade6-210 leu1-32 ura4-D18 imr1L(Ncol):ura4* otr1R(Sphl):ade6*), the latter a tighter ts allele, were transformed by electroporation. As diagrammed in Figure 6-1, cells transformed with the library on a LEU2 plasmid, were plated on minimal media without leucine and incubated for five days at permissive temperature, 25°C. The plates were then replica-plated on minimal media lacking leucine and with phloxin (0.02%v/v) and incubated at 36°C for three days to isolate suppressor of the temperature sensitive phenotype. Phloxin is a pink coloured substance that permits easy discrimination of dead cells (dead cells stain while live cells are able to export phloxin).

60 pale pink leu* colonies out of a total of 32000 clones were recovered at 36°C. 47 plasmids were rescued into E.coli, and minipreps analyzed by restriction pattern analysis using SalI and NotI, two restriction enzymes that cut once in the multiple cloning site of the plasmid. Plasmids were placed into five groups on basis of digestion pattern and they were digested a second time using EcoRI and SacI (Figure 6-2). Partial DNA sequencing of the insert in these five clones using universal 5' and 3' primers annealing with the multiple cloning site from the plasmid was performed. All plasmids contained one of two members of hsp70 protein family: SPAC13G7.02c/ssal on chromosome 1 (clones from group A & B) and SPCC1739.13c/ssa2 on chromosome 3 (clones from groups C,D & E).

All the plasmid clones were screened by PCR. ssa1 and ssp1, an other gene belonging to the hsp70 family, were previously found to complement the temperature sensitivity of csp6 (Karl Ekwall unpublished). Primers that anneal to ssa1 and SPAC110.04c/ssp1 were used: most of the clones amplified a band predicted for ssa1 (Figure 6-3).
Figure 6-1 First strategy to clone csp6* by complementation with a high copy genomic library.
Figure 6-2 Plasmids that suppressed csp6ts phenotype were grouped by restriction pattern.

Complementing plasmids from first transformation of csp6-75 & csp6-95 with Shimoda library were rescued in *E. coli*. Miniprep plasmid DNA was digested with *EcoRI* and *SacI* and separated on a 1% agarose gel and stained with ethidium bromide. Marker M is 1 kb ladder (Biolabs). Plasmids were categorized into groups A-E based on fragment pattern. The empty plasmid pAL-KS was digested as a control.
Figure 6-3 Most of the plasmids that suppress csp6-75/csp6-95 mutation, contain ssa1 gene. Primers designed to detect a 720 bp product within the SPAC13G7.02c/ssa1 open reading frame from hsp70 were used in PCR reaction with csp6-75 & csp6-95 complementing plasmid DNA. Representative plasmids from groups A,B,C,D,&E and all the other complementing clone PCR products, were separated on a 1% agarose gel and stained with ethidium bromide. Marker M is 100 bp ladder (Biolabs).
The same band was detected in the clones encoding for ssa2, using primers for ssa1 because they share a very high sequence identity (94%). In the clones where a band indicative of ssa1 was not amplified, the ssp1 band was detected (data not shown).

In summary from this first attempt I obtained and sequenced two genes that recovered the temperature sensitive phenotype of csp6-75 and csp6-95 alleles: ssa1 and ssa2. The clones that did not contain either of these two hsp70 genes, were found to contain a third hsp70 gene, ssp1.

A second approach was used to try to clone csp6 with the same library: again both the csp6-75 and csp6-95 strains were transformed as represented in Figure 6-4 with the Shimoda library. After transformation, the cells were spread on minimal media lacking leucine but this time incubated at 25°C only for two days. The plates were then incubated directly at 36°C for fifteen days and colonies recovered. This strategy allows the initial expression of all the potential suppressors of the csp6 mutation but might increase the possibility to find csp6 itself by incubating directly at the lethal temperature. The cells are put under conditions that required complementation for growth directly at the restriction temperature. It is possible that such conditions might select for plasmids that contain the real gene.

20 out of 20000 transformants grew at 36°C on plates lacking leucine. Only 8 plasmids out of 20, were recovered after E.coli transformation. Probably the remaining 12 clones growing on plates lacking leucine were the result of plasmid integration in the genome by recombination but this possibility was not analysed further. As before, the plasmids were digested with NotI and Sall to try to group them and then sequenced: again the genes found encoded for members of the hsp70 protein family: 6 of them were ssa2. By PCR on these plasmids with the pair of primers for ssp1 and ssa1 all of them were positive for ssa1.

From previous work performed by Karl Ekwall who tagged ssa1 with GFP and did the knockout in a diploid strain and a genetic analysis by tetrad dissection, it is known that csp6 is not this hsp70 gene. The fact that ssa1 rescued the ts phenotype in the csp6ts mutants in two different complementing screens could be due to its function as molecular chaperone that unfold and refold target proteins and thus allows the formation of silencing complexes unable to form in csp6 mutants.

To determine whether csp6 is an allele of the hsp70 gene ssa2, csp6-75 cells were crossed with cdc11-21ts allele that is linked to ssa2 on chromosome 3 (9.7 Kb apart). The expected result if csp6-75 is ssa2, is that all the progeny would be temperature sensitive. In fact 40% of the resulting progeny were temperature sensitive indicating that csp6 is not linked to ssa2 and thus excluding the possibility that cps6-75 or csp6-95 is an allele of ssa2.
Figure 6-4: Second strategy to clone csp6 by complementation with a high copy genomic library.
6.2 Mapping csp6 to chromosome III.

Vegetatively growing fission yeast cells are normally haploid. Upon nitrogen starvation, zygotes originate by pairwise fusion of cells of the opposite mating type. The two nuclei fuse and the resulting diploid nucleus undergoes meiosis immediately. The zygote develops into an ascus, with four dormant haploid ascospores. Upon return to rich media, these spores can germinate. Occasionally, vegetative growth of diploid cells can be achieved. Diploid heterozygotes can be selected by using intragenic complementation between the ade6 alleles.

Fission yeast has only three chromosomes, chromosome 1 (5.7 Mb), chromosome 2 (4.4 Mb) and chromosome 3 (3.5 Mb).

As cloning csp6 by complementation using libraries had failed, I decided to map the mutation. The first step in this strategy was to determine which chromosome the csp6 maps on. Diploid breakdown was used to determine which chromosome has the csp6 mutation. Because fission yeast rapidly haplodises when one chromosome is lost, it is easy to associate a mutant-phenotype (e.g. temperature sensitivity for csp6-95/csp6-75 allele) with the presence or absence of other auxotrophic markers that distinguish each chromosome.

As illustrated in Figure 6-5, csp6-75 or csp6-95 and csp6+ strains with specific markers on each chromosomes were made. The following marker genes were used to distinguish each chromosomes:

- chromosome 1: arg3-D4/+,
- chromosome 2: leu1-32/+ , his3-1D/+ 
- chromosome 3: ura4-D18/+ , ade6-210 (red)/ade6-216 (pink).

Diploid strains were made in duplicate by crossing the following strains already present in our laboratory collection:

Fy2637 (h csp6-95 arg3-D4 leu1-32 his3-D1 ura4-D18 ade6-210)
Fy93 (h ade6-216)
Fy2638 (h csp6-75 arg3-D4 leu1-32 ura4-D18 ade6-210)
Fy95 (h+ ade6-216)

To obtain a non-sporulating diploid (2n), the crosses were incubated at 25°C on malt extract (ME) plates for one to two days and then cells were plated on minimal media lacking adenine. This forces the cells to stay in a diploid state since ade6+ cells are only found by intragenic complementation on plates that contain no adenine source. Another easy way to monitor their diploidism is by colony-colour assay: ade6-210 cells are red and ade6-216 cells are pink on plates lacking or containing limiting adenine (10 mg/L) because they accumulate the upstream substrate which makes the colony red. This difference in colour is not observed on media containing an excess of adenine.
Figure 6-5 Diploid breakdown experiment to map csp6 to a chromosome.

This scheme represents the chromosome loss experiment performed to assay which chromosome csp6 gene is on. As explained in the text, a diploid was made crossing a wild type strain for arg3 on chromosome 1; his3, leu1 on chromosome 2; ura4 and the allele ade6-216 on chromosome 3 and a csp6-75/csp6-95 strain, deleted for the same markers genes (arg3, his3 and ura4) and with leu1-32 and ade6-210 alleles. After chromosome loss the diploid rapidly haploidizes. csp6-75/csp6-95 alleles (represented as a blue square) will segregate always with the particular marker gene that is on the same chromosome. Three outcomes (A, B, C) are presented.
The non-sporulating state of the diploids was confirmed by exposure to iodine crystals: a sporulating strain becomes black when exposed to iodine vapors because of the presence of starch in the spores.

Non-sporulating diploid genotype cells were then grown overnight in liquid minimal media lacking adenine and treated with 25μg/ml thiabendazole, (TBZ), a drug that destabilizes microtubules and causes chromosomes loss. Cells were subsequently plated on rich media containing limiting adenine (0.6mg/L) and incubated at 25°C. Pink and red haploid colonies were tested for chromosome specific markers and temperature sensitivity. The results are summarized in Figure 6-6A and Figure 6-6B: the temperature sensitive phenotype always segregated with the ura' (ura4-D18) and red (ade6-210) phenotypes. These two markers are carried on chromosome three indicating that csp6+ maps on chromosome III.

DISCUSSION

The attempts to clone csp6+ have been reported in this chapter. The failure to rescue the correct gene using the genomic library, might derive from the use of a high copy library: overexpression of the csp6+ gene itself may be toxic for the cells thus only high copy extragenic suppressors were obtained.

The approach to clone csp6 was based on the rescue of the temperature sensitive phenotype whilst the silencing defect was not used in first instance because csp6 cells variegate and it is difficult to detect absolute complementers with confidence from the background.

The fact that heat-shock protein encoding genes rescue the csp6 mutation, could suggest that Csp6 protein is misfolded in csp6-75/csp6-95 cells and Hsp70 family proteins help Csp6 to fold normally and thus suppresses the mutant phenotype. It could also be that the Csp6 protein in csp6-75/csp6-95 mutants or its complex is not stable and Hsp70 helps somehow its stabilization. However although plasmids containing ssa1, ssp1 suppress the ts phenotype, they did not suppress the silencing defect.

The other attempted strategy to clone csp6, was to map it. Using diploid mapping by an experiment of chromosome loss, csp6 has been found to map on chromosome III. csp6 is rescued by ssa2 and this gene is located on chromosome III. However a cross between csp6 cells and a temperature sensitive allele of a gene linked to ssa2, (cdc11), excluded their linkage. Also, the minichromosome Ch16 that is a derivative of chromosome III (Masumoto et al., 1989) does not complement the csp6 mutant phenotype (Ekwall et al., 1999). Ch16 was derived by chromosome breakage and it contains part of the region nearby the centromere 3 (140 Kb) but its total length is around 530 Kb (Matsumoto et al., 1987).
CROSS between Fy2637(csp6-95) and Fy256(wt)

<table>
<thead>
<tr>
<th>csp6-95</th>
<th>chrI</th>
<th>arg3-</th>
<th>his3-</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>chrII</td>
<td>leu1-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>chrIII</td>
<td>ura4-</td>
<td>ade6-210 (red)</td>
</tr>
<tr>
<td>wt</td>
<td>chrI</td>
<td>arg3+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>chrII</td>
<td>leu1+</td>
<td>his3+</td>
</tr>
<tr>
<td></td>
<td>chrIII</td>
<td>ura4+</td>
<td>ade6-216 (pink)</td>
</tr>
</tbody>
</table>

Results

\[
\begin{align*}
\text{non ts} & : \\
76/76 & : \text{ura}4^+ \text{ ade}6-216 \text{ (pink)} \\
45/76 & : \text{his}3^+ \\
36/76 & : \text{arg}3^+ \\
45/76 & : \text{leu}1^+ \\
\text{ts} & : \\
10/10 & : \text{ura}4^- \text{ ade}6-210 \text{ (red)} \\
6/10 & : \text{his}3^+ \\
6/10 & : \text{arg}3^+ \\
7/10 & : \text{leu}1^+
\end{align*}
\]

Figure 6-6A csp6 maps on chromosome 3.

Non-sporulating diploid cells treated with 25 μg/ml TBZ, were plated on YE 1/10 adenine plates and incubated at 25°C. They were then replica-plated on YES phloxine plates (incubated at 36°C), minimal media (PMG) lacking uracil, histidine, leucine or arginine supplements to follow the segregation of the temperature sensitive phenotype with each marker genes. 86 colonies were analyzed, the temperature sensitive cells were always uracil auxotrophs and red (ade6-210 allele). This clearly indicated that csp6 maps on chromosome 3.
CROSS between Fy2638(csp6-75) and Fy258 (wt)

<table>
<thead>
<tr>
<th></th>
<th>chrI</th>
<th>arg3-</th>
</tr>
</thead>
<tbody>
<tr>
<td>csp6-75</td>
<td>chrII</td>
<td>leu1-</td>
</tr>
<tr>
<td></td>
<td>chrIII</td>
<td>ura4-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>chrI</th>
<th>arg3+</th>
</tr>
</thead>
<tbody>
<tr>
<td>wt</td>
<td>chrII</td>
<td>leu1+</td>
</tr>
<tr>
<td></td>
<td>chrIII</td>
<td>ura4+</td>
</tr>
</tbody>
</table>

Results

<table>
<thead>
<tr>
<th></th>
<th>non ts</th>
<th></th>
<th>ts</th>
</tr>
</thead>
<tbody>
<tr>
<td>22/22</td>
<td>ura4+</td>
<td>ade6-216 (pink)</td>
<td>29/29 ura4- ade6-210 (red)</td>
</tr>
<tr>
<td>17/22</td>
<td>arg3+</td>
<td></td>
<td>14/29 arg3+</td>
</tr>
<tr>
<td>3/22</td>
<td>leu1+</td>
<td></td>
<td>6/29 leu1+</td>
</tr>
</tbody>
</table>

Figure 6-6B csp6 maps on chromosome 3.

Non-sporulating diploid cells treated with 25 μg/ml TBZ, were plated on YE 1/10 adenine plates and incubated at 25°C. They were then replica-plated on YES phloxine plates (incubated at 36°C), minimal media (PMG) lacking uracile, histidine, leucine or arginine supplements to follow the segregation of the temperature sensitive phenotype with each marker genes. 51 colonies were analyzed, the temperature sensitive cells were always ura- and red (ade6-210 allele). This clearly indicated that csp6 maps on chromosome 3.
Chromosome III being the shortest one (3.5 Mb) the strategy to further narrow-down this genomic locus on chromosome III will be to map it on this chromosome. Long-range mapping in *S. pombe* makes use of the swi5-39 mutation, which reduces meiotic intergenic and intragenic recombination without impairing spore viability (Schmidt and Gutz, 1989). Since recombination is decreased 10 fold in swi5-39 cells, it is as if the genetic map contracts and genes that are normally not linked become "linked".

The swi5-39 mutation was isolated in a screen for mutants with a reduced frequency of mating-type switching (Schmidt and Gutz, 1989) and has a pleiotropic effect: swi5-39 also increases the UV radiation sensitivity (Schmidt and Gutz, 1989). UV sensitivity can be used to determine if the swi5-39 mutation has been successfully introduced into temperature sensitive strains.

First it is to be assessed that csp6-95 cells are not sensitive to the amount of radiation (12 Joule/m²) which swi5-39 cells are sensitive to. Once established that csp6-95 cells are not UV sensitive the second step will be to cross double csp6-95 swi5-39 cells to other swi5-39 strains with marked genes on chromosome 3.

Chromosome III with possible marked genes to use to map csp6 is shown in Figure 6-7: ura4, set1, cds1, cdc21, mad3, alp14 are situated on the left half of chromosome III, while rad21, dcr1 and tea1 are located on the right side of cen3.

In swi5-39 background as previously mentioned, the percentage of recombination is reduced 10 fold (1% every 8 kb), therefore it is easier to detect linkage to other markers from the % of recombination obtained after their cross and sporulation. For example if two strains bearing two linked markers are crossed, the possibility that these two genes will recombine is low and their characteristic phenotypes (ts versus ura+ or KAN+)
will always segregate away from each other. For example if csp6+ is linked to a certain marker X, a cross between csp6-95 cells and cells with marker X tagged with ura4, will give very few temperature sensitive cells that are also ura4+.

The efforts in cloning csp6, have been justified by its interesting phenotype: it is the only mutants among the cspts class that loses completely the heterochromatin markers (Chapter3). csp6 cells exhibit also a particular phenotype concerning cell cycle progression and might represent an important protein for centromere architecture and function.
Figure 6-7 Narrowing down csp6 loci.
In the figure the possible tagged genes to map csp6 are represented. Also indicated are the distances in kilo-bases (kb) between one and the other. In blue the minichromosome Ch16 is also represented. It does not complement the temperature sensitivity of csp6 mutation.
INTRODUCTION

The genetic information contained in the nucleus of a eukaryotic cell is made of DNA wrapped around a core histone octamer forming nucleosomes that constitute chromatin. Distinct histone post-translational modifications act sequentially or in combination to regulate the binding of factors that promote transcriptional activity or repression. These chromatin states are proposed to represent an indexing mechanism that could extend the information potential of the genetic code (Strahl and Allis, 2000; Turner, 2000; Jenuwein and Allis, 2001). This code may thus serve as an epigenetic mechanism to store and transmit the genetic information as well as to organize heterochromatin assembly, transcriptional gene activation and genome integrity.

Methylation of specific histone lysine residues has been shown to be a mark for active euchromatin (H3K4), silent heterochromatin (H3K9, H3K27, H4K20) (Lachner et al., 2003; Sims et al., 2003) and only recently for repair of DNA damage (H3K79) (Huyen et al., 2004; Giannattasio et al., 2005). Other histone modifications such as phosphorylation of S129 on H2A, ubiquitination on K123 of H2B (Giannattasio et al., 2005), acetylation or deacetylation of K16 on H4, have been previously linked to the repair of damaged DNA (Nakamura et al., 2004; Fernandez-Capetillo and Nussenzweig, 2004).

Distinct methylation states or combinations between several methylation marks could further discriminate different chromatin regions or entire chromosomes. For example, H3K27me3 in conjunction with H3K9me2 and H4K20me1 are considered epigenetic imprints of the inactive X chromosome (Silva et al., 2003; Okamoto et al., 2004), whereas H3K9me3, H3K27me1 and H4K20me3 are associated with pericentric heterochromatin (Peters et al., 2003; Rice et al., 2003; Schotta et al., 2004). Moreover the combinatorial nature of histone lysine modifications, requires controlled interplay between different histone lysine methylation systems. Loss of a given histone-methyltransferase (HMTase) may also affect methylation on lysine (K) residues for which the enzyme has no intrinsic activity. For example in the absence of the Suv39h1 and Suv39h2, H3K9 methyltransferases, pericentric H3K27me3 rather than H3K27me1 accumulates (Peters et al., 2003).

The fission yeast Schizosaccharomyces pombe provides a convenient, genetically tractable model organism in which to study how histone modifications impact on chromatin assembly. It has only three centromeres, which resemble those of metazoans in that they are composed of
repetitive elements packed in heterochromatin that flank an array of CENP-A containing nucleosomes coating the central domain, and they occupy relatively large regions of chromosomes (40-100 Kb).

The insertion of marker genes into the repetitive outer repeats elements or the central domain results in their transcriptional silencing (Allshire et al., 1994). Fission yeast also has a lysine 9 histone H3 methyltransferase (HMTase) Clr4, that like its orthologue in metazoa (Suv3-9h1/h2), creates a binding site for a chromodomain protein Swi6 (HP1 orthologue) to mediate silent chromatin formation (Rea et al., 2000). The HMTase activity of these enzymes has been narrowed down to a highly conserved "SET" domain. The SET domain was initially identified as a 130-residue motif present in the *Drosophila* PEV-modifier SU(VAR)3-9 (Tschiersch et al., 1994), the *Polycomb*-group protein Enhancer of zeste [E(z)] (Jones et al., 1993) and in the *trithorax*-group protein Trithorax (TRX) (Stassen et al., 1995). SET domains consist of two non-contiguous regions formed by N- and C-terminal ends of the primary sequence, known as SET-N and SET-C, respectively. As shown in Figure 7-1, each region contains three to four short β-strands, a short helix and several loops that connect these secondary structural elements. SET-C forms a topologically unusual knot-like structure in which a strand threads through a loop region. The residues that form the knot (NHSCXPN) are the most conserved within the SET proteins, implying that it is of functional importance (Marmorstein, 2003).

In mammals 73 SET domain genes have been identified (Lachner and Jenuwein, 2002; Kouzarides, 2002) of which a subset is evolutionary conserved in *Drosophila melanogaster* and *S. pombe*.

Apart from Clr4, fission yeast genome encodes 13 other SET domain proteins but only Set1 has been characterized and its substrate found to be H3K4 (Noma et al., 2002; Roguev et al., 2003; Kanoh et al., 2003). Set2 in fission yeast has low similarity to Set2p, the H3K36 specific methyltransferase in budding yeast (Strahl et al., 2002; Krogan et al., 2003; Landry et al., 2003). Recently also fission yeast Set2 has been shown to methylate H3K36 (Morris et al., 2005).

Some of the other SET domain proteins would be expected to methylate other lysine residues on histones and therefore may play a role in silent chromatin formation. To address this I chose four of these SET domain proteins as being of potential interest and tested them for roles in transcriptional silencing (Table1 and Figure 7-2). Although none of these affect silent chromatin formation, Set9 is required for maintaining genome integrity after DNA damage.
Figure 7-1 Structure of SET domain proteins.
Structures of Neurospora DiM-5, Schizosaccharomyces Cir4, the garden pea Rubisco large subunit methyltransferase bound to the cofactor AdoHcy and human SET7/9 bound to AdoHcy are represented. In each protein the SET-N and SET-C subdomains are in blue, the SET-I is in green, the unusual knot structure is in orange, bound Zn atoms are in purple and the N- and C-terminal regions flanking the SET domain are in gray. The AdoHcy cofactor is in red.
Figure 7-2A Diagram of the SET domain proteins in *S. pombe*.
Schematic representation of *Schizosaccharomyces pombe* SET domain protein. The length of each protein in aa is annotated on the right. Conserved domains are indicated as follows: RRM RNA recognition motif; Chromo, Chromodomain; PHD, PHD finger; Cys-rich, Cysteine-rich domain; Ser-rich, Serine-rich domain; MYND, cluster of Cysteines and Histidines. Spbc16c6.01c is named as Set10; Spac688.14 is named as Set11; Spcc1223.04c is named Set12; Spbc1709.13c is named as Set13.
## Table 7-2B

<table>
<thead>
<tr>
<th><strong>S. pombe</strong></th>
<th><strong>S. cerevisiae</strong></th>
<th><strong>Human</strong></th>
<th><strong>Substrate</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Set1</td>
<td>Set1</td>
<td>Set1 family</td>
<td>H3-K4</td>
</tr>
<tr>
<td>Set2</td>
<td>Set2</td>
<td>Set2 family</td>
<td>H3-K36</td>
</tr>
<tr>
<td><strong>Set3</strong></td>
<td>Set3/4</td>
<td>MLL5</td>
<td>unknown</td>
</tr>
<tr>
<td>Cln4</td>
<td>-</td>
<td>Suv3-9 family</td>
<td>H3-K9</td>
</tr>
<tr>
<td>Set5</td>
<td>Set5</td>
<td>-</td>
<td>unknown</td>
</tr>
<tr>
<td>Set6</td>
<td>Set6</td>
<td>SMYD3</td>
<td>unknown</td>
</tr>
<tr>
<td><strong>Set7</strong></td>
<td>-</td>
<td>-</td>
<td>unknown</td>
</tr>
<tr>
<td>Set8</td>
<td>YPL208W</td>
<td>FLJ21148</td>
<td>unknown</td>
</tr>
<tr>
<td><strong>Set9</strong></td>
<td>YPL208W</td>
<td>FLJ21148</td>
<td><strong>H4-K20</strong></td>
</tr>
<tr>
<td>SPBC16C6.01c</td>
<td>-</td>
<td>Suv4-h20 family</td>
<td>unknown</td>
</tr>
<tr>
<td>SPAC688.14</td>
<td>RMS1</td>
<td>FLJ21148</td>
<td>unknown</td>
</tr>
<tr>
<td>SPCC1223.04c</td>
<td>YDR198c</td>
<td>C21orf18</td>
<td>unknown</td>
</tr>
<tr>
<td>SPBC1709.13c</td>
<td>YPL208W</td>
<td>FLJ21148</td>
<td>unknown</td>
</tr>
</tbody>
</table>

**Figure 7-2B, Table 1.** Blast search analysis was performed using protein sequence from *Schizosaccharomyces pombe* SET domain proteins, as defined by SMART database (left column). The most related proteins identified from *Saccharomyces cerevisiae* and humans and known substrates are listed across the rows. The names of SET domain families is defined as previously published (Kouzarides, 2002; Schotta et al., 2004). The SET domain protein deleted in this study are indicated in red.
RESULTS

7.1 Deletion of four genes encoding SET domain proteins.

Analysis of mouse cells lines generated from the T. Jenuwein laboratory in Suv3-9h1/2 double KOs, indicated that H3K9me1 persists at pericentric heterochromatin and suggested the presence of a H3K9mono-methyltransferase (Peters et al., 2003; Rice et al., 2003). Preliminary data (not shown) also suggested this could have been the case in fission yeast. In order to try to identify the putative monomethyltransferase for H3K9 in fission yeast, I deleted three "SET" domain encoding genes which had clear homologs in mammalian cells but that lacked orthologs in *S.cerevisiae* where there is no evidence of H3K9 methylation. In addition I deleted the ORF related to *S.cerevisiae* Set3, the substrate of which is unknown. The four chosen genes were: SPAC22E12.11C, known as *set3*, SPBC16C6.01C, called *set10* in this thesis; SPCC297.04C, known as *set7*, SPCC4B3.12 known as *set9* for the remainder of the thesis.

*set3* encodes for a protein of 859 amino acids that has 24% identity and 42% similarity over a stretch of 679 amino acids to Set3p in *S.cerevisiae*. Set3p is part of a complex with a histone deacetylase activity that represses meiotic specific sporulation gene expression (Pijnappel W et al., 2001). Set3 also has a PHD finger domain that is present in chromatin regulators and transcriptional cofactors (CBP1p300, ATRX, DNMT3, TIF1/KAP-1)(Capili et al., 2001). Its substrate and function in fission yeast are still unknown. The presence of both SET and PHD finger domains in fission yeast Set3 is reminiscent of *Drosophila* Trithorax and may be important for its role in some chromatin remodelling activity. Trithorax is involved in epigenetic regulation of developmentally regulated genes and works to maintain active transcription states through cell division. It is histone methylation that serves as a mark for Trithorax complexes. Moreover the presence of the budding yeast counterpart in a complex involved in inhibiting the expression of genes involved in a sort of "differentiation" mechanism for yeast cells (meiosis), was another good reason to make its deletion and analyze its phenotype.

*set10* encodes for a protein of 473 aa that has 25% identity and 41% similarity over a stretch of 270 aa with human FLJ21148, a member of the SET domain protein family. It has been found to interact with Pmt1 by TAP-tagged pull down (Ladurner A. personal communication). Pmt1 is the inactive DNA methyltransferase present in fission yeast. So far no evidence of DNA methylation has been found in fission yeast and neither *pmt1*Δ cells present silencing defect phenotypes (Grace Goll & Bestor T., 2005). Despite this, the fact that Pmt1 interacts with Set10 may be important for some still unknown histone modifications catalized by the two together in a complex.

*set7* encodes for a very short protein of 147 aa that has 29% identity and 46% similarity to CG2995, a *Drosophila melanogaster* protein with unknown function.

*set9* encodes for a protein of 441 aa that has ortholog both in human cells and in Drosophila. It has 27% identity and 43% similarity over a stretch of 379 aa with Suv420h1 and 30% identity.
and 48% similarity over 234 aa with Suv420h2, a protein with high similarity to suppressor of variegation 4-20 homolog 2 (mouse Suv420h2), which is a heterochromatic histone H4 lysine 20 trimethylating histone methyltransferase.

**set3, set10, set7 and set9** were deleted using PCR-based gene targeting with the kanamycin module as described in Materials and Methods, which renders cells resistant to G418. To allow immediate assessment of effects on silent chromatin function, each gene was deleted in a haploid wt strain (FY1180) with the ade6° marker gene inserted within the outer repeats of centromere 1 cen1otr1R(SphI):ade6° (Ekwall et al., 1997). Resulting transformants were selected on plates containing 0.1mg/ml G418 and incubated at 25°C. G418R transformants were screened for deletion of each gene using the following pairs of primers: the forward primer specific for the 5' region of each SET domain gene (located ~180 bp upstream of the ATG) and the reverse primer common for all, complementary to the KAN gene cassette. A specific band of size ~930 bp was expected for transformants in which the **set3, set10, set7** or **set9** genes had been deleted. In each case, such band was detected, indicating that the four genes had been individually deleted (Figure 7-3A). A second PCR test was performed to check for the deletion of these genes. The forward primers described before were used with reverse primers complementary to the gene sequence and located approximately 600 bp 3' of the ATG. As expected, bands approximately of about 730 bp were amplified only using genomic DNA from a wild type strain and not when DNA from SET gene deleted strains was used as template (Figure 7-3B).

This confirmed that **set3, set10, set7 and set9** genes had been successfully deleted and that they were not essential. If one of them had been essential, the deletion would have been carried out in a diploid strain and analysis of tetrads performed.

### 7.2 Viability and phenotypes of the SET deleted strains.

To uncover mutant phenotypes associated with **set3, set10, set7** and **set9** gene deletions, the SET deleted strains were tested for growth under different conditions. In the laboratory the normal temperature for growth of wild type fission yeast cells is 32°C. At 36°C the growth of fission yeast temperature-sensitive mutants is impaired whilst 18°C and 25°C are the restrictive temperature for cold-sensitive mutants.

Serial dilution assay on rich media (YES) or complete minimal media (PMG) at different temperatures; 18°C, 25°C, 32°C and 36°C, was done to assess growth defects in any of the SET protein deleted cells. As shown in Figure 7-4 only **set3** deleted cells exhibited slight cold sensitivity at 18°C, 25°C and slow growth at 36°C on rich and minimal media compared to wild type cells. This slower rate of growth was similar to that seen in a **clr4Δ** strain. No general defects in cell growth for strains lacking **set10, set7** and **set9** genes was found both on rich and minimal media at the different temperatures.
Figure 7-3 Deletion of set3, set10, set7 and set9 genes with kanamicyn (KAN) cassette.
Integration of the kanamicyn (KAN) cassette at the set loci (set3, set10, set7 and set9) was verified by 2 different PCRs on yeast genomic DNA.
In the first PCR (panel A) the following primers were used: forward primer upstream of each SET domain locus set ctrl 5'; reverse primer within the KAN cassette kan 3' ctrl. The presence of a specific band of 930 bp was amplified in the knock out cells. In the second PCR (panel B) the following primers were used: forward primer upstream of each SET domain locus set ctrl 5'; reverse primer within the SET gene set ctrl 3'. The absence of the wild type band around 730 bp in the knock out cells compared to wild type cells is shown. M indicates 1 kb ladder; m is the 100 bp ladder.
Figure 7-4 Testing cold and temperature sensitivity of set3, set7, set9 and set10 deletion strains.

To test viability at different temperatures, serial dilution of cells were plated on rich media YES (A) or minimal media PMG (B) at 18°C, 25°C, 32°C and 36°C. A wt strain (FY1180) was plated as a control together with a clr4Δ (FY2009) strain. Pictures were taken after 7 days incubation at 18°C, 5 days at 25°C, 3 days at 32°C and 36°C. set3Δ cells grow slower at 18°C, 25°C and 36°C on both rich and minimal media compared to wild type cells. Mutant in H4 on K20 residue (h4.2K20R) was also tested. It showed similar growth at all temperatures to the wild type h4.2K20 control.
7.3 Monitoring chromatin silencing defects in SET deleted strains.

The centromere, telomere and mating type loci of fission yeast display features of heterochromatin. These regions are characterized by repeat sequences occupied by protein complexes that result in transcriptional repression of marker genes artificially inserted within these DNA sequences (Allshire et al., 1994, 1995; Ekwall et al., 1995, 1996). At the centromere the formation of silent chromatin appears to reflect the formation of a fully functional centromere/kinetochore. Experiments were carried out to determine whether set3Δ, set10Δ, set7Δ and set9Δ cells showed defective silencing at these heterochromatic loci.

7.3.1 Silencing at the outer repeats and central domain of centromere 1.

Two distinct domains are present within fission yeast centromeres: the outer repeats and the central domain.

The ability of the SET deleted strains to repress the ade6+ marker gene inserted into the Sphi site at the otr/imr junction of centromere 1 (cen1otr1R(Sphi):ade6+) was assessed by a colony colour assay. Due to the heterochromatic nature, genes artificially inserted within the outer repeats of centromere are normally repressed. Repression of ade6+ gene renders cells ade6. This leads to the accumulation of a substrate of adenine metabolism that confers red colour to colonies grown on limiting adenine plates. When heterochromatin is disrupted, such as in a clr4Δ background, the ade6+ gene is expressed and colonies are white. Variegation of silencing results in a mixture of red, pink and white colonies that arise when in genetically identical cells, the transcription of the marker gene flips between on and off states. As shown in Figure 7-5A wild type cells (FY1180)(cen1otr1R(Sphi):ade6+) form red colonies due to silencing of ade6+, clr4Δ cells (FY2009) (Ekwall et al., 1997) alleviate this silencing, resulting in white colonies. All the SET gene deleted cells (cen1otr1R(Sphi):ade6+) show red colonies indicating they do not affect silencing in the outer repeats of cen 1.

Similar results were obtained upon examination of silencing of the ura4+ marker gene at the same site in centromere 1 cen1otr1R(Sphi):ura4+ (Figure 7-5B). set3Δ, set10Δ, set7Δ and set9Δ cells (cen1otr1R(Sphi):ura4+) grow as poorly as wild type cells (FY501) on plates lacking uracil and they grow well on plates with counter-selective FOA indicating that silencing is intact. In contrast, clr4Δ cells (FY6084) alleviate otr:ura4+ and the resulting increase in the expression allows faster growth on plates lacking uracil and loss of growth on plates with FOA (Figure 7-5B).

The silencing of the ura4+ marker gene inserted at the innermost repeats of centromere 1(imr1) was also assessed. set3Δ, set10Δ, set7Δ and set9Δ cells grow to the same extent as wild type cells bearing centromeric ura4 (cen1imr1L(NcoI):ura4+) and less than clr4Δ cells (FY2154) on plates lacking uracil. Wild type and all the set protein deleted cells grow much more than clr4Δ cells on FOA plates (Figure 7-6).
SET gene deleted strains containing the *ade6* marker gene inserted at the outer repeats of centromere 1 were plated at 32°C on YES and YE media containing 1/10 of the normal supplement of adenine to reveal their colour. Red colour indicates repression of the *ade6* gene (wt/FY1180) whereas white colour indicates *ade6* expression (*clr4ΔFY2009*), thus alleviation of silencing. Cells were plated at 32°C on YES or YES supplemented with 15 μg/ml TBZ and growth was compared to the TBZ sensitive mutant *clr4* and the control wt. Mutants in H4 on K20 residue (*h4.2K20R*) did not show any defect in silencing of *cen1:otr1Lade6* and were not TBZ sensitive when compared with the wild type control *h4.2K20*.

Figure 7-5A Silencing at centromere 1 outer repeats in *set3Δ, set10Δ, set7Δ* and *set9Δ* cells remains intact and they are not sensitive to TBZ.
Figure 7-5B set3, set10, set7 and set9 are not required for silencing of ura4 in outer repeats of the centromere 1.

set deleted strains containing the ura4+ marker gene inserted at the outer repeats (otr1L) of centromere 1 were tested for alleviation of silencing on PMG medium lacking uracil or with counterselective 5-fluoroorotic acid (FOA) at 32°C. Cells were plated on non-selective (N/S) plates as a control for growth and number. Wild type (Fy4841) cells and control clr4Δ (Fy6084) cells were plated for comparison.
Figure 7-6 Silencing in the imr region of cen1 remains intact in set3Δ, set10Δ, set7Δ and set9Δ cells.

set deleted strains containing the ura4+ marker gene inserted at the immermost repeats (imr1R) of centromere 1 were tested for alleviation of silencing on PMG medium lacking uracil or with counterselective 5-fluoroorotic acid (FOA) at 32°C. Cells were plated on non-selective (N/S) plates as a control for growth and number. Wild type (FY500) cells and control clr4Δ (FY2154) cells were plated for comparison.
The central domain consists of the inner part of imr and the central core (cnt). Silencing across central core (cnt1) was analysed using a strain bearing arg3<sup>+</sup> at the central core (cnt1(Ncol):arg3<sup>+</sup>) in the central core. While the temperature sensitive mutant sim4-193 (FY6372) can grow on – arginine plates, no increase in growth could be detected for any of the deleted SET domain strains compared to wild type cells (FY3027) when grown at 25°C (Figure 7-7).

Many mutants defective in centromere/kinetochore structure have defects in interactions with spindle microtubules and are often sensitive to microtubule destabilizing drugs such as thiabendazole, TBZ and Benomyl, which bind tubulin dimers and shift the equilibrium between polymerization and degradation towards the degradation. These drugs, interfering with microtubule polymerization, affect this process also during mitosis. The centromeres of strains bearing mutations of proteins involved in centromere function, such as the c1r4ii mutant, are often impaired in their interaction with microtubules. The presence of TBZ exacerbates this defect, reducing cell viability.

set3Δ, set10Δ, set7Δ and set9Δ cells like wild type cells displayed no increased sensitivity to TBZ at a concentration of 15 μg/ml at 32°C whereas the c1r4Δ control cells clearly were sensitive (Figure 7-5A).

7.3.2 Silencing at the mating type locus and telomeres.

As discussed previously in the Introduction (Chapter 1), heterochromatin in fission yeast is assembled not only at the centromere but also over the silent mating type loci (mat2-mat3) and adjacent to telomeres. To investigate if any of these SET domain proteins has a specific role in silent chromatin formation at these loci, silencing of the ura4 marker gene close to mat3 (mat3(EcoRV):ura4<sup>+</sup>), and his3<sup>+</sup> inserted 300 bp from the end of chromosome 1 (tellL:his3<sup>+</sup>) at tel1 was assessed. Strains bearing these markers (FY511, FY3027) were crossed to strains bearing deletion of set3, set10, set7 and set9 genes. The presence of the marker genes, ura4<sup>+</sup> and his3<sup>+</sup> was verified by PCR. No silencing defects in any of the SET protein deleted strains, assayed by plating on plates lacking uracil or histidine was evident compared to wild type cells, whereas swi6Δ cells (FY620 or FY1954) lose silencing at the mating type loci and at the telomere and can grow on plates lacking uracil or histidine respectively (Figure 7-8).

7.4 Analysis of Set9.

7.4.1 H4-K20 is methylated in fission yeast and Set9 is the methyltransferase responsible for this modification.

In mammalian cells and in Drosophila melanogaster, H4K20 is methylated and in particular H4K20me3 is enriched across pericentric heterochromatin of mammalian chromosomes and in the chromocenter of Drosophila (Schotta et al., 2004). One characteristic of these regions is that they have high levels of H3me3K9 and bound the chromo-domain protein HP1 (Peters et
Figure 7-7 Silencing at the central core remains intact in set3Δ, set10Δ, set7Δ and set9Δ cells.

set deleted strains containing the arg3+ marker gene inserted at the central core (cnt1) of centromere 1 were tested for alleviation of silencing on PMG medium lacking arginine at 25°C. Cells were plated on non-selective (N/S) plates as a control for growth and number. Wild type (FY 3027) cells and control sim4-193 (FY 6372) cells were plated for comparison.
Figure 7-8 set3Δ, set10Δ, set7Δ and set9Δ cells maintain silencing at the mating type locus and at telomeres.

Top panel. SET gene deleted strains containing the ura4+ marker gene adjacent to mat3 (mat3:ura4+) were plated on selective minimal media lacking uracil or containing counterselective 5-fluoroorotic acid (FOA) and incubated at 32°C. In swi6Δ cells (FY620) silencing of mat3:ura4+ is alleviated and they provide a positive control. Cells were plated also on non-selective plates (N/S) as a control for growth and number.

Bottom panel. set deleted mutants containing the his3+ marker gene at telomere 1 (tel1:his3+) were plated on selective minimal media lacking histidine and incubated at 32°C. In swi6Δ cells (FY1954) silencing of tel1:his3+ is alleviated and they provide a positive control. Cells were plated on non-selective plates (N/S) as a control for growth and number.
al., 2003; Rice et al., 2003). The enzyme responsible for H4K20me3 belongs to the Suv4-h20 family. Suv4-h20 is a histone lysine methyltransferase (HKMT) related to Set9 in S. pombe. It has been shown that Suv4-h20 is necessary for heterochromatin formation in mammalian cells and Drosophila melanogaster (Schotta et al., 2004).

The analysis presented above demonstrated that Set9 deleted strain did not display any defects in its ability to form heterochromatin.

Data from our collaborators, Steve Sanders and Tony Kouzarides at the CRUK in Cambridge, had shown that purified TAP tagged Set9 exhibits a methyltransferase activity on nucleosomal histone H4, but not on free H4. Homology to Suv4-20 and literature concerning lysine methylation on H4 suggested that H4K20 would be a good candidate for the substrate of Set9 (Schotta et al., 2004; Sanders et al., 2004). Antibodies specific for H4K20me1, H4K20me2 and H4K20me3 obtained from Abcam were tested by western analysis of histones prepared from wild type cells and of recombinant unmodified fission yeast H4: a specific band was only detected on native fission yeast H4 (performed by Sanders).

To rigorously test the specificity of these antibodies, K20 residue in H4 of fission yeast cells was mutated to arginine (R) in a strain bearing only one copy of H4 out of the three present in wild type cells (Fy 4756) (Mellone et al., 2003). The construction of this strain is reported in Chaper 2 (Materials and Methods). Western analysis was performed on histones extracted from strains bearing only the mutated *h4.2K20R* or wild type *h4.2K20*. As shown in Figure 7-9 (right panel) the signals detected by the three antibodies against H4K20mel, H4K20me2 and H4K20me3 were dependent on the presence of lysine at position 20 whereas control antibodies that recognize unmodified H4 and H3 revealed the presence of H4 and H3 in *h4.2K20R* mutant extracts. We concluded that fission yeast H4 can be specifically mono, di- or tri-methylated on K20.

The *h4.2K20R* mutated cells were tested for defects in growth, silencing and TBZ sensitivity as described for *set9Δ* cells. These analyses detected no apparent defect in *h4.2K20R* cells compared to *h4.2K20* control wild type cells. Both *h4.2K20R* and *h4.2K20* wild type cells exhibit slightly growth impairment at 18°C, 25°C and 36°C. It is likely that this is due to the presence of only one copy out of the three pairs of histone genes in the strain used to create cells expressing a single mutant H4 gene (Figure 7-5A and Mellone et al., 2003).

To establish whether Set9 is responsible for this methylation, western analysis on protein preparations from *wt, set3Δ, set10Δ, set7Δ, set9Δ* and *clr4Δ* cells was performed using antibodies specific for H4K20me1, H4K20me2 and H4K20me3. Their signals are strictly Set9 dependent and they are not influenced by any of the other genes encoding SET domain proteins. Equal amounts of histones were loaded in each sample as revealed by the western analysis using antibodies specific for unmodified H4 and H3 (Figure 7-9 left panel). Together these data and the results relative to the *h4.2K20R* mutant confirmed that H4 in fission yeast is methylated on K20 and Set9 is the protein responsible for H4K20 methylation.
Figure 7-9 Histone H4 is methylated on K20 in fission yeast and Set9 is responsible for this modification.

In the left panel it is presented a Western analysis on protein preparations from wt, set3Δ, set10Δ, set7Δ, set9Δ, cfr4Δ cells probed with specific antibodies for H4K20me1, H4K20me2, H4K20me3. The same protein preparations were tested with antibodies that recognize unmodified H3 and H4 as loading control. Only in set9Δ cells H4K20 modification is lost.

In the right panel protein extracts derived from fission yeast cells containing only a single h4.2K20 or h4.2K20R allele were subjected to immunoblot with specific antibodies for H4K20me1, H4K20me2, H4K20me3. The same preps were tested with antibodies that recognize unmodified H3 and H4 as loading control. The antibodies recognized specifically H4K20me1, H4K20me2 and H4K20me3 in a wild type strain (h4.2K20) and the signals is lost in point mutants (h4.2K20R). Thus histone H4 is methylated on K20 in fission yeast.
7.4.2 H4K20me3 localization.

As discussed in the Introduction, H4K20me3 is concentrated at pericentric regions in mammalian cells and a pathway to establish this heterochromatin has been proposed: initially H3K9me3 occurs, allowing the binding of the chromodomain protein HP1. Subsequently H4K20me3 appears and acts to lock in place this state of silent chromatin. Consistent with this, H4K20me3 is lost in Suv3-9h1/h2 double knock out mouse embryo fibroblast cells where H3K9 is not trimethylated and HP1 is not bound (Schotta et al., 2004).

In mammalian cells, although Suv4-20 is responsible for H4K20me3 modification, H4K20me1 and H4K20me2 are catalyzed by PRSet7I8, SET domain protein. These modifications are implicated in other functions: H4K20me1 is involved in X-chromosome inactivation and H4K20me2 is localized predominantly in euchromatin. Their localization is cell cycle regulated (Kohlmaier et al., 2004; Julien and Herr, 2004; Nishioka et al., 2002).

To investigate the localization of H4K20 methylation in fission yeast, fixed cells were stained with the anti-H4K20me3 antibody and DAPI. The immunolocalization pattern indicates that H4K20me3 is distributed over the entire genome and its localization is lost in set9A and h4.2K20R mutants (Figure 7-1A and 7-10B). H4K20me3 is not a specific centromere marker. In fact costaining of cells with anti-H4K20me3 and anti-Cnpl (kinetochore specific protein) demonstrates that there is no clear concentration of H4K20me3 at centromeres as detected by anti-Cnpl (Figure 7-1A). Furthermore, H4K20me3 localization is not dependent on the HP1 ortholog, Swi6, or the Suv3-9 ortholog, Clr4: a lagging chromosome in swi6A cells does not lose H4K20me3 localization and it is completely decorated. Similar results were obtained in cells with clr4A gene deleted (Figure 7-11B). The absence of H4K20me3 does not interfere with Cnpl localization at the centromere in set9A cells (Figure 7-11A).

These results together with the data collected from the western analysis, indicate that S. pombe Set9 specifically modifies K20 of H4 in vivo.

The immunolocalization data indicate that H4K20me3 is distributed on all the chromatin. It is possible that H4K20me3 is over the entire genome and it is excluded in centromere regions. To assess the association of H4K20me3 with centromeric sequences, Chromatin Immunoprecipitation (ChIP) experiments with the same H4K20me3 antiserum were performed. Wild type, set9A and clr4A cells were fixed. Sonicated extracts were immunoprecipitated and enrichment of centromeric sequences (otr, imr, cnt) relative to euchromatic genes (fbp1) was assessed (Figure 13A). These analyses suggest that nucleosomes coating centromeric DNA lack H4K20me3. The fact that in clr4A cells, where H3K9me2 and heterochromatin are lost, H4K20me3 can still be detected only at euchromatin (fbp1) and not at centromeric DNA, excludes the possibility that failure to IP centromeric sequences is due to centromeric chromatin inaccessibility.

As H4K20me3 is absent from the centromere (does not colocalize with Cnpl) and the nucleolus by immunofluorescence, the same immunoprecipitated DNA was tested in a PCR using specific primers for rDNA sequences compared to the euchromatic gene fbp1. H4K20me3 does not
Figure 7-10A H4K20me3 localizes to the nucleus and its nuclear localization is lost in set9Δ cells. Wild type cells (panel A) and set9Δ cells (panel B) were stained with rabbit anti-H4K20me3 antibody (Abcam)(green), and DAPI (red).
Figure 7-10B H4K20me3 staining is lost from chromosomes in H4K20R mutant cells. Wild type (panel A) and h4.2K20R cells (panel B) were fixed and processed for immunolocalization with rabbit anti-H4K20me3 (green) antibody and DAPI (blue).
Figure 7-1A H4K20me3 does not colocalize with Cnp1 and Cnp1 is still localized in set9Δ cells. Wild type cells (panel A) and set9Δ cells (panel B) were stained with rabbit anti-H4K20me3 antibody (Abcam)(green), sheep anti-Cnp1 antibody (red) and DAPI (blue).
Figure 7-11B H4-K20me3 remains localized with chromatin in *swi6Δ* and *clr4Δ* cells.

*swi6Δ* cells (panel A) and *clr4Δ* cells (panel B) were stained with rabbit anti-H4K20me3 antibody (Abcam)(green), sheep anti-Cnp1 antibody (red) and DAPI (blue). A lagging chromosome is shown by an arrow in *swi6Δ* cells.
appear to be associated with rDNA although this analysis only samples a small region in the
10.4 kb rDNA unit (between -28S-18S ribosomal DNA) (Figure 13B). The same result was
obtained by immunofluorescence where it was clear that H4K20me3 staining colocalized only
with DAPI staining DNA but was absent from the nucleolar sector of the nucleus (Figure 7-11A).
Immunostaining with antibodies specific for H4K20me1 and H4K20me2 was performed but the
signal obtained was considered non specific: there was no difference between wild type cells
and in set9Δ cells while by western analysis the difference was clear.

7.4.3 Set9 immunolocalization.
The localization of Set9 was assessed in a strain in which endogenous Set9 was fused to GFP
at its C-terminus. This fused protein is functional: protein preparation from Set9-GFP expressing
cells, assayed by western analysis, retains H4K20 methylation (Steve Sanders).
Set9-GFP appears to be distributed over the entire nucleus including chromatin and nucleolus.
Staining is punctuate with brighter signals within the nucleolus (Figure 12). However, as shown
by immunofluorescence in Figure 11A and by ChIP in Figure 12, H4K20me3 staining associated
with chromatin and appears to be excluded from the nucleolus.
To further investigate Set9 nucleolar localization, Set9-GFP expressing cells were costained
with anti-GFP and anti-Nopi (a marker for the nucleolus) antibodies. The resulting images
demonstrate overlapping of some Set9-GFP punctuate staining with Nop1 indicating that Set9-
GFP could associate with the nucleolus. This localization could result from distribution of Set9
throughout the nucleoplasm rather then association with chromatin.
Costaining with anti-GFP and anti-Cnp1 (a centromeric specific marker) suggests that the Set9
protein may be excluded from clustered centromeres in interphase cells since none of the Set9-
GFP spots were localized at the centromere (Figure 15). However more cells would need to be
analysed in detail to allow strong conclusions.
From these experiments we concluded that Set9 is distributed throughout all the nucleus while
H4K20me3 covers only chromatin; neither Set9 nor H4K20me3 are specifically associated with
centromeric DNA.

7.4.6 Set9 and Methyl-K20H4 have a role in genome integrity.
Set9 and H4K20 are not involved in heterochromatin formation or in transcriptional gene
activation. From microarray data performed on RNAs from set9Δ cells compared to wt cells, by
Juan Bata (Sanger institute Hinxton-Cambridge), of the 4641 genes scored in two out of two
experiments, 99% displayed a change (either up or down) of <30%. The expression of only one
gene ipk1+ was altered >2-fold, a change considered significant; this gene is located just 2.5 Kb
from set9+ on chromosome III. The induced expression of ipk1+ may be indirectly due to the
presence of kanamicyn gene at set9 loci in set9Δ cells (Juan Bata).
In order to determine the role of Set9 and H4K20me in fission yeast, Steve Sanders tested the
response to different stresses such as osmotic, heavy metal, oxidative stress: set9Δ cells
Figure 7-12 The H4K20me3 modification occurs on euchromatin and is not a mark for centromeric sequences. H4K20me3 does not occur on rDNA.

ChIP was performed with anti-H4K20me3 antibody on wt, set9Δ and clr4Δ cells.

Panel A: the immunoprecipitated DNA was analyzed by multiplex PCR using four primer pairs designed to amplify the outer repeats (otr), the central core (cnt), the innermost repeat (imi) of centromere1 and fbp1 gene as an euchromatic control.

Panel B: the immunoprecipitated DNA was analyzed by multiplex PCR using two primer pairs designed to amplify the rDNA loci (in the nucleolus) and the fbp1 gene as an euchromatic control.
displayed viability comparable to wild type cells under all these different conditions. Phenotypic analysis did reveal, however, hypersensitivity of set9Δ cells and K20R mutants to DNA damaging agents that cause DNA double-strand breaks (DSBs) such as ultraviolet light (UV), ionizing radiation (IR) and the topoisomerase I poison, camptothecin. Loss of Set9 or mutation of K20 to R did not increase sensitivity of cells to hydroxyurea (HU) a drug that cause replication fork stalling.

From genetic analysis (Sanders et al., 2004), Set9 and H4K20me are not involved in DNA repair since the two pathways usually activated, homologous repair (HR) and excision repair, are intact and functional. Instead, Set9 and H4K20 appear to act in the DNA damage or replication checkpoints. Genetic interactions and the fact that set9Δ and h4.2K20R mutants are not sensitive to hydroxyurea (HU), linked Set9 and h4.2K20 function with the DNA damage checkpoint. In particular Crb2, the BRCT domain protein that becomes phosphorylated after DNA damage and activates Chk1 kinase, is strictly dependent for its functionality on the presence of wild type Set9 and H4K20 (Sanders et al., 2004). From studies by Nakamura et al., 2004 another post-translation modification, phosphorylation on H2A (S129), appears at the DNA double-strand (DSBs) breaks induced by IR and is required for the action of Crb2 to sustain the DNA damage checkpoint in the process of damage. However after IR treatment, set9Δ cells induced H2A phosphorylation at the same levels as wild type cells by western analysis. However, Crb2 recruitment at the DNA foci is compromised in set9Δ cells (Sanders et al., 2004).

Wild type and set9Δ cells treated with 127 Gy IR and fixed to perform immunofluorescence using H4me3K20 antibody did not show any redistribution of H4K20me3. The same experiment performed on Set9-GFP expressing cells did not reveal any change in the localization of Set9 after IR exposure (data not shown).

From these series of experiments we concluded that Set9 and H4K20me have a role in genome integrity via DNA damage checkpoint activation.

**DISCUSSION**

From this initial study, deletion of none of the SET genes affect silent chromatin formation. Set3, Set7, Set10 whose targets have not been identified yet, could be involved in the methylation of the following lysines in H3 and H4 known to be methylated in human nucleosomes: H3K14, H3K23, H3K27, H3K79, H4K12, H4K59, H4K79. A preliminary analysis could be done by western analysis on the SET deleted strains to check for the absence of a particular H3Kme or H4Kme compared to wild type strain. Mass spectrometry on histone preparations from wild type and SET deleted cells could be also done to determine if the modification pattern of histones is altered in wild type versus SET mutants. This could suggest which SET protein is required.
Figure 7-13 Set9 localizes to the nucleus.
Cells expressing Set9-GFP were fixed and processed for immunolocalization with rabbit anti-GFP antibody (1:100) (green) and DAPI staining (Blue).
Figure 7-14 Set9 localizes to the nucleolus.
Set9-GFP expressing cells were fixed and processed for immunolocalization with rabbit anti-GFP antibody (green), monoclonal Nop1 antibody (a nucleolar marker)(red) and DAPI staining (blue). Overlapping signals between Nop1 and Set9-GFP staining are indicated by the arrows.
Figure 7-15 Set9 does not localize at centromeres.
Cells expressing Set9-GFP were fixed and processed for immunolocalization with anti-GFP antibody (green), anti-Cnp1 antibody (red) and DAPI staining.
It is known that H3K4 is methylated by Seti (Noma et al., 2002) and H3K9 is methylated by Clr4 (Bannister et al., 2001). Recently it has been shown that Set2 is responsible for H3K36 methylation (Morris et al., 2005) but the substrates of all the other set domain proteins are still unknown.

Another important point to consider is that each lysine can have different methylation states (mono-, di- and trimethylated states). These modifications can be catalyzed by the same or different enzymes. Having very specific antibodies for the different methylation states of a particular lysine is an absolute requirement to allow such investigations.

**Set3**

In fission yeast Set3 depleted cells exhibit a cold sensitive phenotype when grown on minimal media. Its role could be similar to Set3p in *S. cerevisiae*: Set3p is in a complex with two histone deacetylases, Hos2 and Hst1 which function in the repression of specific sporulating meiotic genes (Pijnappel et al., 2001). Its role during meiosis in fission yeast has not been analysed in this study. Both the SET and PHD finger domains in spSet3 and scSet3 proteins are present in mammalian MLL5. Myeloid/lymphoid or mixed-lineage leukaemia 5 is a trithorax *Drosophila* homolog that may regulate transcription through protein associations. In *Drosophila melanogaster*, Trithorax proteins are activators that maintain *Hox* gene transcription during development of anterior-posterior axial structures (McGinnis et al., 1992; Duboule et al., 1994). An involvement of Set3 in regulating the transcription of a specific set of genes during meiosis in *S. pombe* could be investigated by microarray analysis in wild type versus set3Δ cells.

**Set10**

No phenotype was found for set10 deleted cells. Its mammalian homolog has not been characterized yet and there are no hints on what its substrate may be. It has been found to interact with Pmt1, the fission yeast inactive DNA methyltransferase homolog (Ladurner personal communication). This suggests a possible role for Set10 in mediating chromatin modification only when it is in a complex with Pmt1. Further studies need to be done to investigate this interaction and Set10 activity.

**Set7**

set7 deleted cells are not affected in growth or silencing at the heterochromatic loci in fission yeast. Set7 has an homolog in *Drosophila* whose function is still unknown. Set7 is a very small protein of just 147 aa and it could be acting in a complex together with other proteins.

**Set9**

Set9 has been found to be responsible for H4K20 methylation in fission yeast. Several experiments confirm this: Set9tap-tagged exhibits activity on H4, western blot analysis with specific antibodies recognizing H4K20me1, H4K20me2 and H4K20me3 revealed the disappearance of these modifications in set9Δ cells and in strains where H4K20 residue was mutagenized to arginine (H4K20R).
Set9 is not required for heterochromatin formation in S. pombe, in contrast to its homologs in mammalian cells and Drosophila, Suv4-20 (Schotta et al., 2004). set9Δ cells do not present any defect in growth at different temperature on rich or minimal media.

Instead Set9 has a role in maintain DNA damage checkpoint activated. In set9Δ cells the localization of Crb2 to DSBs is compromised and this impairs the ability of cells to maintain cell cycle arrest after IR. In fact by genetic analysis between Set9 and checkpoint proteins, Set9 functions in a checkpoint Rad-protein dependent pathway downstream the activation of the protein kinase Rad3.

In mammalian cells and in S. cerevisiae this role has been attributed to Dot1 (a histone methyltransferase without a SET domain) and its substrate, H3K79: after DNA damage occurs, H3K79me contributes to the maintenance of the recruited 53BP1/Rad9/Crb2 to the dsDNA breaks (Huyen et al., 2004; Giannattasio et al., 2005). Why such a different function in S. pombe and metazoans?

In S. pombe there is only one SET domain protein, Set9, that has specific activity on H4K20 and it probably catalyzes mono, di, trimethylation whilst in mammalian cells and Drosophila another SET domain family, PRSET7/SET8 catalyzes mono and di-methylation of H4K20. These modifications have all different functions: H4K20me1 marks inactive X chromosome (Kohlmar et al., 2004) and H4K20me2 is present on euchromatin in mammalian cells (Rice et al., 2002; Nishioka et al., 2002, Fang et al., 2002).

However another modification involved in genome integrity is conserved in all organisms, from S. cerevisiae to S. pombe to mammalian cells: H2A phosphorylation. In contrast to H4K20me3 that is constitutively present and widely spread on the entire genome as shown in the immunofluorescence data and in the ChIP experiments, H2AS129p occurs immediately after double strand DNA damage. Both H4K20 and H2A are indispensable for a correct response to DNA damage and Crb2 recruitment to the DNA foci but they act independently: H2A becomes phosphorylated after IR treatment despite the presence of Set9 and H4K20me3. The role of H4K20 in genome integrity in fission yeast is probably to be attributed to an evolutionary early function that then developed in higher eukaryotes with more complex genomes to "lock" pericentromeric heterochromatin (Schotta et al., 2004). In S. cerevisiae, where H3K79me is important for the DNA damage response (Giannattasio et al., 2005), an ortholog of Set9 does not exist.

Though present in the N-terminal tail region of H4, lysine 20 is not exposed but buried in the context of stacked nucleosomes (Dorigo et al., 2003; White et al., 2001). Based on this and on the fact that H4K20 is present constitutively throughout all the genome, it might be that when a DSB occurs, a region of open chromatin exposes this pre-existing methylated H4K20 residue. Recognition of this modification together with H2AX phosphorylation by Crb2 leads to the recruitment of more Crb2 protein and activation of the checkpoint (Figure 7-16). A possible experiment to rule out this hypothesis would be to perform Chromatin Immunoprecipitation with α-H4K20me3 and α-Crb2 on an inducible DSB. Inducible DSB can be caused by site-specific
endonucleases, like HO endonuclease. The use of this technique is advantageous because the site of damage is known and the signal emanates from a single lesion. This same technique could also be used to check whether Crb2 binds directly H4K20: Crb2 has a noncanonical tudor fold, possibly involved in binding methyl-lysine. Its homolog in mammalian cells, 53BP has two tudor domains and it has been recently shown that these two domains are indispensable in recognition and binding of H3K79 after DSB formation (Huyen et al., 2004).

Set9-GFP protein and H4K20me3 are differently localized: Set9 seems to be distributed throughout the nucleoplasm whilst H4K20me3 is on the chromatin but not in the nucleolus.

The fact that Set9-GFP immunolocalization present overlapping signals with Nop1, a nucleolar marker, suggests the possible association of Set9 with the nucleolus. However further and more detailed analysis are required to assess the association of Set9 with chromatin and rDNA: ChIP with α-GFP in set9-GFP expressing cells, reveals a 1:1 ratio between rDNA and fbp1 (not shown), a euchromatic region: Set9 could be associated with both regions or simply distributed in all the nucleus and in this case the ChIP simply might not have worked.

To further investigate the function of Set9 in DNA damage checkpoint, its localization was examined after DNA damage induction by IR irradiation but no change in its localization was observed.

The discovery of the involvement of this SET domain protein and its substrate, H4K20 in maintaining DNA damage checkpoint, reinforces the idea of the existence of a histone code also in processes different from gene activation/repression. Again the importance of histone modifications in determining particular chromatin structures becomes the base for a histone code that may be read and recognized by proteins with particular domains that can bind these modified histones. Another challenge becomes the discovery of protein domains involved in recognition of modified histone on "damaged" DNA.
Figure 7-16 Diagram of H4K20me function in maintaining DNA damage checkpoint arrest.

Introduction of a DSB generates a region of open chromatin exposing a preexisting H4K20me that can be recognized by Crb2. Me_x represents the possible mono-, di- or trimethylated state of H4K20. Phosphorylation of H2A on Serine 129 is also required to recruit Crb2 but the mechanism is still unknown (Sanders et al., 2004). This model does not show all the proteins activated during the checkpoint activation step.
CONCLUSION AND PERSPECTIVES

NEW COMPONENTS OF THE RNAi PATHWAY

Rpb7, a component of RNA pol II, is involved in centromeric transcription.

Fission yeast is an excellent model system in which to study centromere structure. So many features are well conserved with respect to its structure and organization relative to metazoans centromeres. Furthermore as in plants and metazoans, fission yeast possesses an intact RNAi pathway which is required for the formation of silent chromatin at centromere and at mating type locus (Volpe et al., 2002; Hall et al., 2002). The advantage of fission yeast is that each of the genes encoding for the components of the RNAi pathway is present in single copy and these genes are not essential. This facilitates both genetic and biochemical analyses of the RNAi. Consequently, many recent studies have utilized S. pombe to dissect the mechanism by which RNAi pathway brings about chromatin modifications. It has been demonstrated that transcripts from centromeric repeats accumulate in the RNAi mutants and siRNAs complementary to centromeric sequences have been cloned (Volpe et al. 2002; Reinhart & Bartel, 2002). Moreover the RITS complex, that is formed by a chromo-domain protein Chp1, Ago1 and an unknown protein Tas3, has been purified. The RITS complex has been shown to contain centromeric siRNAs and represents the link for RNAi directing chromatin modification at the centromere (Verdel et al., 2003). Another complex called RDRC is formed by Rdp1, a helicase Hrr1 and Cid12 (a putative polyA polymerase) has been purified and demonstrated to localize at the centromeric repeats (Motamedi et al., 2004). Together these data have provided additional details about the components involved and the mechanism by which the RNAi pathway functions. It appears that once established, RNAi acts in cis in a closed loop to maintain heterochromatin over centromeric repeats (Sugiyama et al., 2005). However other details remain unknown: which RNA polymerase is responsible for producing these centromeric transcripts and how it is recruited in this "silent" chromatin enviroment. In addition it is still not known how Ctr4 is recruited by RNAi to methylate H3K9. In plants it has been found that some repeats (5S rDNA and AtSN1 SINE retrotransposons) are transcribed by a fourth RNA polymerase, Pol IV but homologs of this polymerase are absent in fission yeast and metazoans (Onodera et al., 2005; Herr et al., 2005).

In this thesis, during the course of this analysis, other components have been characterized that act upstream or within the RNAi-chromatin modification pathway. These components are three subunits of RNA pol II: Rpb2 (Kato et al., 2005), Rpb1 (Schramke et al., 2005 and unpublished data) and Rpb7 (Djupedal et al., 2005). Interestingly even though these are three subunits of RNA pol II, they have very different phenotypes and functions with respect to RNAi-chromatin modification pathway.

In rpb2-m203 cells, centromeric transcripts accumulate like in RNAi mutants and centromeric siRNAs are not produced. In addition, as in RNAi mutants, H3K9me2 is not present over a
marker gene inserted in the centromeric outer repeats and Swi6 association is lost. These phenotypes suggests that somehow the transcripts produced in a rpb2-m203 strain are “defective”, they cannot be diced by Dicer1 to produce siRNAs and thus chromatin modifications cannot occur and spread on a marker gene inserted in the centromeric repeats. Alternatively the loop to maintain heterochromatin is “broken” and loss of H3K9me2 derives from inefficiency of RNAi-chromatin modification pathway.

In rpb1-11 cells, that have a truncation of the CTD domain of RNA pol II, centromeric transcripts do not accumulate and siRNAs are produced at the same levels as in wild type cells. However this mutant presents a decrease of H3K9me2 and Swi6 localization over a marker gene inserted into the repeats. These phenotypes are suggestive of a problem that affects the RNAi pathway to direct chromatin modifications.

Finally the mutant rpb7-1, characterized in this thesis, previously found to affect silencing at the outer repeats of centromere 1 (Ekwall et al., 1999), is involved specifically in the production of centromeric transcripts: no accumulation of centromeric transcripts appears in this mutant and few centromeric siRNAs are made. All this data together with the fact that even in combination with dcr1Δ, relatively little centromeric transcripts accumulate, suggest that Rpb7 acts upstream Dicer and in fact is involved in the production of Dicer substrates, the centromeric transcript. The heterochromatic features, such as H3K9me2 and Swi6 localization, are lost on a gene inserted in the outer centromeric repeats. However, on the centromeric repeats themselves, H3K9me2 and Swi6 localization are retained and this could suggest that the few centromeric siRNAs produced in rpb7-1 cells, are enough to initiate but not spread H3K9me2 and Swi6 heterochromatin formation. Another possible explanation could be that some H3K9 methylation occurs on the centromeric repeats and thus C1r4 can be recruited at the centromeric repeats in a RNAi-independent mechanism, like at the mating type loci via a mechanism similar to the stress-activated ATF/CREB transcription factors (Jia et al., 2004).

Since rpb2-m203, rpb1-11 and rpb7-1 are all mutations in subunits of RNA pol II, this suggests that RNA pol II is transcribing the centromeric outer repeats in fission yeast. In fact, ChIP analysis performed with a specific antibody directed against the CTD domain of RNA pol II, confirmed that RNA pol II is associated with the centromeric repeats (this thesis Chapter 4 and Djupedal et al., 2005; Cam et al., 2005).

Recruitment of RNA pol II to heterochromatic domains is paradoxical and suggests the need for a very refined mechanism to control the expression of centromeric transcripts and at the same time the establishment of silent chromatin on repeat elements (centromeric or mating type repeats). The RNAi-directed chromatin modification machinery probably represents such a mechanism since it must simultaneously interacts with both heterochromatin and transcription machinery. It generates siRNAs both for assembly of higher-order chromatin structures and to silencing repeat elements, destroying their transcripts in a feed back loop. Physical interactions between the RNAi pathway (Ago1) and the transcription apparatus (RNA pol II) have been detected but it is not known which components of RITS contacts RNA pol II subunit and whether
rpb1-11 and rpb2-m203 mutations affect these interactions. (Schramke et al., 2005 and unpublished results).

What remains to be revealed is how the RNAi pathway and the heterochromatin initiation-machinery such the key H3K9 methyltransferase Clr4, “communicate” to establish silent chromatin formation. From recent work it seems that Rik1 and some novel proteins (Dos1, Dos2) that interact with it (Li et al., 2005; Horn et al., 2005) play an important role in the connection between RNAi pathway and histone modification. One possibility is that the interaction between RNAi components and RNA pol II alters the status of RNA pol II somehow recruiting Clr4/Rik1 and establishing the initiating H3K9me2 modification event which subsequently spreads with the histone deacetylase Sir2 along the fibre.

However, as previously mentioned, there is still the possibility of the existence of an RNAi independent mechanism, like the ATF/CREB stress activated factors that nucleate silent chromatin at the mating type loci (Jia et al., 2004) to recruit Clr4 and establish silent chromatin at nucleation loci also at the centromeres. This is suggested by the fact that not all H3K9me2 is lost in RNAi mutants at the centromeric repeats (Sadaie et al., 2004).

Moreover recent data from our lab indicate that H3K9R mutants can produce a low level of centromeric siRNAs despite the fact that siRNAs are not detected in cells harbouring a catalitically dead allele of the H3K9 methyltransferase Clr4 (S. White and R. Allshire). This may suggests that there are additional Clr4 substrates whose modification is required to retain these siRNAs. These siRNAs may represent putative primary siRNAs formed from the centromeric transcripts prior to the establishment of heterochromatin modifications. Further investigation of this will be necessary to build a more detailed model of how RNAi brings about H3K9 methylation.

csp mutants: a link between general RNA metabolism and RNAi pathway.

Among the csppts mutants, that alleviate centromeric silencing, two splicing factors, Cwf10/Csp4 and Prp39/Csp5 were identified (Ekwall et al., 1999). These mutants represent factors involved in general RNA metabolism and probably in the RNAi pathway: the major defect they present relative to silent chromatin formation, is that centromeric siRNAs production is greatly reduced compared to wild type cells by northern analysis presented in Chapter 3. In addition, many splicing factors (29) associated in pull down assays with Cid12, a component of the RDRC complex, particularly in cells lacking Rdpl (Motamedi et al., 2005). This together with the fact that some splicing mutants are defective in silencing (Ekwall et al., 1999), indicates a possible involvement of some splicing factors in complexes or mechanisms that are common also to the RNAi pathway.

In fact many different proteins have been identified as essential for RNAi or as a components of the RISC complex but only a few have been functionally characterized at the molecular level. Dicer and Argonaute proteins have received the most attention. Among other relevant RNAi factors are small dsRBD-protein partners of Dicer proteins, such as R2D2 and Loqs of
Drosophila, and RDE-4 of C. elegans (Forstemann et al., 2005; Saito et al., 2005), proteins with RNA helicase/ATPase domains, and a Tudor staphylococcal nuclease, Tudor-SN. Tudor-SN, a component of RISC was recently shown to bind and possibly degrade dsRNAs hyperedited by adenosine deaminases (ADARs), pointing to a connection between editing and RNAi pathways (Scadden et al., 2005). Another protein, like Vigilin appears to be involved in the formation of heterochromatin and it is found in complexes with RNA helicase A activity and Ku86/70 but also with ADAR1, editing activity (Wang et al., 2005).

To discover more about the mechanism of the RNAi pathway, and the possible role of the csp ts mutants in the RNAi pathway and in the formation of silent chromatin, some biochemical analysis will be required. This will facilitate the detection of interactions that occur between RNAi pathway proteins and factors such as splicing components involved in general RNA metabolism. Eventually it will be interesting to find which step of the RNAi pathway they affect. It is not known whether Cid12 (a component of the RDRC complex; Motamedi et al., 2005) has poly(A)polymerase activity but if it has, it could be involved in the polyadenylation of the centromeric transcripts (I. Djupedal/Ekwall lab personal communication and S. White) or it might act to recruit RNA degradation enzymes such as the exosome (La cava et al., 2005; Vanacova et al., 2005).

Other RNA processing and RNA pol II interacting factors, polyadenylation factors, capping and splicing factors could have a role during the centromeric transcripts processing and RNAi directed chromatin modifications. In S. cerevisiae capping, splicing and polyadenylation appear to occur co-transcriptionally. The csp 1, csp2 or csp6 have not been cloned. All these csp mutations are in essential genes which could encode other factors involved in RNA metabolism such as capping, splicing factors or polyadenylation factors. However, considering that csp 1 and csp2 cells do not have such a strong defective silencing phenotype in terms of defects in the RNAi pathway, H3K9 methylation and Swi6 localization, it may be that they have indirect effects on the pathway.

In contrast to all of the other cspts mutants, csp6 loses H3K9me2 completely Swi6 delocalized and no centromeric siRNAs produced. This suggests that it is required somehow for Clr4 recruitment thus being at the interface between the RNAi pathway and chromatin modification. From recent data, it is known that both Clr4 and Swi6 are sumoylated and this modification is important for silent chromatin formation at centromeres (Shin et al., 2005). Csp6 might be important in a similar pathway, to guarantee the stability of Clr4.

CHROMATIN STRUCTURE AND DNA REPAIR

SET domain protein and the role of Set9 in DNA damage checkpoint.

Histone modifications determine in part chromatin structure. Different methylation or acetylation states on the same residue can induce binding of different proteins and thus mediate transcription or repression of specific genomic regions (Sthral & Allis, 2000). A family of enzymes that specifically catalyze histone methylation are the SET domain proteins. Usually
these enzymes exhibit a common motif (130 amino acids) that is necessary but not sufficient for their activity.

Fission yeast has only 13 SET domain proteins whereas in human there are over 70. In this regard it could be that, each SET protein have a specific target while in mammalian cells there is a higher degree of redundancy and it becomes difficult to find which residue is methylated by which SET domain protein on histones. For example histone H3K9 methylation is catalyzed by Suv39 and by G9a in mouse cells and this residue can be mono-, di and trimethylated. The deletion of genes encoding three different SET domain proteins that are present in S. pombe but absent in S. cerevisiae (Chapter 7) was performed with the intention of determining whether a possible H3K9 monomethyltransferase exists in fission yeast. No H3K9 methylation has in fact been found in S. cerevisiae. H3K9me1 appears to persist at pericentromeric heterochromatin in mouse cell line lacking both Suv39h1 and Suv39h2 (Thomas Jenuwein personal communication). It is likely that the SET domain protein catalyzing H3K9 methylation in S. pombe, Clr4, is responsible for the dimethylation but it is still unclear whether it catalizes also monomethylation and also trimethylation. From the data collected to date there is no indication of a monomethyltransferase in fission yeast. Deletion of none of the set3, set7, set9 and set10 genes affect silencing at any of known heterochromatic loci: centromeres, telomeres and mating type. If one of these set genes encoded a monomethyltransferase for H3K9 and prepared the substrate for dimethylation by Clr4, a defect in silent chromatin formation would have been expected.

Strains harbouring combinations of set3Δ, set7Δ, set9Δ and set10Δ mutants have not been constructed: they could enhance defects that are not detectable in single mutant. There are other phenotypes that have not been analysed and should be tested on set3Δ, set7Δ and set10Δ cells, like sensitivity to DNA damaging agents. In fact the only phenotype observed within set9Δ cells was the sensitivity to DNA damage agents. This together with other studies concerning H3K79 and H3K56 histone modifications involved in checkpoint activation, have strengthened the idea of a role for histone modifications in genome integrity and the existence of a histone code for DNA damage (Sanders et al., 2004; Hayen et al., 2004; Giannattasio et al., 2005; Masumoto, 2005).

In a collaborative study with Sanders in Kouzarides laboratory (CRUK/Cambridge), the deletion of Set9 has been demonstrated to cause the disappearance of H4K20me1, H4K20me2 and H4K20me3 modification. This, together with the creation of H4K20R mutants, allowed us to show the existence of H4K20 methylation in S. pombe as in mammalian cells and Drosophila (Schotta et al., 2004). However, although H4K20 is methylated in S. pombe, it has a very different function since neither Set9 HMTase or the H4K20 residue play a role in the formation of heterochromatin. Instead this methylation event is important to maintain intact the DNA damage checkpoint response. In set9Δ cells, the checkpoint protein Crb2 does not localize to ds DNA breaks foci anymore. Previously another modification, the phosphorylation on H2AS129
was shown to be important for the localization of Crb2 (Nakamura et al., 2004) but this modification transmits independently from H4K20 methylation the signal to Crb2 and it is unaffected in set9Δ cells and H4K20R mutants. Further analysis are required to try to identify which protein binds to H4K20me3 in fission yeast when DNA damage occurs. It would be interesting to find a novel motif that recognizes chromatin structures altered by DNA damage.
ACKNOWLEDGEMENTS

I would like to thank Robin for discussing my results, assisting with the writing of my thesis and giving me the opportunity to undertake my PhD in his lab.
I am really grateful to Alison for also helping me with the writing, discussing my results and her continual advice on how to work with fission yeast.
Thanks to all the lab members for their support, in particular to Ara, Sarah, Sharon and Ricardo for showing me how to use End-note and the alignment programs. I would also like to thank Barbara Mellone for helping me and showing me how to work with S. pombe when I arrived in the lab.
It was very good to collaborate with Steve Sanders (CRUK/Cambridge) whom I want to thank for his advice on the Set9 project.
Many thanks also to Karl Ekwall whom I worked with on Rpb7.
LITERATURE CITED


centromeric region in Beta species. Genome Res 11, 253-265.


histone H3 lysine 9 creates a binding site for HP1 proteins. Nature 410, 116-120.


Centromere silencing and function in fission yeast is governed by the amino terminus of histone H3. Curr Biol 13, 1748-1757.


130
Dev 9, 158-163.


465-476.


nondisjunction and cell cycle arrest at mitosis. Genes Dev 9, 573-586.


van Nues, R. W., and Beggs, J. D. (2001). Functional contacts with a range of splicing proteins suggest a central role for Brr2p in the dynamic control of the order of events in spliceosomes of Saccharomyces cerevisiae. Genetics 157, 1451-1467.


siRNA accumulation and DNA and histone methylation. Science 299, 716-719.