This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
Role of DNA sequence in CENP-A^{Cnp1} assembly
at fission yeast centromeres

Sandra Catania

Thesis presented for the
Degree of Doctor of Philosophy
University of Edinburgh
December 2012
PREFACE

This thesis was composed by myself and the research presented is my own unless otherwise stated.

Sandra Catania

December 2012
Table of Contents

ABSTRACT V
Abbreviations VII
LIST OF FIGURES X

CHAPTER 1 1
INTRODUCTION 1
1.1 CENTROMERE STRUCTURE 1
   1.1.1 Point centromeres: S. cerevisiae centromeres 3
   1.1.2 Holocentric centromeres 5
   1.1.3 Drosophila melanogaster centromeres 5
   1.1.4 Mammalian centromeres 6
1.2 Schizosaccharomyces pombe centromeres 8
   1.2.1 Centromere structure 8
   1.2.2 Sequence requirements for centromere function 10
   1.2.3 Chromatin domain at fission yeast centromeres 10
   1.2.4 Centromeric heterochromatin domain 11
   1.2.5 Heterochromatin and centromere function 14
   1.2.6 Central domain and kinetochore structure 15
1.3 EPIGENETIC REGULATION OF CENTROMERES 21
   1.3.1 Centromeric DNA is not sufficient for centromere function 21
   1.3.2 Centromeric DNA is not required for centromere function 22
1.4 CENP-A NUCLEOSOME STRUCTURE 23
1.5 CENP-A ASSEMBLY 26
   1.5.1 Timing of CENP-A assembly 26
   1.5.2 Factors affecting CENP-A deposition 28
   1.5.3 Chromatin environment and CENP-A deposition 30
   1.5.4 Sequence specificity and centromere function 31
1.6 TRANSCRIPTION AND CENP-A DEPOSITION 33
   1.6.1 Transcription at fission yeast centromeres 34
   1.6.2 Is there a link between transcription and CENP-A localisation? 35
1.7 Aims of the project 37

CHAPTER 2 39
MATERIALS AND METHODS 39
2.1 Fission Yeast Protocols 39
   2.1.1 Mating and random spore analysis 39
   2.1.2 S. pombe transformations through electroporation 39
   2.1.3 Spotting assay 39
   2.1.4 Centromere silencing assay 40
   2.1.5 Minichromosome transformation 40
   2.1.6 Establishment assay 40
   2.1.7 Minichromosome loss assay 41
   2.1.8 Media and Growth 41
2.2 DNA protocols 43
   2.2.1 Genomic DNA isolation from fission yeast 43
   2.2.2 Yeast genomic DNA isolation for colony-PCR 44
2.2.3 Agarose gel electrophoresis 44
2.2.4 Polymerase Chain Reaction (PCR) 44
2.2.5 Fusion PCR 45
2.2.6 Minichromosome construction 45
2.2.7 Sequencing 46
2.2.8 Southern analysis 46

2.3 RNA protocols 47
2.3.1 Total RNA isolation 47
2.3.2 Poly(A) RNA isolation 48
2.3.3 Northern Blotting 48
2.3.4 RT-PCR 49

2.4 Protein Protocols 49
2.4.1 S. pombe protein extraction 49
2.4.2 Western analysis 50
2.4.3 Chromatin Immunoprecipitation (ChIP) 50

2.5 ENZYMATIC ASSAY 52
2.5.1 Spotting assay for β-galactosidase 52
2.5.2 Liquid assay for β-galactosidase 52

2.6 Bacterial Protocols 53
2.6.1 Competent cell 53
2.6.2 Transformation 54
2.6.3 Plasmid miniprep 54
2.6.4 Bacteria Media 55

2.7 Antibodies 55
2.7.1 ChIP 55
2.7.2 Western Analysis 55

2.8 Primers used in this thesis 56
2.9 Strains used in this thesis 63

CHAPTER 3
INVESTIGATION OF THE MINIMAL SEQUENCE REQUIREMENT FOR CENP-A_Cnp1 CHROMATIN ASSEMBLY 65
3.1 INTRODUCTION 65
3.2 RESULTS 66
3.2.1 The central core 2 sequence shows functional redundancy 67
3.2.2 A 2 kb- repeated sequence from cc2 is sufficient to assemble CENP-A_Cnp1 69
3.2.3 The 3xLM sequence is able to assemble a functional centromere in wild type cells 71
3.2.4 A second 2 kb-repeat sequence from cc2 is able to confer centromeric function when inserted into a plasmid. 75
3.3 DISCUSSION 75

CHAPTER 4
GENERATING FISSION YEAST WITH NO CENTRAL CORE DOMAIN AT CENTROMERE 2 TO ALLOW ANALYSES OF SEQUENCE REQUIREMENTS FOR CENP-A ESTABLISHMENT 78
4.1 INTRODUCTION 78
4.2 RESULTS 79
4.2.1 Removing the central core 2 from endogenous cen2 79
4.2.2 Step1: Insertion of the central core sequence from cen1 into cen2 82
4.2.3 Step2 & 3: Removal of central core 2 sequences from cen2 82
4.2.4 Replacement of cc2 with cc1 DNA does not affect cell viability 84
4.2.5 Assessing the ability to establish a functional centromere on a plasmid 84
4.2.6 Assessing the performance of cc2 fragments in CENP-A<sup>Cnp1</sup> establishment assays in cells lacking most of cc2 at cen2 87

4.3 DISCUSSION 92

CHAPTER 5
A ROLE FOR NUCLEOTIDE COMPOSITION IN SPECIFYING REGIONAL CENTROMERE 95
5.1 INTRODUCTION 95
5.2 RESULTS 96
5.2.1 A random A/T rich sequence cannot establish CENP-A<sup>Cnp1</sup> in a H3K9R mutant. 96
5.2.2 A synthetic AT-rich sequence with the same predicted nucleosome as endogenous centromeric DNA cannot establish CENP-A<sup>Cnp1</sup> chromatin. 98
5.2.3 Native centromeric DNA sequence can drive CENP-A<sup>Cnp1</sup> deposition on adjacent synthetic DNA sequence. 101
5.3 DISCUSSION 101

CHAPTER 6
INVESTIGATING CANDIDATE DNA BINDING PROTEINS FOR A ROLE IN DEFINING CENTROMERE DNA . 105
6.1 INTRODUCTION 105
6.2 RESULTS 106
6.2.1 Analysis of in vivo binding of candidate proteins with potential nucleotide binding motifs 106
6.2.2 Determining if the BAH domain of the origin recognition complex 1 protein plays a role in defining centromeres 109
6.3 DISCUSSION 116

CHAPTER 7
ROLE OF TRANSCRIPTION IN CENTROMERE FUNCTION 120
7.1 INTRODUCTION 120
7.2 RESULTS 121
7.2.1 Analysis of transcripts from the central domain of centromere 2 121
7.2.2 The detection of transcripts using plasmid borne central domain DNA 122
7.2.3 Analysis of LM transcripts by RT-PCR 127
7.2.4 Characterisation of transcriptional start sites and upstream promoter regions from cc2 127
7.2.5 Investigating the role of transcription in CENP-A<sup>Cnp1</sup> function 135
7.3 DISCUSSION 141

CHAPTER 8
DISCUSSION 143
8.1 Establishment of CENP-A<sup>Cnp1</sup> chromatin in fission yeast 143
8.2 MODEL 1: Sequence-specific binding factors promote CENP-A establishment 145
8.3 MODEL 2: Chromatin remodelling promotes CENP-A establishment 148
8.3.1 Transcription as a chromatin remodeller 150
8.3.2 Other chromatin remodelling processes involved in centromere function 152
8.4  MODEL 3: Imposition of a precise level of transcription promotes CENP-A establishment 153
8.5  Further investigation of the role of transcription in CENP-A establishment 156
8.6  Concluding remarks 158
Acknowledgements 159
APPENDIX 1 160
APPENDIX 2 161
APPENDIX 3 162
References 163
ABSTRACT

The centromere is the site of kinetochore assembly that ensures proper chromosome segregation. Active centromeres are formed at chromosomal locations that do not appear to share homology between different species; this and other analyses has lead to the conclusion that centromeres are epigenetically determined. In all organisms, centromere location is specified by the assembly of unusual nucleosomes containing the histone H3 variant CENP-A in place of H3. However, an apparent paradox is that CENP-A in most organisms generally occurs on certain preferred sequences. The analyses presented focuses on the influence of DNA sequence on the selection of the locus where CENP-A chromatin are formed and whether there are any particular DNA features that promote CENP-A\textsuperscript{Cnp1} assembly at centromeres in \textit{Schizosaccharomyces pombe}. \textit{S. pombe} provides an excellent model to study the structure and function of a complex eukaryotic centromere as it possesses epigenetically regulated centromeres that are structurally related to those of metazoa. Furthermore, plasmid-based circular minichromosomes provide a useful tool for studying these centromeres and the inactivation of minichromosome centromeres does not affect cell viability. The main aim of this study was to identify those features that enable centromeric sequences to assemble CENP-A chromatin. Manipulated circular minichromosomes were utilised to investigate the minimal central core sequence requirement for establishment of CENP-A\textsuperscript{Cnp1} chromatin. These analyses showed that a minimal 2kb region from the central core of \textit{cen2} could form a functional centromere. A second aim was to analyse the DNA sequence requirements for centromere function on this minimal 2 kb region. To facilitate this, the endogenous central core region of endogenous \textit{cen2} was replaced with the central core region of \textit{cen1}. This modified strain allows the structural and biological properties associated with plasmid borne central core 2 sequences to be analysed. Transcription of central domain sequences has been proposed to play a role in CENP-A establishment and/or maintenance. To explore the contribution of transcription potential promoters were mapped within the minimal 2 kb sequence and their regulatory elements investigated. Mutation of the minimal DNA element impedes its ability to assemble CENP-A chromatin. Therefore the primary DNA sequence of fission yeast centromeres is important for establishing functional centromeres and thus centromere location not entirely epigenetically regulated. It remains to be
determined if the characteristics associated with these sequences, and their mode of action, are conserved at other centromeres.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-FOA</td>
<td>5-fluoro-orotic-acid</td>
</tr>
<tr>
<td>ade</td>
<td>adenine</td>
</tr>
<tr>
<td>adh</td>
<td>alcohol dehydrogenase</td>
</tr>
<tr>
<td>AHT</td>
<td>anhydrotetracycline</td>
</tr>
<tr>
<td>arg</td>
<td>arginine</td>
</tr>
<tr>
<td>ARS</td>
<td>autonomously replicating sequence</td>
</tr>
<tr>
<td>BAH</td>
<td>bromo-adjacent homology</td>
</tr>
<tr>
<td>bp</td>
<td>base pair</td>
</tr>
<tr>
<td>C-terminal</td>
<td>carboxy terminal</td>
</tr>
<tr>
<td>cc</td>
<td>central core</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>CENP</td>
<td>centromere protein</td>
</tr>
<tr>
<td>CFP</td>
<td>Cyan fluorescent protein</td>
</tr>
<tr>
<td>ChIP</td>
<td>chromatin Immunoprecipitation</td>
</tr>
<tr>
<td>CTD</td>
<td>carboxy-terminal domain</td>
</tr>
<tr>
<td>dH₂O</td>
<td>distilled water</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxy-nucleotide triphosphate</td>
</tr>
<tr>
<td>dsRNA</td>
<td>double stranded ribonucleic acid</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene di-amine tetra acetic acid</td>
</tr>
<tr>
<td>FACT</td>
<td>facilitates chromatin transcription</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>H3K36me</td>
<td>histone H3 methylated on lysine 36</td>
</tr>
<tr>
<td>H3K9me2</td>
<td>histone H3 dimethylated on lysine 9</td>
</tr>
<tr>
<td>H3K9R</td>
<td>lysine 9 of histone H3 mutated to arginine</td>
</tr>
<tr>
<td>H4K4me2</td>
<td>histone H3 dimethylated on lysine 4</td>
</tr>
<tr>
<td>HA</td>
<td>haemagglutinin</td>
</tr>
<tr>
<td>HDAC</td>
<td>histone deacetylase</td>
</tr>
<tr>
<td>imr</td>
<td>inner most repeats</td>
</tr>
<tr>
<td>Kb</td>
<td>kilobase</td>
</tr>
<tr>
<td>KDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>LacZ gene</td>
<td>β-galactosidase gene</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertani medium</td>
</tr>
<tr>
<td>LINE</td>
<td>long interspersed nuclear elements</td>
</tr>
<tr>
<td>ME</td>
<td>malt extract</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MNase</td>
<td>micrococcal nuclease</td>
</tr>
<tr>
<td>mRNA</td>
<td>messanger RNA</td>
</tr>
<tr>
<td>N-terminal</td>
<td>amino-terminal</td>
</tr>
<tr>
<td>NFR</td>
<td>nucleosome free region</td>
</tr>
<tr>
<td>nmt</td>
<td>no message in thiamine</td>
</tr>
<tr>
<td>nt</td>
<td>nucleotide</td>
</tr>
<tr>
<td>ONPG</td>
<td>Ortho-nitrophenyl-β-D-galactopyranoside</td>
</tr>
<tr>
<td>ORC</td>
<td>origin of replication complex</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>otr</td>
<td>outer repeats</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered 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>qPCR</td>
<td>quantitative polymerase chain reaction</td>
</tr>
<tr>
<td>RACE</td>
<td>Rapid amplification of cDNA end</td>
</tr>
<tr>
<td>RDRC</td>
<td>RNA-directed RNA polymerase complex</td>
</tr>
<tr>
<td>RITS</td>
<td>RNA-mediated 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>RNAPII</td>
<td>RNA polymerase II</td>
</tr>
<tr>
<td>rpm</td>
<td>rotation per minute</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>reverse transcription polymerase chain reaction</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulphate</td>
</tr>
<tr>
<td>sim</td>
<td>silencing in the middle</td>
</tr>
<tr>
<td>sir</td>
<td>silent information regulator</td>
</tr>
<tr>
<td>SSC</td>
<td>saline-sodium citrate</td>
</tr>
<tr>
<td>SynAT</td>
<td>synthetic AT-rich</td>
</tr>
<tr>
<td>SynRLM</td>
<td>synthetic randomised LM</td>
</tr>
<tr>
<td>TBE</td>
<td>tris-borate EDTA</td>
</tr>
<tr>
<td>TBZ</td>
<td>thiobendazole</td>
</tr>
<tr>
<td>tcs</td>
<td>TEV recognition site</td>
</tr>
<tr>
<td>TetO</td>
<td>tetracycline operator</td>
</tr>
<tr>
<td>TetR</td>
<td>tetracycline repressor protein</td>
</tr>
<tr>
<td>TEV</td>
<td>Tobacco Etch Virus</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>ts</td>
<td>temperature sensitive</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>TSS</td>
<td>transcription start site</td>
</tr>
<tr>
<td>Tween</td>
<td>polyoxyethylenesorbitan monolaurate</td>
</tr>
<tr>
<td>ura</td>
<td>uracile</td>
</tr>
<tr>
<td>wt</td>
<td>wild type</td>
</tr>
<tr>
<td>X-Gal</td>
<td>5-bromo-4-chloro-indolyl-β-D- galactopyranoside</td>
</tr>
<tr>
<td>YES</td>
<td>yeast extract supplemented</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1-1.
Schematic representation of centromeric DNA and CENP-A domain organisation

Figure 1-2.
Schematic representation of budding yeast kinetochore

Figure 1-3.
DNA structure of fission yeast centromeres

Figure 1-4.
Centromeric central core chromatin structure of centromere-inactive and centromere-active minichromosomes in S. pombe

Figure 1-5.
Model for heterochromatin formation in S. pombe

Figure 1-6.
Overview of the molecular organisation of vertebrates and fungi kinetochores.

Figure 1-7.
Structure of vertebrate kinetochore

Figure 1-8.
CENP-A structure and models for CENP-A containing nucleosomes.

Figure 1-9.
5’-RACE/PCR and Northern analyses of S. pombe cnt1 transcripts

Figure 3-1.
Central core 2 shows functional redundancy

Figure 3-2.
A 2 kb repeated sequence from cc2 is sufficient to assemble CENP-A^{Cnp1}.

Figure 3-3.
Schematic representation of the establishment assay.

Figure 3-4.
pH-3xLM is able to establish a functional centromere in wild type.

Figure 3-5.
pH-3xOP is able to establish a functional centromere in wild type.

Figure 4-1.
Schematic representation of replacement of cc2 with cc1 by traditional homologous recombination
Figure 4-2.
Schematic representation of replacement of cc2 by insertion of cc1 sequence followed by deletion of cc2.

Figure 4-3.
Southern analysis of strain created in STEP1

Figure 4-4.
Southern analysis of the deletion of central core 2 sequence

Figure 4-5.
Deletion of centromere 2 is not affecting centromere structure.

Figure 4-6.
Central core 2-deleted strain can establish a functional kinetochore on a minichromosome

Figure 4-7.
Analysis of full-length cc2 on a plasmid in H3K9R-cc2Δ cells.

Figure 4-8.
Analysis of the level of CENP-A^{Cnp1} on plasmids containing fragments of cc2 sequence

Figure 4-9.
Levels of H3 on plasmids containing fragments of cc2 sequence.

Figure 5-1.
SynAT can not establish CENP-A^{Cnp1} in H3K9R cells

Figure 5-2.
Predicted nucleosome positions on centromeric DNA

Figure 5-3.
Predicted nucleosome positions on centromeric DNA and SynAT is distinct.

Figure 5-4.
SynRLM is not able to recruit CENP-A^{Cnp1}

Figure 5-5.
A combination of endogenous and synthetic sequence can assemble CENP-A^{Cnp1}.

Figure 6-1.
ChIP to test in vivo binding of centromeric proteins to central domain DNA.

Figure 6-2.
Function of Sir proteins in S. cerevisiae

Figure 6-3.
Orc1 depleted of its BAH domain is able to rescue the orp1-4 ts allele
**Figure 6-4.**
Effects of BAHΔ on centromere structure

**Figure 6-5.**
Deletion of BAH domain does not affect centromere establishment on a minichromosome

**Figure 7-1.**
Northern of transcripts originating from endogenous centromeres.

**Figure 7-2.**
Northern of transcripts originating from a cc1-containing plasmid.

**Figure 7-3.**
Northern of transcripts originating from pcc2 and p-3XLM.

**Figure 7-4.**
Detection of transcripts originating from 3xLM

**Figure 7-5.**
Schematic representation of TSS identified within 3xLM.

**Figure 7-6.**
LacZ assay on *S. pombe* promoters.

**Figure 7-7.**
LacZ assay using cc2 sequences as promoter.

**Figure 7-8.**
Dissection of sequence required for the promoter function of p200b

**Figure 7-9.**
Targeted cleavage of Rpb1 CTD using TetR-TEV.

**Figure 7-10.**
Targeted cleavage of Rpb1 CTD to endogenous centromeres

**Figure 8-1.**
Heterochromatin and central domain DNA are required for establishment of CENP-A chromatin.

**Figure 8-2.**
MODEL 1- Sequence-specific binding factors promote CENP-A establishment

**Figure 8-3.**
Schematic representation of plasmids used for plasmid-pulldown experiment.

**Figure 8-4.**
MODEL 2: Transcriptional remodelling as a trigger for CENP-A establishment
Figure 8-5.
MODEL 3: Regulation of transcription is required for CENP-A deposition.

Figure 8-6.
Inhibition of transcription on plasmid-based minichromosomes. A)

Appendix 1.
Quantification of the CENP-A levels at the endogenous centromeres of wild type (wt*) or mutated H3K9R cells transformed with plasmids containing different regions of cc2.

Appendix 2.
Quantification of the H3 levels at the endogenous centromeres of wild type (wt*) or mutated H3K9R transformed with plasmids containing different regions of cc2 (A-E).
Quantification for CENP-A enrichment in wild type (grey)(n=3)

Appendix 3.
Agarose gels of 5'-RACE PCR and sequence of TSS identified.
INTRODUCTION

Cell division is a critical biological process that ensures the transmission of the genetic information to the daughter cells. Defects in this process can result in aneuploidy, cell death and can drive tumour formation. To faithfully segregate, each sister chromatid must be connected to the spindle microtubules. This interaction is mediated by a specific multi-protein complex called the kinetochore that is assembled on centromeric DNA. As centromeres and kinetochores are responsible for chromosome segregation, their regulation is essential for cell survival. It is therefore important that each chromosome possessess one and only functional kinetochore. Thus, the kinetochore must be assembled on one single site on the chromosomes since the presence of multiple kinetochores may result in formation of multicentric chromosomes leading to chromosomes breakage and unequal division.

The centromere is cytologically identified as the primary constriction of the chromosome and is the locus at which the kinetochore is assembled. In most organisms, centromeres are specified by the assembly of unusual nucleosomes that contain the histone H3 variant CENP-A in place of H3. As CENP-A specifies the region where kinetochore is assembled, regulation of its incorporation is crucial for chromosome segregation.

1.1 CENTROMERE STRUCTURE

The inheritance of a stable centromere location on chromosomes suggests that a specific DNA sequence may be required to direct centromere assembly to specific regions. Therefore, initially many studies focused on the identification of DNA regions that could specify centromere identity and function. However, only the simplest point centromeres in budding yeast are specified by the specific DNA sequence while in other organisms epigenetic mechanisms are involved in the selection of the location for kinetochore formation.

Three main types of centromeres have been described so far: point centromeres, holocentric and monocentric regional centromeres (Figure 1-1). Why an element that is essential for cell viability has evolved into so many different types of structure
Figure 1-1. Schematic representation of centromeric DNA and CENP-A domain organisation (from Allshire and Karpen, 2008). A) The simplest “point” centromere of budding yeast is dependent on 125 bp sequence which is composed by the three elements CDEI- II- III. The unique CENP-A\textsuperscript{Cae4}\textsubscript{C} containing nucleosome is assembled on CDEII. B) C. elegans does not possess a specific sequence. Centromeres assemble and CENP-A\textsuperscript{HCP-3} is deposited along the entire length of chromosomes. C) D. melanogaster centromeres are composed of interspersed blocks of CENP-A\textsuperscript{CID} and histone H3. The centromere from the Dp1187 minichromosome consists of AATAT and CTCTT satellites repeats. D) Human centromeres consist of modified H3 nucleosomes linearly interspersed with CENP-A nucleosomes. Centromeric DNA is composed of a higher order array of 171 bp α-satellite DNA, arranged in a tandem and frequently interrupted by retrotransposon elements. Some of the α-satellite monomers contain a CENP-B box that is recognised by CENP-B. E) Fission yeast centromeres are composed by two outer repeat regions (otr) that surround a central domain (imr and core). In S. pombe centromeres CENP-A\textsuperscript{Cnp1} is deposited in the central domain.
whilst maintaining a similar overall mechanism of function is an intriguing question for chromosome biology.

1.1.1 Point centromeres: *S. cerevisiae* centromeres

Point centromeres are the only centromeres whose function is strictly dependent on the underlying DNA sequence. The minimal functional centromere of the budding yeast *S. cerevisiae* consists of a conserved 125 bp DNA sequence organized in three Centromere DNA Elements (CDEs): CDEI, CDEII and CDEIII (reviewed in Cheeseman et al., 2002; Fitzgerald-Hayes et al., 1982; Hegemann and Fleig, 1993; Westermann et al., 2007). The central element CDEII is an 80 bp AT-rich sequence and represents the site where the unique CENP-A<sup>Cse4</sup>-containing nucleosome is deposited. CDEII is flanked by two palindromic sequences CDEI (8 bp) and CDEIII (25 bp). CDEIII is important for centromere function since it serves as a binding site for the CBF3 complex which is required for proper chromosome segregation and centromere function (Ndc10p, Cep3p, Ctf13p, and Skp1p) (Goh and Kilmartin, 1993; Lechner and Carbon, 1991; Sorger et al., 1994). The binding of CBF3 is sequence-specific and indeed mutations of the DNA sequence in CDEIII abolish CBF3 binding and centromere function *in vivo* (Lechner and Carbon, 1991). Unlike elements II and III, CDEI is not absolutely required for centromere function during mitosis, but rather confers a small enhancing effect on the efficiency of chromosome segregation (Cumberledge et al 1987, Sears et al 1995). CDEI is bound by the centromere binding factor 1 (Cbf1), a basic helix-loop-helix leucine zipper (bHLHzip) protein that in addition to CDEI it is recruited to several promoters where it acts as a transcription factor (Mellor et al., 1990). Although Cbf1p is not essential for kinetochore function, it induces bending of the CDEI motif allowing a direct physical interaction between Cbf1 and the CBF3 complex and therefore contributes to higher order kinetochore structure (Kent et al., 1994; Niedenthal et al., 1993). In addition, Cbf1 is required for transcription at budding yeast centromeres (Ohkuni and Kitagawa, 2011). In fact, chromosome instability of *cbf1Δ* cells is suppressed by transcription driven from an artificial promoter suggesting that although not essential, transcription may be involved in ensuring proper centromere function (Ohkuni and Kitagawa, 2011). The Ctf19 complex (Ctf19p, Mcm21p, and Okp1p) is found at the interface between DNA and the outer kinetochore and links the inner kinetochore complex Ctf3 (Ctf3p, Mcm16p, and Mcm22p) to the Ndc80 complex (Ndc80p, Spc24p, Spc25p, and Nuf2p (Wigge and Kilmartin, 2001). The interaction between the kinetochore complexes and the single microtubule is
Figure 1-2. Schematic representation of budding yeast kinetochore (from Cheeseman et al., 2002). Budding yeast centromere is genetically determined and is specified by a conserved 125 bp DNA sequence organized in three Centromere DNA Elements (CDEs): CDEI, CDEII and CDEIII. Each element is bound by specific kinetochore proteins. The central element CDEII represents the site where the unique CENP-A Cse4-containing nucleosome is deposited. CDEII is flanked by two palindromic sequences CDEI (8 bp) and CDEIII (25 bp). CDEI is bound by Cbf1 and although dispensable it cooperates to chromosome stability. CDEIII is important for centromere function since it serves as a binding site for the CBF3 complex which is required for proper chromosome segregation. The Ctf19 complex links the inner kinetochore complex Ctf3 to the Ndc80 complex. The interaction between the kinetochore and the single microtubule is mediated by the Dam1 complex, which physically interact with Ctf3 and Ndc80 complexes.
mediated by the Dam1 complex which physically interact with Ctf3 and Ndc80 complexes (Cheeseman et al., 2001; Measday et al., 2002).

1.1.2 *Holocentric centromeres*

Holocentric eukaryotes possesses kinetochore assembled along the entire length of chromosomes (reviewed in (Maddox et al., 2004). Holocentricity can be found in plant, insects as well as in nematodes. The holocentric centromeres from *C. elegans* are the one that have been better characterised. Despite this diversity in the architecture, the molecular structure and the composition of its kinetochores are similar to those of other organisms. The *C. elegans* CENP-A homologue, HCP-3, localises along the entire length of chromosome arms during mitosis (Buchwitz et al., 1999). ChIP-chip analyses have shown that HCP-3 occupies ~50% of the genome, indicating that specific DNA sequences are not required for the assembly of functional kinetochore (Gassmann et al., 2012). In addition, the incorporation of HCP-3 is inversely correlated with transcribed genes, suggesting that transcription may inhibit but define regions for HCP-3 deposition (Gassmann et al., 2012).

Consistent with the phenotypes associated with CENP-A deletion in other organisms, embryos depleted of HCP-3 show defects in mitotic chromosome segregation and are characterised by the typical kinetochore-null (KNL) phenotype (Buchwitz et al., 1999; Desai et al., 2003; Oegema et al., 2001). In *C. elegans*, HCP-3 is required for the localisation of the other known kinetochore proteins, such as the CENP-C homologue HCP-4 (Oegema et al., 2001). In contrast, depletion of other conserved kinetochore proteins does not affect HCP-3 incorporation into chromosomes indicating that HCP-3 directs centromere formation in mitosis (Desai et al., 2003; Oegema et al., 2001). In *C. elegans, de novo* centromere formation on injected extrachromosomal DNA occurs at high frequency and is not dependent on the DNA sequence (Mello et al., 1991; Yuen et al., 2011). In fact, LacO arrays injected into *C. elegans* germline are able to assemble centromeric chromatin and kinetochore proteins and autonomously segregate (Yuen et al., 2011). Interestingly, *C. elegans* heterochromatin is not required for *de novo* centromere formation and instead seems to antagonize HCP-3 deposition on extrachromosomal DNA (Yuen et al., 2011).

1.1.3 *Drosophila melanogaster centromeres*

In *Drosophila melanogaster*, studies on the *Dp1187* minichromosome have
identified a 420 kb region sufficient for centromere function (Murphy and Karpen, 1995; Sun et al., 1997). The main components of this region are the AATAT and AAGAG satellites which represent more than 85% of the total sequence and interspersed transposable elements which constitute the 10% of the total 420 kb (Sun et al., 1997). Analysis of transmission behaviour of minichromosome deletions have suggested that D. melanogaster centromeres are composed of two domains responsible for sister chromatid cohesion and kinetochore formation (Murphy and Karpen, 1995). Minichromosomes with deletion of the AAGAG block (Maupiti region) are completely lost indicating that the Maupiti element is absolutely required for centromere function and may include the site for kinetochore formation (Murphy and Karpen, 1995). In contrast, minichromosome derivatives with deletion of the AATAT main (named Bora Bora) display diminished stability and high loss rate, suggesting that this region may be involved in sister chromatin cohesion (Murphy and Karpen, 1995). However, the AATAT and AAGAG elements are not specific for centromere sequence since they can be found on other chromosomal regions that do not function as centromeres (Lohe et al., 1993). Therefore, the distribution of these satellite repeats within the genome indicates that these elements are not sufficient for centromere function (Sun et al., 1997).

Centromeric chromatic is composed of interspersed blocks of D. melanogaster CENP-A\textsuperscript{CID}, and histone H3 (Blower and Sullivan, 2002). CENP-A\textsuperscript{CID} localizes at the inner kinetochore and its localization correlates with centromeric DNA and function (Blower and Karpen, 2001). In addition, overexpression of CENP-A\textsuperscript{CID} leads to the formation of ectopic centromeres and multicentric chromosomes, which causes chromosome missegregation and growth defects, suggesting that CENP-A is a key epigenetic mark for centromere identity (Heun et al., 2006). Furthermore, tethering CENP-A\textsuperscript{CID} as a CID-GFP-LacI fusion protein to stably integrated \textit{lac} operator (lacO) arrays, can assemble functional and stable kinetochores which are maintained through generations even after removing the tethered CID-GFP-LacI (Mendiburo et al., 2011). These observations suggest that CENP-A\textsuperscript{CID} is both necessary and sufficient to serve as an epigenetic mark for centromeres formation.

1.1.4 Mammalian centromeres

Human centromeres are characterised by a chromosome-specific higher order array of 171 bp A/T-rich unit termed α-satellite, arranged in a tandem, head-to-tail configuration (Willard, 1985). Sequencing of a part of the human X-chromosome centromere indicated that centromere sequences in humans are homogeneous
repeats of \( \alpha \)-satellite DNA (Schueler et al., 2001). These higher order arrays can extend for 3-5 Mb and consists of the same \( \alpha \)-satellite unit repeated hundreds or thousands times within the same centromere. They are frequently interrupted by long interspersed element (LINE), short interspersed element (SINE) and long terminal repeat (LTR) retrotransposons (Schueler et al., 2005; Willard, 1989). Some of the \( \alpha \)-satellite monomers near the array edge contain a conserved binding site for the specific DNA binding protein CENP-B (Masumoto et al., 1989). Interestingly, mutation of the CENP-B box reduces the efficiency of centromere formation on a mammalian artificial minichromosome suggesting that the CENP-B box and CENP-B may be required for centromere establishment on \( \alpha \)-satellite DNA (Ohzeki et al., 2002).

In human centromeres, CENP-A nucleosomes associate with \( \alpha \)-satellite repeats however, CENP-A is present on only half to two-thirds of the entire centromere-specific \( \alpha \)-satellite array (Blower and Sullivan, 2002; Vafa and Sullivan, 1997). Similarly to *Drosophila* centromeres, the remaining part of the centromeric satellite array contains blocks of modified H3 nucleosomes linearly interspersed with CENP-A nucleosomes suggesting that this particular chromatin is an evolutionarily conserved aspect of centromere structure (Blower and Sullivan, 2002; Ribeiro et al., 2010). Although the existence of neocentromeres on non-satellite DNA suggest that alphoid DNA sequence is dispensable for centromere formation, the \( \alpha \)-satellite array is sufficient to establish *de novo* CENP-A chromatin when inserted into a human artificial chromosome (HAC) (reviewed in Bergmann et al., 2012b). Analyses of the frequency of centromere formation on human artificial chromosomes (HACs) have shown that kinetochore formation and the nucleation of CENP-A chromatin are dependent on the presence of \( \alpha \)-satellite and CENP-B boxes within the \( \alpha \)-satellite (Ohzeki et al., 2002; Okamoto et al., 2007). On the other hand, neocentromeres such as mardel(10) form on regions devoid of \( \alpha \)-satellite repeats and where CENP-B is not detected and these kinetochores are completely functional and are stably maintained over many cell divisions, suggesting that CENP-B may play a more significant role in establishment of centromeres on naked DNA (Voullaire et al., 1993). However, the role of CENP-B in HAC centromere formation remains unclear. It has recently been proposed that CENP-B may possess a dual role in centromere function. It may modulate the chromatin structure of the \( \alpha \)-satellite, promoting CENP-A chromatin establishment and it may also be involved in maintaining an heterochromatic environment over inactive HAC centromere DNA (Okada et al., 2007). In addition to the presence of CENP-B boxes within the \( \alpha \)-satellite, the length of the \( \alpha \)-satellite array plays an important role in the establishment of a functional
centromere on the HAC. Indeed, a yeast artificial chromosome (YAC) containing 10 kb of alphoid DNA does not recruit CENP-A while longer arrays of satellite repeats are able to form functional centromeres with an efficiency that is dependent on the length of satellite DNA present on the YAC, suggesting that a minimal length of centromeric DNA is required to recruit the factors responsible for centromere establishment (Okamoto et al., 2007).

1.2 Schizosaccharomyces pombe centromeres

1.2.1 Centromere structure

Fission yeast centromeres are composed of two principal domains: two outer repeats region (otr) that surround central core (cc/cnt) flanked by 2 innermost repeat regions (imr) (Figure 1-3) (Pidoux and Allshire, 2004; Steiner et al., 1993; Takahashi et al., 1992). Together the cc regions and the imr form the central domain of the fission yeast centromere where the centromeric histone variant CENP-A<sup>Cnp1</sup> is deposited (Takahashi et al., 2000). In contrast, the flanking otr repeats are embedded in heterochromatin and characterized by methylation of histone H3 on lysine 9 (H3K9me).

The central core is composed of ~72% AT-rich non-repetitive DNA. Centromere 1 (cen1) and cen3 are 99% identical over a ~4 kb region (the TM element) while the equivalent central region of cnt2 is more divergent with only 48% identity to the TM element (Pidoux and Allshire, 2004; Takahashi et al., 1992; Wood et al., 2002). The imr sequences are unique and specific for each centromere while all the centromeres share highly homologous otr repeats (Baum et al., 1994; Clarke et al., 1993; Fitzgerald-Hayes et al., 1982; Hegemann and Fleig, 1993; Takahashi et al., 1992; Wood et al., 2002). The otr regions are composed of dg and dh elements, which are present in variable numbers and arrangements in the three centromeres. In addition, one or more tRNA genes are found symmetrically arranged within the imr repeats and in the central domain of cnt3 (Goh and Kilmartin, 1993; Lechner and Carbon, 1991; Sorger et al., 1994; Takahashi et al., 1991). The function of this tRNA clusters is unclear however, it has been proposed that they may function as boundary elements between different types of chromatin – kinetochore, heterochromatin, euchromatin (Lechner and Carbon, 1991; Partridge et al., 2000; Scott et al., 2006).
Figure 1-3. DNA structure of fission yeast centromeres. The three fission yeast centromeres possess the same overall structure where two outer repeats region (otr) surround a central core (cc/cnt) flanked by 2 innermost repeat regions (imr). Together the cc regions and the imr form the central domain of the fission yeast centromere. cc1 and cc3 contain a TM element (white rectangle) that is identical and which has some similarity within cc2. The inverted imr repeats are unique for each centromere while the otr contain identical elements between the three chromosomes. tRNA genes are indicated with small black bars within the imr. The otr repeats together with part of the imr are embedded in heterochromatin and characterized by methylation of histone H3 at lysine 9 (H3K9me). The presence of heterochromatin is directed by the RNAi machinery. The central domain (cc region and part of the imr) is characterised by the presence of the centromeric histone variant CENP-A^{Cnp1} and corresponds to the region where kinetochore is assembled (adapted from (Pidoux and Allshire, 2004)).
1.2.2 Sequence requirements for centromere function

The DNA sequence requirements for centromere function have been investigated using circular and linear minichromosome assay systems (Baum et al., 1994; Folco et al., 2008a; Hahnenberger et al., 1989; Niwa et al., 1986; 1989; Polizzi and Clarke, 1991). In these studies, the stability of minichromosomes containing different fragments of otr and cc sequence was assessed by segregation frequencies. However, in contrast to budding yeast, where the minimal functional centromere consists of a conserved 125 bp sequence, in fission yeast it is not possible to isolate a small specific element that is sufficient to confer chromosome stability. Indeed, in *S. pombe* the establishment of functional centromere on a minichromosome requires otr repeats and a portion of the central domain (*imr* and *cc*) (Baum et al., 1994; Folco et al., 2008a; Marschall and Clarke, 1995). The otr sequences allows formation of heterochromatin, which is known to recruit cohesin and ensure proper chromosome segregation (Bannister et al., 2001; Bernard et al., 2001). Furthermore, the presence of heterochromatin has been shown to be required to initiate CENP-A<sup>Cnp1</sup> deposition over the central domain (Folco et al., 2008a; Kagansky et al., 2009) (see below). The central domain is the only sequence known to attract CENP-A and together with flanking otr DNA is required to assemble functional centromeres (Takahashi et al., 2000). Although the central domain sequences are not highly conserved at the three centromeres, this sequence is required to form a functional centromere on a minichromosome (Baum et al., 1994; Marschall and Clarke, 1995; Polizzi and Clarke, 1991). However, the reason that make the central domain the priming site for CENP-A and kinetochore assembly are not known. Previous analysis indicated that the cc2 sequence consists of modular elements since deletion of several different parts of its sequence did not affect centromere functionality (Baum et al., 1994). Therefore, it is possible that the presence of these multiple modular elements in the sequence may direct CENP-A<sup>Cnp1</sup> incorporation on the central domain.

1.2.3 Chromatin domain at fission yeast centromeres

Fission yeast centromeres consist of two distinct chromatin domains, the central CENP-A chromatin region and the flanking heterochromatin domain (Partridge et al., 2000). Initial analyses revealed that the chromatin structure within the central
domain is distinct from that of the flanking outer repeats; digestion with microccocal nuclease (MNase) yields a uniform smeared pattern for the central domain instead of the regular nucleosome ladder which is present over the outer repeats and most of the other chromatin loci (Figure 1-4A) (Polizzi and Clarke, 1991; Takahashi et al., 1992). This unusual nuclease cleavage pattern within the central domain may be caused by absence of a regular periodic nucleosomal array or the regular ladder is masked by the presence of a functional kinetochore assembled at the central core. Interestingly, cells with defective kinetochore formation and CENP-A^Cnp1 loading, display a ladder pattern typical of bulk chromatin (Goshima et al., 1999; Hayashi et al., 2004; Ohkuni and Kitagawa, 2011; Pidoux, 2003; Takahashi et al., 2000). Analyses of MNase nuclease pattern allowed specialised centromeric chromatin structure (the unusual MNase nuclease digestion pattern) and centromere function to be correlated. A non-functional circular minichromosome is characterised by a typical nucleosome-packaging pattern resulting in a regular nucleosome ladder (Figure 1-4B). In contrast, functional minichromosomes show a smeared pattern that is characteristic of the functional endogenous centromere (Figure 1-4C) (Marschall and Clarke, 1995; Ohkuni and Kitagawa, 2011).

1.2.4 Centromeric heterochromatin domain

Pericentromeric repeats in many species are assembled in heterochromatin. The flanking outer repeats at S. pombe centromeres are coated in heterochromatin where chromatin is underacetylated and histone H3 is specifically methylated on lysine 9 (H3K9). This specific modification allows the recruitment of Swi6, a heterochromatin protein 1 (HP1)-like protein. It has been shown that Swi6 is involved in cohesion recruitment at centromeres and consequently contributes to accurate chromosome segregation (Bannister et al., 2001; Bernard et al., 2001; Lachner et al., 2001; Maddox et al., 2004; Nakayama et al., 2000). Although marker genes inserted at the heterochromatic repeats are transcriptionally silenced, transcription of the otr repeats is required for the establishment of heterochromatin (Allshire et al., 1994; 1995; Djupedal et al., 2005; Kato et al., 2005; Volpe et al., 2002). Indeed, the presence of heterochromatin at fission yeast centromeres is dependent on the RNAi machinery and on the creation of a positive-feedback loop that ensures the maintenance of the heterochromatic domain (Figure 1-5) (reviewed in Lejeune et al., 2011; Volpe et al., 2002). The centromeric otr sequences are transcribed.
Figure 1-4. Centromeric central core chromatin structure of centromere-inactive and centromere-active minichromosomes in *S. pombe*. A) (Left panel) Ethidium bromide-stained gel of total DNA purified from chromatin partially digested with MNase. (Right panel) Autoradiogram of the genomic DNA (left) hybridised with a probed specific from *cen2*. B) DNA extracted from fission yeast transformed with a non functional minichromosome. A nucleosomal DNA ladder from the inactive centromeres on the minichromosomes is superimposed to the smear originating from the genomic copy of *cen2*. C) (Left) Ethidium bromide-stained gel of total DNA extracted from fission yeast transformed with a minichromosome that has a functional kinetochore assembled. (Right) Autoradiogram of the same genomic DNA hybridised with a probed specific from *cen2*. In this example, a smear is present and it is not possible distinguish the signal derived *cen2* on the minichromosome and the genomic *cen2* (adapted from Marschall and Clarke, 1995).
Figure 1-5. Model for heterochromatin formation in *S. pombe* (from Lejeune et al., 2011). Centromeric outer repeats are transcribed by RNA Pol II and converted into double stranded RNAs (dsRNA) that are processed by the RNase III–like enzyme Dcr1. The siRNA are loaded into Ago1 which targets the RITS complex to the otr repeats. RITS recruits the RDRC complex, generating additional dsRNA via Rdpl RNA polymerase. RITS also recruits the CLRC complex to chromatin, allowing the enzymatic subunit Clr4 to methylate the lysine9 of histone H3 (H3K9me). This mark is recognised by the chromodomain protein Swi6 that, in turn, ensures the recruitment of cohesin, which is required for accurate chromosome segregation.
during S-phase by RNA polymerase II and converted into double stranded RNA (dsRNA) (Chen et al., 2008; Zhang et al., 2008). This dsRNA is processed by the RNA III-like ribonuclease Dicer (Dcr1) into siRNA and subsequently loaded into the argonaute Ago1. Ago1 is part of the RITS complex (Ago1, Tas3, Chp1) and targets the complex to the chromatin through base pairing of the siRNA with the nascent transcripts. The RITS complex recruits the RNA-dependent RNA polymerase complex (RDRC; Rdp1, Cid12 and Hrr1) whose function is to amplify the signal by generating additional dsRNA (Motamedi et al., 2004). RITS is also involved into the recruitment of the CLRC complex (Clr4, Rik1 and Cul4) via an interaction with the Stc1 (siRNA to chromatin protein 1) (Bayne et al., 2010). The CLRC complex contains Clr4, the only histone methyl transferase (HMT) present in fission yeast specific for methylation of H3K9. Therefore, the recruitment of the CLRC complex induces the methylation of histone H3K9 and the binding of the chromodomain proteins such as Swi6, Chp1, Chp2 and Clr4 itself (Bannister et al., 2001; Zhang et al., 2008).

1.2.5 Heterochromatin and centromere function

Although defects in heterochromatin integrity affect chromosome segregation due to the reduce sister-centromeric cohesion, the level of CENP-A<sup>Cnp1</sup> at endogenous centromeres does not change. However, several studies indicate that heterochromatin flanking the central core is required to promote CENP-A<sup>Cnp1</sup> recruitment over the central domain when newly introduces on minichromosomes (Baum et al., 1994; Folco et al., 2008a; Kagansky et al., 2009). In fact, minichromosomes containing just the central region but not outer repeats we shown not be able to form functional centromeres (Folco et al., 2008b). Similarly, CENP-A<sup>Cnp1</sup> is not detected on minichromosomes containing out repeats flanking cc transformed in fo cells with defective heterochromatin. These observations indicate that pericentric heterochromatin is required for the establishment of a functional centromere, possibly by providing the appropriate chromatin environment for CENP-A<sup>Cnp1</sup> deposition (Folco et al., 2008a). Interestingly, targeting the methyl-transferase Clr4 to a minichromosome is sufficient to recruit CENP-A<sup>Cnp1</sup> to the adjacent central domain DNA, indicating that the RNAi machinery is not involved in the assembly of CENP-A<sup>Cnp1</sup> chromatin and the repeats themselves are not required (Kagansky et al., 2009). In addition, fission yeast neocentromeres have been shown to form in the
vicinity of telomeric heterochromatin and the frequency of neocentromere formation decreases in the absence of heterochromatic factors, again this suggests that the de novo establishment of centromeres requires a certain chromatin environment (Ishii et al., 2008). However, once CENP-A\(^{Cnp1}\) chromatin is established, heterochromatin becomes dispensable for its maintenance, suggesting that heterochromatin is exclusively involved in de novo centromere establishment by creating the proper environment or through the recruitment of loading factors specific for CENP-A\(^{Cnp1}\) deposition (Folco et al., 2008a).

1.2.6 Central domain and kinetochore structure

The central domain of fission yeast centromeres is the region where CENP-A\(^{Cnp1}\) is recruited and kinetochore is formed (Partridge et al., 2000; Takahashi et al., 2000). A large number of proteins are required for kinetochore assembly and most of these factors are conserved amongst eukaryotes (Figure 1-6). In vertebrates, the large multi-protein complex responsible for kinetochore assembly is termed the CCAN (constitutive centromere-associated network) and is composed of 16 different centromeric proteins (Figure 1-6) (Foltz et al., 2006; Okada et al., 2009). In S. pombe, several genetic screens have permitted the identification of a number of kinetochore proteins affecting CENP-A\(^{Cnp1}\) deposition. Many of these factors play similar conserved functions in higher eukaryotes. For example, the centromeric protein Mis6 and Mis12 were identified in screens for minichromosome loss and are enriched over the central domain. Although less robust than the heterochromatic regions, transcriptional silencing also occurs within the central domain and is dependent on the formation of functional centromeres and kinetochores. Therefore, mutations in genes encoding for factors that affect centromere structure alleviate the silencing of a marker gene inserted at the central core (Pidoux, 2003). Indeed, most of kinetochore mutants such as scm3, mis6 and sim4 alleviate silencing at cc (Dunleavy et al., 2007; Pidoux, 2003; Pidoux et al., 2009). In addition, most of these kinetochore mutants alter the normal chromatin structure associated with the central domain so that the typical smear pattern is lost and replaced by the nucleosomal ladder pattern present in other genomic regions (Figure 1-4) (Goshima et al., 1999; Pidoux, 2003; Saitoh et al., 1997; Takahashi et al., 2000).
Figure 1-6. Overview of the molecular organisation of vertebrates and fungi kinetochores. Schematic representation of A) vertebrate kinetochore, B) Fission yeast kinetochore and C) Budding yeast kinetochore. Proteins are represented in boxes with color representing the conservation among organisms. Physical interaction between protein/complexes is represented by dotted lines or overlap of between boxes. Large complexes are represented with their subunits as a single shape (adapted from Chan et al., 2005).
CENP-A. The centromeric histone H3 variant CENP-A\(^{Cnp1}\) is enriched over the central domain region and like the CENP-A orthologues in other organisms, is required for proper kinetochore formation (Takahashi et al., 2000). Indeed, the temperature-sensitive (ts) allele \(cnp1-1\) exhibits unequal mitotic segregation and increase incidence of lagging chromosomes when grown at restrictive temperature (36°C). In addition, the smeared MNase pattern characteristic of central domain chromatin at functional centromeres is abolished in the \(cnp1-1\) mutant, indicating that CENP-A\(^{Cnp1}\) is required to maintain a functional centromere-specific chromatin structure (Takahashi et al., 2000).

CENP-C. Although it is not absolutely necessary for kinetochore assembly, CENP-C is associated with functional centromeres (Earnshaw et al., 1989; Fukagawa et al., 1999; Tanaka et al., 2009). In \(S.\) pombe, deletion of the gene encoding for fission yeast CENP-C homologue, CENP-C\(^{Cnp3}\) is not lethal, thus CENP-C\(^{Cnp3}\) is not essential for kinetochore function (Tanaka et al., 2009). However, cells lacking CENP-C\(^{Cnp3}\) show a high incidence of lagging chromosomes and unequal nuclear division indicating impaired kinetochore-microtubule attachment (Tanaka et al., 2009). These defects in chromosome segregation are explained by the function of CENP-C\(^{Cnp3}\) as a scaffold for localisation of other kinetochore proteins. Indeed, two-hybrid screens have identified Pcs1 and CENP-L\(^{Fta1}\) (part of the CENP-H-I complex) as interactors of CENP-C\(^{Cnp3}\). Both Psc1 and CENP-L\(^{Fta1}\) facilitate the mono-oriented attachment of spindle microtubules to the kinetochore and play an essential role in kinetochore assembly (Gregan et al., 2007; Tanaka et al., 2009). In human cells, CENP-C is recruited by direct interaction with the C-terminal region of CENP-A (Carroll et al., 2010). In addition, biochemical analyses have shown that the C-terminus of CENP-C associates with Mis18BP1, which is required for the correct deposition of CENP-A. The N-terminus of CENP-C contacts the outer kinetochore Mis12 complex, suggesting that CENP-C may act as a linker between inner kinetochore and microtubule binding components (Dambacher et al., 2012; Screpanti et al., 2011).

The CENP-T/-W/-S/-X complex. In human cells, CENP-T and CENP-S were initially identified by mass spectrometry of CENP-A containing nucleosomes while CENP-W and CENP-X were found associated with CENP-T and CENP-S, respectively (Amano et al., 2009; Foltz et al., 2006; Hori et al., 2008). CENP-T and CENP-W are conserved between species and interact though their histone-fold domains to form a stable complex that shows DNA binding ability (Hori et al., 2008; Nishino et al.,
Also CENP-S and CENP-X possess a histone-fold domain and form a tight complex similar to CENP-T/-W. The interaction between the two complexes leads to the formation of a very stable heterotetramer in vitro that possess a nucleosome-like structure able to bind and assemble into plasmids in a manner that induces supercoiling similar to nucleosomes (Nishino et al., 2012). Therefore, it was suggested that the CENP-T-W-S-X complex may form an non-canonical nucleosome-like structure that might mark centromeric sequences and direct kinetochore establishment (Figure 1-7) (Nishino et al., 2012). Interestingly, CENP-T directly interacts with the Ndc80 complex suggesting that, like CENP-C, CENP-T also connects the centromeric chromatin to the outer kinetochore (Schleiffer et al., 2012). Although the fission yeast CENP-T/-W/-S/-X complex is not characterised, homologous proteins have been identified (Schleiffer et al., 2012). Similarly to vertebrate kinetochores, the S. pombe CENP-T homologue CENP-T\textsuperscript{Cnp20} seems to interact with the outer kinetochore proteins. Indeed, in cells expressing a ts allele of CENP-T\textsuperscript{Cnp20}, the localisation of CENP-L\textsuperscript{Fta1} to centromeres is reduced without affecting CENP-C\textsuperscript{Cnp3} localisation suggesting that CENP-T\textsuperscript{Cnp20} and CENP-C\textsuperscript{Cnp3} act in parallel to contact the outer kinetochore by recruiting CENP-L\textsuperscript{Fta1} (Tanaka et al., 2009).

The CENP-H/-I/-K complex. The fission yeast kinetochore protein Mis6 (homologue of CENP-I) was originally identified in a minichromosome loss screen where mutation of the mis6 gene caused defective chromosome segregation and increase minichromosome missegregation frequency (Goshima et al., 1999; Takahashi et al., 2000). ChIP experiments showed that Mis6 is enriched over the central domain and mutation of mis6 alleviates silencing in the central core, suggesting that Mis6 is required to maintain a normal centromeric chromatin structure. Indeed, mis6 mutants have reduced level of CENP-A\textsuperscript{Cnp1} within the central core and fail to incorporate newly synthesised CENP-A\textsuperscript{Cnp1} (Saitoh et al., 1997; Takahashi et al., 2000). These observations suggested that Mis6 may act as a centromeric loading factor for CENP-A\textsuperscript{Cnp1}. In contrast, the budding yeast homologue, Ctf3, is not required for the deposition of budding yeast CENP-A\textsuperscript{Cse4} (Measday et al., 2002).

Mis6 is found in complex with Sim4 (the homologue of CENP-K), an other kinetochore protein identified in a screen for factors that alleviate silencing at the central core (Pidoux, 2003). Sim4 is also enriched over the central domain and
Figure 1-7. Structure of vertebrate kinetochore. A large number of proteins are required for kinetochore assembly and most of these factors are conserved among eukaryotes. The large multi-protein complex responsible for kinetochore assembly is composed of several proteins conserved between different species. In the figure, CENP-T-W-S-X complex forms an non-canonical nucleosome-like structure that might mark centromeric sequences and direct kinetochore establishment (from Takeuchi and Fukagawa, 2012).
sim4 mutants display disrupted chromatin structure as observed in MNase chromatin digestion assay (Pidoux, 2003). Mis6 and Sim4 are part of the Sim4 complex that includes Mal2 (CENP-O), Mis15 (CENP-N), Mis17 (CENP-U), the DASH component Dad1 and the Fta1-7 proteins (Liu et al., 2005). Fta1 (CENP-L), Fta2 (CENP-P), Fta3 (CENP-H) and Fta4 were found to localise at the central domain of centromeres and play a role in microtubule-kinetochore attachment (Liu et al., 2005; Okada et al., 2006). Mis15 and Mis17 are found in complex with Mis6 and both localise at the centromere (Hayashi et al., 2004). mis15 and mis17 mutants show altered centromeric chromatin structure in MNase chromatin digestion assay and reduced CENP-A Cnp1 level at centromeres. Thus, Sim4, Mis6, Mis15 and Mis17 may act together in a pathway required for incorporation of CENP-A Cnp1 at centromeres (Hayashi et al., 2004).

Similarly, in human cells the deposition of newly synthesised CENP-A is impaired in cells lacking the CENP-H-I-K complex (Okada et al., 2006). In addition, CENP-H seems to facilitate CENP-A deposition by acting together with the chromatin remodelling factor FACT (facilitates chromatin transcription) (Okada et al., 2009).

The Mis12 complex. The Mis12 complex is part of the KMN (KNL1, Mis12 and Ndc80 complex) network that is required for mediating a proper attachment between kinetochores and microtubules (Cheeseman et al., 2004). In fission yeast, Mis12 was identified together with Mis6 in a screen for minichromosome loss (Goshima et al., 1999). Although Mis12 is not required for CENP-A Cnp1 localisation at the centromeres, Mis12 is enriched at the central domain and is required for proper chromosome segregation (Goshima et al., 1999). Mis12 is found in a complex with Mis13 (Dsn1), Mis14 (Nsl1), Nnf1 and Spc7 (KNL1) (Obuse et al., 2004). Spc7 localises to the kinetochore and is bound to the central domain of centromeres (Obuse et al., 2004). Spc7 interacts with microtubules through its interaction with Mal3 and therefore may facilitate the contact between the Mis12 and Sim4 complexes with the spindle microtubules (Kerres et al., 2007).

Similarly in human, the Mis12 complex contains four proteins (Nnf1, Mis12, Dsn1, and Nsl1) and acts as interaction hub between kinetochore and microtubules (Cheeseman et al., 2004; Kline et al., 2006; Obuse et al., 2004; Petrovic et al., 2010). In fact, on one hand, although the Mis12 complex does not bind directly to microtubules, Dsn1 interacts with KNL1, which does bind microtubules (Cheeseman et al., 2007). On the other hand, Mis12 contacts the inner kinetochore through a direct interaction with the N-terminus of CENP-C (Screpanti et al., 2011). In addition, cells depleted of Dsn1 display decreased levels of CENP-A at centromeres.
suggesting that the complex may affect the incorporation of CENP-A at centromeres (Kline et al., 2006). Mis12 forms a stable complex with the heterochromatin components HP1α and HP1γ suggesting that the Mis12 complex may bridge heterochromatin and kinetochore while HP1 may anchor the Mis12 complex to the pericentromeric heterochromatin (Obuse et al., 2004; Petrovic et al., 2010).

1.3 EPIGENETIC REGULATION OF CENTROMERES

Although a large number of proteins are required for kinetochore assembly and the function of centromeres is essential and conserved, surprisingly the primary DNA sequence underlying the centromere is not conserved between different species. The simplest point centromeres of budding yeasts are the only known centromeres that are completely specified by DNA sequence since point mutations in their DNA sequence destroy centromere function (Lechner and Carbon, 1991; Mello et al., 1991; Pidoux, 2003). However, this simple centromere structure is not conserved in other organisms and instead epigenetic mechanisms play a role in determining the region where centromeres are formed.

1.3.1 Centromeric DNA is not sufficient for centromere function

In many organisms, although centromeres are assembled on preferred sequences, the sequence itself is not always sufficient for centromere establishment and instead epigenetic mechanisms play a role in centromere specification (Allshire and Karpen, 2008). In fact, the presence of centromeric DNA alone does not ensure kinetochore formation and function. For example, the satellite repeats and the transposon components of Drosophila centromeres are not specific for centromeres since they can be found at other chromosomal regions which are not associated with kinetochore proteins (Murphy and Karpen, 1995; Sun et al., 1997). The acquisition of centromeric formation also seems to be epigenetically regulated in fission yeast. When plasmid based minichromosomes containing only part of the centromere sequence are transformed into S. pombe, the plasmids adopt either a “stable” or “unstable” segregation states by establishing functional centromeres stochastically. Once the functional state is acquired it is maintained through numerous divisions (Steiner and Clarke, 1994).

Centromeres can be inactivated without alteration of the DNA sequence. Human dicentric chromosomes can contain two distinct regions with centromeric alphsatellite DNA. However, frequently only one centromeric region possesses a
functional kinetochore. In these dicentric chromosomes, CENP-A is only localised at the active centromere while the other centromere is inactivated and maintained in a non functional state (Earnshaw et al., 1989; Murphy and Karpen, 1995; Sullivan and Schwartz, 1995). Stable dicentric chromosomes have been recovered in flies and mammals and they appear to be transmitted normally (Murphy and Karpen, 1995; Sullivan and Schwartz, 1995; Therman et al., 1974). However, it remains unclear how discrimination between the two alpha-satellite regions is achieved and how one of the centromeres is inactivated. A recent study on artificially generated dicentric chromosomes in fission yeast suggested that the inactivation of one of the centromeres is epigenetically regulated since there are no alterations in the DNA sequence. Furthermore, this inactivation is triggered by dissociation of kinetochore components and is followed by heterochromatisation of the centromeric sequences, indicating that the formation of the heterochromatic domain prevents reactivation of the inactivated centromere (Sato et al., 2012).

1.3.2 Centromeric DNA is not required for centromere function

The fact that kinetochores can form on non-centromeric DNA sequence provides further supports on the epigenetic nature of centromeres. Neocentromeres form on regions that do not possess any sequence similarity with centromeric DNA yet are able to assemble a completely functional kinetochore, and chromosomes bearing neocentromeres are mitotically and meiotically stable (Ishii et al., 2008; Lohe et al., 1993; Murphy and Karpen, 1995; Voullaire et al., 1993). In human cells, the mardel(10) chromosome lacks centromeric DNA but remains mitotically stable since it possesses a neocentromere capable of normal kinetochore assembly and function (Sun et al., 1997; Voullaire et al., 1993). In Drosophila, targeting of CENP-A^{CID} to ectopic loci that do not possess homology to centromeric DNA, leads to the formation of functional centromeres indicating that it is CENP-A^{CID} and not the DNA sequence that regulates the place where centromeres are formed (Mendiburo et al., 2011). In C. albicans, centromeres are strongly regulated by epigenetic mechanisms. The deletion of centromeres allows the formation of neocentromeres in 100% of cases analysed on DNA sequences that do not possess any similarity (Ketel et al., 2009). In fission yeast, the frequency of neocentromere formation is much lower than that observed in C. albicans and upon deletion of endogenous centromere neocentromeres are formed exclusively at the subtelomeric regions, suggesting that the subtelomeric heterochromatin may be required for kinetochore assembly (Ketel et al., 2009). Therefore, in C. albicans neocentromere formation
might also occur upon activation of the neocentromeric region by a specific chromatin environment, although there is no evidence for this. Alternatively, endogenous centromeres may somehow inhibit the formation of neocentromeres, which could then arise when the repressive mechanism is removed.

1.4 CENP-A NUCLEOSOME STRUCTURE

It is clear that all functional centromeres are characterised by the presence of CENP-A and instead epigenetic mechanisms play a role in centromere specification. Therefore, CENP-A is the best candidate for the epigenetic mark that specifies the site of kinetochore assembly. In fact, tethering CENP-A to a ectopic locus on a chromosome results in recruitment of kinetochore proteins and assembly of a functional kinetochore that capture microtubules. This indicates that CENP-A alone is sufficient to specify the site where the kinetochore is assembled (Mendiburo et al., 2011). In addition, analysis of histone dynamics shows that there is almost no turnover of CENP-A at centromeres strongly suggesting that CENP-A marks the region where centromeres are assembled (Heun et al., 2006; Jansen et al., 2007; Mendiburo et al., 2011; Shelby et al., 2000).

Although CENP-A is a highly divergent histone H3 variant, its overall structure does not differ from canonical histone H3. In fact, like all histones, CENP-A contains a globular histone fold domain (HFD) that is composed of two flanking alpha helixes (α1 and α3) and one central alpha helix (α2) connected by loops (Black et al., 2004; Sullivan et al., 1994) (Figure 1-8A). The CENP-A targeting domain (CATD) consists of loop1 and α-helix 2 of the histone fold and is sufficient for efficient localisation of CENP-A at centromeres. Indeed, when substituted into conventional H3, the CATD converts H3 into a centromeric histone and is sufficient to target the CATD-H3 chimeric protein to centromeres (Black et al., 2007). Despite the overall structural similarity, CENP-A shows no homology at its N-terminus with histone H3. Indeed, CENP-A contains a N-terminal domain that is highly divergent between different organisms in both length and amino acid composition, suggesting that CENP-A evolves more rapidly than the conserved canonical histone H3. It is
Figure 1-8. CENP-A structure and models for CENP-A containing nucleosomes (adapted from Black and Cleveland, 2011). A) Schematic representation of the structure of the centromeric histone H3 variant CENP-A. The structure contains globular histone fold domain (HFD) that is composed of two flanking alpha helixes (α1 and α3) and one central alpha helix (α2) connected by loops. The CENP-A targeting domain (CATD) consists of loop 1 and α-helix 2 of the histone fold and is sufficient for efficient targeting of CENP-A. Post-translational modifications have been identified in human (black) and budding yeast (blue). B) Possible model for CENP-A nucleosomes. Octameric nucleosomes: CENP-A replaces histone H3 in a canonical nucleosome structure containing two copies of each H2A/H2B/H4/CENP-A. Hemisome: CENP-A containing nucleosomes contain just one copy of each histone H2A/H2B/H4/CENP-A. Hexasome: the tetramer formed by CENP-A<sup>Cse4</sup>/H4 is in complex with the CENP-A chaperon Scm3 that displaces H2A/H2B dimers.
probable that this fast evolutionary rate reflects the ability to adapt to the of the centromeric DNA that in any organisms is amongst the most rapidly evolving DNA sequences (Malik and Henikoff, 2009).

Canonical histones are modified by different post-translational modifications that are required for the regulation of chromatin states and the establishment of functional domains. However, only few modifications on CENP-A have been reported (Figure 1-8A). The phosphorylation of serine 7 by Aurora B has been detected in human cells and seems to have a role in cytokinesis. However, the serine at position seven is not a well-conserved feature of CENP-A, even within vertebrates (Zeitlin et al., 2001a; 2001b). Bui et al, (2012) have also reported the acetylation of lysine 124 (K124) although its biological function has not yet been investigated (Bui et al., 2012). In S. cerevisiae, the levels of CENP-A\textsuperscript{Cse4} are controlled through the ubiquitination of its CATD domain by the ubiquitin ligase Psh1 and prevents the mislocalisation of CENP-A\textsuperscript{Cse4} to ectopic loci (Hewawasam et al., 2010a; Ranjitkar et al., 2010) . In addition, CENP-A\textsuperscript{Cse4} has been recently found methylated on arginine 37 and this modification seems to regulate the recruitment of kinetochore components to centromeres (Samel et al., 2012).

It is generally assumed that CENP-A is incorporated into nucleosome composed of eight subunits. However, the actual composition of CENP-A-containing nucleosomes is a matter of debate (Figure 1-8B) (reviewed in Black and Cleveland, 2011; and Henikoff and Furuyama, 2012). Some analyses support the simplest assumption that CENP-A, like canonical H3, forms octameric particles with two copies of each histones H2A, H2B, H4 and CENP-A (Figure 1-8B). Indeed, \textit{in vitro} and \textit{in vivo} analysis have shown that human CENP-A can replace histone H3 nucleosomes and CENP-A nucleosomes contain two copies of histone CENP-A rather than one copy of each H3 and CENP-A (Foltz et al., 2006; Yoda et al., 2000). The crystal structure of a CENP-A-containing nucleosomes has been solved and showed characteristics similar to H3 nucleosomes supporting the idea of octameric CENP-A nucleosomes (Tachiwana et al., 2011). Similarly, in \textit{D. melanogaster} two CENP-A\textsuperscript{CID} molecules can be cross-linked within the same nucleosome particle nucleosomes, indicating that CENP-A\textsuperscript{CID} nucleosomes are octameric \textit{in vivo} (Zhang et al., 2012). Moreover, point mutations of the residues required for the interaction between two CENP-A\textsuperscript{CID} molecules result in mislocalitation of CENP-A\textsuperscript{CID} nucleosomes suggesting that dimerisation within the same nucleosome is essential for correct centromere formation (Zhang et al., 2012).

However, observations in Drosophila S2 cells and more recently in mammalian cells, have led to propose an alternative “hemisome” model (Bui et al., 2012; Dalal...
et al., 2007; Dimitriadis et al., 2010). Atomic force microscopy (AFM) analysis of nucleosome particles isolated from in culture cells showed that CENP-A-containing nucleosomes are half the height of canonical H3 chromatin suggesting the presence of a nucleosome containing just one copy of each histone H2A/H2B/H4/CENP-A at the centromere. Further support for the hemisome model is provided from analysis in budding yeast where CENP-A<sup>Cse4</sup> nucleosomes can wrap less DNA in a right hand super-helix (Furuyama and Henikoff, 2009).

A third model, suggested by analysis of budding yeast centromeres, is the “hexosome” in which a tetramer formed by the CENP-A<sup>Cse4</sup>/H4 dimer is proposed to be in a complex with Scm3, a specific chaperone for CENP-A<sup>Cse4</sup> required for centromere function and able to displace H2A/H2B dimers in the centromeric nucleosome (Mizuguchi et al., 2007; Xiao et al., 2011). However, this model is likely to be specific for budding yeast, since S. pombe Scm3 and its homolog in human HJURP localise at centromeres independently of CENP-A<sup>Cnp1</sup> (Dunleavy et al., 2009; Foltz et al., 2009; Pidoux et al., 2009; Williams et al., 2009).

Recently, a new model has been proposed that takes in consideration both the “octameric model” and the “hemisome model”. Using quantitative imaging and biochemical analysis, it was suggested that CENP-A nucleosomes are generally present in the hemisome form. However in this model, hemisomes undergo structural transitions during the cell cycle assuming an octameric conformation during anaphase (in yeast) or early S-phase (in human cells) (Bui et al., 2012; Shivaraju et al., 2012). Since several contrasting models have been proposed, further studies are required to improve analysis of the complexes formed by CENP-A nucleosomes at centromeres at different stages of the cell cycle.

1.5 CENP-A ASSEMBLY

1.5.1 Timing of CENP-A assembly

In all organisms, the deposition of CENP-A at centromeres occurs during specific phases of the cell cycle. The regulation of the timing of deposition as well as the availability of CENP-A play a key role in preventing uncontrolled CENP-A deposition that leads to chromosomes missegregation. Indeed, the overexpression of CENP-A<sup>CID</sup> causes mislocalisation of CENP-A<sup>CID</sup> and formation of ectopic centromeres in Drosophila (Heun et al., 2006). Similarly in budding yeast, degradation of CENP-A<sup>Cse4</sup> by the E3 ubiquitin ligase Psh1 prevents the misincorporation of CENP-A<sup>Cse4</sup> in non-centromeric regions, suggesting that the level of CENP-A and the timing of its
deposition are tightly regulated throughout the cell cycle (Hewawasam et al., 2010b; Ranjitkar et al., 2010). Although in different species the deposition of newly synthesized CENP-A occurs at different stages of the cell cycle, a very conserved feature is that CENP-A assembly occurs independently of replication. In human and Drosophila, centromeric DNA is replicated in mid-to-late S phase while the level of CENP-A are not maximal until the end of S-phase (Hagen et al., 1990; Shelby et al., 2000). In human cells, elegant pulse-chase experiments using SNAP-tagged CENP-A demonstrated that CENP-A loading occurs once per cell cycle, in particular during telophase/early G1 and requires passage through mitosis (Jansen et al., 2007). In Drosophila, CENP-A\textsuperscript{CID} incorporation at centromeres initiates during metaphase and the CENP-A loading machinery is kept inactive by Cdk5 activity during S, G2, and M phase and the assembly of CENP-A requires degradation of cyclin A (Silva et al., 2012). In contrast, in S. pombe newly synthesised CENP-A can be deposited both during S phase and in G2 (Takayama et al., 2008). The S phase deposition is dependent on the GATA-like transcription factor Ams2, which regulates histone expression (Takayama and Takahashi, 2007). Indeed, deletion of ams2 leads to impaired CENP-A loading during S phase and CENP-A accumulates on centromeres via the G2 assembly machinery which is in turn inhibited by Hip1, a homologue of the HIRA histone chaperone (Takayama et al., 2008).

The fact that CENP-A is deposited during mitosis and G1 implies that the amount of CENP-A at centromeres is diluted during S phase and at half its maximal amount during most of the cell cycle (Jansen et al., 2007). Indeed, analysis of extended chromatin fibers in Drosophila showed that G1 fibers contain 2-fold more endogenous CENP-A compared to S-phase fibers (Dunleavy et al., 2011). Therefore, if no new CENP-A is deposited, what happens to the chromatin state of centromeres after replication? One possibility is that CENP-A nucleosomes are distributed to the daughter chromosomes with formation of nucleosome-free regions (Allshire and Karpen, 2008; Black and Cleveland, 2011; Hayashi et al., 2004). The “gaps” could then be filled with H3-containing nucleosomes. Indeed, recent analyses suggest that histone H3 are deposited in place of CENP-A during replication (Dunleavy et al., 2011). Immunobridisation studies on chromatin fibers showed that histone H3 variants H3.1 and H3.3 are both deposited at centromeres during replication. However, the reduced level of H3.3 (but not H3.1) in G1 compared to S-phase and the corresponding deposition of CENP-A suggest that H3.3 acts as a “placeholder” for CENP-A in S-phase (Dunleavy et al., 2011). A second model
Chapter 1: Introduction

1.5.2 Factors affecting CENP-A deposition

A large number of proteins are required for kinetochore assembly and most of those factors are conserved among eukaryotes. In *S. pombe*, several genetic screens combined with biochemical approaches have permitted the identification of a number of proteins affecting CENP-A deposition. Many of these factors play similar conserved function in higher eukaryotes.

The Mis16-Mis18 complex is conserved between species and acts most upstream in CENP-A targeting to centromeres and its localization does not require other kinetochore proteins (Fujita et al., 2007; Hayashi et al., 2004). In *S. pombe*, Mis18 localizes at centromeres in late anaphase to telophase just before the S-phase loading of CENP-A (Hayashi et al., 2004). Similarly, in human the two isoforms hMis18α and hMis18β accumulate at centromeres from telophase to early G1, the stage in which the recruitment of newly synthesized CENP-A occurs, suggesting a role of the Mis18 complex in priming centromeres for the incorporation of CENP-A at centromeres (Fujita et al., 2007; Shelby et al., 2000). Recently, Mis18 was reported to interact with the DNA-methyl transferase 3 (DNMT3A/B) and this interaction is critical to maintain the level of DNA methylation. In fact, deletion of Mis18 leads to a decreased level of methylated DNA with correspondent alteration of the chromatin state and centromeric transcription (Kim et al., 2012). In human cells, the two isoforms of Mis18 interact with Mis18 Binding Protein 1 (Mis18BP1-KNL2), which is required for CENP-A localization. The cell cycle-dependent phosphorylation of Mis18BP1-KNL2 plays a key role in the control of the timing of CENP-A assembly (Fujita et al., 2007; Silva et al., 2012).

In human cells, the hMis18 complex cooperates with RbAp46 and RbAp48 (Mis16 in fission yeast), the histone H3/H4 chaperones that are part of the CAF-1 (chromatin
assembly factor-1) complex (Fujita et al., 2007). CAF-1 functions are tightly associated with DNA replication and DNA repair and include chromatin remodelling and histone deacetylation (Loyola and Almouzni, 2004). Interestingly, in human cells and in fission yeast, RbAp48/Mis16 and Mis18 appear to be required for maintaining the central core domain in hypo-acetylated state (Hayashi et al., 2004).

In *S. pombe*, Mis16 and Mis18 are also required for the recruitment of Mis6, Mis15 and Mis17 to centromeres, which are also involved in CENP-A\(^{Cnp1}\) loading at centromeres (Hayashi et al., 2007; Saitoh et al., 1997). Mis6, the homologue of CENP-I, forms a complex with Sim4 (CENP-K), Mis15 (CENP-N), Mis17 (CENP-U) and it localizes at the centromere throughout the cell cycle (Hayashi et al., 2004; Pidoux, 2003). Moreover, Mis6 is required for the incorporation of newly synthesised CENP-A\(^{Cnp1}\) (Takahashi et al., 2000). However, the budding yeast homologue Ctf3 is not required for loading of CENP-A\(^{Cse4}\) (Measday et al., 2002). The vertebrate homologues CENP-I and CENP-H (respectively Mis6 and Sim4 in fission yeast), are required for efficient incorporation of newly synthesized CENP-A in cooperation with FACT (facilitates chromatin transcription) and CHD1 (chromodomain-helicase-DNA-binding protein 1) (Okada et al., 2006; 2009). Interestingly, deletion of the fission yeast homolog Hrp1, results in reduction of CENP-A\(^{Cnp1}\) associated at the centromeres suggesting that chromatin remodelling is required for CENP-A incorporation (Walfridsson et al., 2007). In addition, Hrp1 chromatin remodelling occurs at some cryptic RNAP II promoters and a low level of CENP-A\(^{Cnp1}\) associates with gene promoters where histone H3 is depleted by the activity of Hrp1 (Choi et al., 2011).

Scm3 is CENP-A specific chaperone and is essential for the localization CENP-A\(^{Cnp1}\) (Pidoux et al., 2009; Williams et al., 2009). Scm3 interacts with CENP-A\(^{Cnp1}\) and with the Mis16/Mis18 complex and since a direct interaction between Mis16/Mis18 and CENP-A\(^{Cnp1}\) has not been detected, *S. pombe* Scm3 may represent the connection between Mis16/Mis18 and CENP-A\(^{Cnp1}\) (Pidoux et al., 2009; Williams et al., 2009). In human cells, a similar function has been observed for HJURP, an histone chaperone related to Scm3 (Sanchez-Pulido et al., 2009). HJURP was found to physically interact with CENP-A and its depletion causes mislocalization of CENP-A at centromeres with consequent chromosome segregation defects (Dunleavy et al., 2009; Foltz et al., 2009). HJURP localises at centromeres during G1 phase when newly synthesised CENP-A is loaded but later than Mis18. These observations suggest that the Mis18 complex may alter chromatin structure by modifying histone
acetylation or by promoting histone H3 removal and HJURP/Scm3 may act as delivery system for CENP-A assembly.

A soluble chaperone homologous to the human histone-binding protein NASP, Sim3, has been shown to be required for the deposition of newly synthesized CENP-A<sup>Cnp1</sup> at fission yeast centromeres (Dunleavy et al., 2007). Indeed, cells defective for Sim3 function show reduced levels of CENP-A<sup>Cnp1</sup> accompanied with an increase in H3 at centromeres and defects in chromosome segregation. The Sim3 homolog in Xenopus, N1/N2, is involved in storing histone H3 and H4 in the cytoplasm, therefore it might be possible that in fission yeast, Sim3 chaperones CENP-A on route to centromeres, preventing its mislocalisation (Dunleavy et al., 2007).

In fission yeast, Ams2 is a GATA-like transcription factor isolated as a multicopy suppressor of the <i>cnp1-1</i> mutant (Chen et al., 2003; Takayama et al., 2008). In <i>ams2</i> deleted cells, CENP-A<sup>Cnp1</sup> is not deposited at centromeres in S-phase while it starts to accumulate when cells enter in G2 phase via a Ams2-independent pathway (Takayama et al., 2008). <i>ams2Δ</i> cells exhibit reduction of CENP-A<sup>Cnp1</sup> within centromeres and aberrant mitosis with unequal segregation. However, it is not clear how Ams2 mediates the S-phase incorporation of CENP-A<sup>Cnp1</sup>. Ams2 has also been shown to be a regulator of all core histone genes but not the gene encoding CENP-A<sup>Cnp1</sup> and it binds directly the promoter region by a consensus sequence found upstream of histone genes (Takayama and Takahashi, 2007). It has also been shown that the amount of histones present in the cells affect centromere function (Castillo et al., 2007). Therefore, it is possible that the effect of Ams2 deletion on centromeres is caused by histone imbalance.

1.5.3 Chromatin environment and CENP-A deposition

The site where centromeres are formed in some organisms, such as budding yeast, is determined by a short DNA sequence that recruits specific DNA binding factors and kinetochore proteins (Westermann et al., 2007). However, in other organisms the regions where centromeres form is not specified by any particular DNA sequences and instead, epigenetic mechanisms may contribute to the selection of the site of kinetochore assembly. In many species, centromeres and kinetochores are often embedded in pericentromeric heterochromatin. For example, the CENP-A<sup>Cnp1</sup> chromatin of fission yeast in surrounded by heterochromatic regions that are
characterized by underacetylated and specific modified histone H3, in particular methylated histone H3 at lysine 9 (H3K9) (Bannister et al., 2001). In *S. pombe*, although the absence of heterochromatin does not alter the amount of CENP-A<sup>Cnp1</sup> at endogenous centromeres, heterochromatin is necessary to allow the assembly of CENP-A<sup>Cnp1</sup> on minichromosomes possibly by creating a proper environment for centromere formation (Folco et al., 2008b; Kagansky et al., 2009). Interestingly, heterochromatic regions are also found at the subtelomeric regions in close proximity of the fission yeast neocentromeres, suggesting a possible role in the selection of the region where centromeres are formed (Ishii et al., 2008; Takayama and Takahashi, 2007). Similarly, in human neocentromeres the detection of the heterochromatin-associated protein HP-1 at neocentromeres indicates that some heterochromatic elements are required for centromeres formation (Saffery et al., 2003). However, not all neocentromeres possess these heterochromatic domains, suggesting that they are not necessary for centromere function (Alonso et al., 2010). In *D. melanogaster*, the induction of synthetic heterochromatin by targeting HP1-LacI fusions to a stably integrated Lac operator arrays produces a hotspot for CENP-A<sup>CID</sup> deposition indicating that some characteristics of heterochromatin promote de novo kinetochore assembly and contribute to centromere identity (Olszak et al., 2011). On the other hand, in other organism such as *C. elgans*, heterochromatin is not required for de novo centromere formation and instead seems to antagonise centromere formation on a extrachromosomal DNA (Yuen et al., 2011).

A specific pattern of histone modification may contribute to the specification of centromere position. In human and *D. melanogaster*, immunolocalisation analysis of two-dimensional extended chromatin fibres have demonstrated that centromeres are enriched in H3K4me2, H3K36me2, and H3K36me3 suggesting that a particular combination of histone marks may contribute to the selection of the region where the kinetochore assembles (Cam et al., 2005; Ribeiro et al., 2010; Sullivan and Karpen, 2004). Similarly, centromeres formed on human artificial chromosomes (HAC) are characterised by a typical pattern of histone modifications (Nakano et al., 2008). Interestingly, manipulation of this pattern through the targeting of histone modifier enzymes affects the centromere structure of the HAC and is accompanied by reduction in the levels of CENP-A and CENP-C being deposited (Bergmann et al., 2012a; 2011; Nakano, 2003; Nakano et al., 2008).

1.5.4 *Sequence specificity and centromere function*
Although centromeres are assembled in regions that do not appear to share homology between different species and several analysis points to the conclusion that centromeres are epigenetically determined, the fact remains that CENP-A assembly generally occurs on certain preferred sequences. Therefore, it is possible that within the same species or between different species centromeres and neocentromeres share some underlying features (e.g. biological process) that result in their choice as the site where functional centromeres are assembled.

A particular elevated A/T nucleotide content seems to be one of the major characteristics shared between centromeres of different organisms. For example, in human centromeres are assembled on the A/T rich tandemly repeated called alpha-satellite (Heun et al., 2006; Willard, 1985). Similarly, D. melanogaster centromeres are composed of stretches of A/T-rich satellite repeats together with several interspersed transposable elements (Mendiburo et al., 2011; Murphy and Karpen, 1995). The presence of these long stretches of A/T nucleotides may serve to select for the assembly of a certain type of nucleosomes (e.g. CENP-A). In fact, it has been shown that sequence composition and DNA sequence preferences can direct canonical H3 nucleosome positioning \textit{in vivo} (Kaplan et al., 2010; Segal et al., 2006). \textit{In vitro} analyses have demonstrated that nucleosomes can wrap a DNA molecule with different affinities and these data are confirmed by nucleosomes position preferences \textit{in vivo}. The comparison between \textit{in vivo} and \textit{in vitro} data allowed the creation of an algorithm capable of predicting nucleosome positioning \textit{in vivo} based on the DNA sequence alone (Kaplan et al., 2009). Nevertheless, the poly(dA:dT) elements are known to possess nucleosome-excluding properties and are favoured at promoter regions in S. cerevisiae where they generate a zone of nucleosome depletion (Segal and Widom, 2009a). Therefore, it is possible that H3 nucleosomes also have a lower affinity for centromeric DNA due to its nucleotide composition, this may influence nucleosome stability and favour the replacement of H3 with CENP-A containing nucleosomes which may have in turn a higher affinity for centromere DNA sequences.

Another possibility is that centromeric DNA sequences could contain binding sites for specific centromeric proteins. For example, in human centromeres, some of the alpha-satellite monomers contain a 17-bp motif recognised by CENP-B which is involved in centromere establishment (Ando et al., 2002). In addition, other kinetochore proteins such us HJURP$^{\text{Sor3}}$ in budding yeast and CENP-C in human and maize, possess DNA binding ability with preferences for A/T-rich DNA.
sequences (Du et al., 2010; Sugimoto et al., 1994; Xiao et al., 2011; Yang et al., 1996).

The DNA sequence at centromeres may also promote some biological activity (e.g. transcription, replication) that facilitates chromatin remodelling and CENP-A deposition. For example, a relationship between replication and neocentromere formation was recently observed in C. albicans. In this organism, neocentromeres form at several loci that share some property of low gene density surrounded by repeat elements and the inheritance of centromere position is correlated with an active origin of replication that functions at early S-phase (Ketel et al., 2009; Koren et al., 2010). Alternatively, other biological processes such as transcription may be involved in centromere function (see section below).

Thus, specific DNA sequences do not seem to be essential for CENP-A and kinetochore assembly but could provide a favourable environment for establishing centromeric chromatin. Once functional centromeres are established, epigenetic mechanisms may then take over and ensure maintenance and propagation of CENP-A chromatin at that site, irrespective of the surrounding chromatin/DNA context.

1.6 TRANSCRIPTION AND CENP-A DEPOSITION

Despite the fact that centromeres are generally located in poorly transcribed regions, increasing evidence suggests that transcription is associated with centromeres and centromeric transcripts might play important roles in centromere function. In many organisms, RNA polymerase II (RNAPII) and its associated transcription factors localise at centromeres and alteration of RNAPII activity leads to increase of chromosome missegregation during cell division (Chan et al., 2012; Nakano et al., 2008; Ohkuni and Kitagawa, 2011; Willard, 1985). Moreover, actively transcribed genes are also found within neocentromere regions, and their expression does not affect the presence of a functional kinetochore (Ishii et al., 2008; Saffery et al., 2003; Schueler et al., 2001). Centromeric satellite and retrotransposon transcripts have been identified in different organisms, including human, mouse and maize (Carone et al., 2009; Chen et al., 2008; Schueler et al., 2005; Topp et al., 2004; Willard, 1989). In maize, transcripts originating from centromeric DNA elements are part of the centromeric structure and may be involved in CENP-A recruitment as they co-immunoprecipitate with CENP-A nucleosomes (Masumoto et al., 1989; Topp et al., 2004). Similarly, transcripts
originating from satellite repeats of marsupials are bound by centromere proteins and may perform similar functions (Carone et al., 2009; Ohzeki et al., 2002). Further evidence suggesting a role of transcripts in centromere formation is provided by studies on human neocentromeres (Chueh et al., 2009; Nakano et al., 2008; Voullaire et al., 1993). The mardel(10) neocentromere is formed over a domain with higher than average prevalence of LINE retrotrasposons in the DNA sequences where CENP-A is deposited. The LINEs are transcribed and RNAi-mediated knockdown of the RNA products leads to reduced CENP-A levels at centromeres and instability of the chromosome carrying the neocentromere, suggesting that these retrotrasposon transcripts contribute to centromere function (Blower and Sullivan, 2002; Chueh et al., 2009; Vafa and Sullivan, 1997).

Although generally permissive for transcription, centromeric DNA is generally transcribed at a lower level compared to that of other chromosomal regions and maintaining the right level of centromeric transcription seems to be important for centromere function. For example, in budding yeast, altering the level of centromeric transcription results in an increased number of lagging chromosomes (Blower and Sullivan, 2002; Chueh et al., 2009; Ohkuni and Kitagawa, 2011). Similar effects have been described by manipulating the chromatin state of a human artificial chromosome (HAC) by targeting a transcriptional activator or a transcription repressor to the HAC centromere causing missetegration and loss of the HAC (Bergmann et al., 2011; Nakano et al., 2008). This suggests that either a highly transcriptional environment or the repression of the transcription are not compatible with a functional kinetochore (Nakano et al., 2008; Ohkuni and Kitagawa, 2011; Pidoux and Allshire, 2004; Wood et al., 2002).

In Drosophila and human cells, centromeres show a distinct chromatin structure enriched in H3K4me2, H3K36me2, and H3K36me3, modifications normally associated with transcriptionally active regions (Cam et al., 2005; Sullivan and Karpen, 2004). Interestingly, H3K4me2 is required for the recruitment of the CENP-A chaperone HJURP and new CENP-A loading on a HAC (Bergmann et al., 2011).

1.6.1 Transcription at fission yeast centromeres

Genes placed within S. pombe centromeres are transcriptionally silenced (Allshire et al., 1994; 1995). Mutants have been identified that specifically alleviate silencing within the heterochromatic repeats or in the central domain. An ura4 marker gene inserted in the central core is transcribed at lower level but incorporates CENP-
A\textsuperscript{Cnp1}, indicating that a certain level of transcription is compatible with centromere function (Allshire et al., 1994; Castillo et al., 2007).

In fission yeast, the otr repeats surrounding the central domain are transcribed and this process is responsible for the establishment and maintenance of the heterochromatic domain (Djupedal et al., 2005; Kato et al., 2005). In addition, transcripts originating from within the central CENP-A\textsuperscript{Cnp1} chromatin domain of cen1 have been recently detected and their transcriptional start sites mapped (Choi et al., 2011). These transcripts have been termed TUKs (transcripts underlying kinetochores) and are normally degraded in wild-type cells by the exosome. In cells with a mutation in a major subunit of the exosome (dis3 mutants), a smear of transcripts is detected by northern using strand specific probes (Figure 1-9). In contrast, mutants that affect CENP-A\textsuperscript{Cnp1} loading at centromeres such as mis6-302 and cnp1-1, allow the detection of discrete sized poly-adenylated transcripts (Choi et al., 2011). In this condition, when CENP-A\textsuperscript{Cnp1} chromatin is disrupted, H3 replaces CENP-A\textsuperscript{Cnp1} at centromeres and this correlates with an increase in centromeric transcription (Castillo et al., 2007; Choi et al., 2011). This suggests that transcription is distinct in H3 and CENP-A\textsuperscript{Cnp1} chromatin.

1.6.2 Is there a link between transcription and CENP-A localisation?

It is known that during transcription nucleosomes are disassembled from their original position in front of RNAPII allowing it to elongate through the sequences contained in the nucleosomes. Analyses, mainly in S. cerevisiae, suggests the presence of a mechanism that controls the disassembling of the nucleosomes in front of the elongating polymerase II and the reassembly of the nucleosomes behind the passing RNA Pol II (Boeger et al., 2003; Rocha and Verreault, 2008; Williams and Tyler, 2007).

During the reassembly mechanism, it is possible that other histone variants may replace canonical H3 nucleosome. This is the case in higher eukaryotes where transcribed genes are enriched for the histone variant H3.3 (Ahmad and Henikoff, 2002). In D. melanogaster, regular H3 nucleosomes are lost during transcription and the histone variant H3.3 is deposited at sites of nucleosomal displacement throughout the transcribed genes (Wirbelauer et al., 2005). Furthermore, a recent study suggested that transcription at human centromeres might be essential since
Figure 1-9. 5’-RACE/PCR and Northern analyses of *S. pombe* cnt1 transcripts. A) schematic representation of transcription start sites determined by 5’-RACE/PCR in wild type (wt) and *cnp1-1* cells. Black arrows, transcription start sites identified in wt and *cnp1-1* cells. B) Northern analysis of total RNAs and Poly(A) purified RNA from exosome (*dis3-54*) and kinetochore mutants (*mis6-302* and *cnp1-1*). An RNA probe complementary to *cnt1* was used. Figure shows blot for RNA extracted at restrictive temperature (36 °C for *mis6-302* and *cnp1-1* and 18 °C for *dis3-54* cells). *act1+* was used as a loading control. *: signal from rRNA (adapted from Choi et al. 2011).
it deposits histone H3.3 that in turn acts as a “temporary placeholder” for CENP-A assembly at centromeres (Dunleavy et al., 2011). The disassembly and re-deposition mechanisms require several chromatin-remodelling factors. The histone chaperones Spt6 and the FACT complex by physically interacting with histone H3 and H2A/H2B respectively, are involved in the reassembly mechanism (Belotserkovskaya et al., 2003; Kaplan et al., 2000). Recently, the FACT complex was found to localise at centromeres in a CENP-H dependent-manner and to physically interact with CENP-A nucleosomes (Foltz et al., 2006; Okada et al., 2009). In addition, deletion of one of the FACT complex subunits decreases the efficiency of incorporation of newly synthesized CENP-A at human centromeres (Foltz et al., 2006). In contrast, FACT has little or no effect on CENP-A\textsuperscript{Cnp1} assembly at endogenous centromeres in fission yeast. Indeed, mutations of the FACT complex impair the reassembly of H3 chromatin on transcribed regions and as result, allow CENP-A\textsuperscript{Cnp1} nucleosomes to be deposited in its place (Choi et al., 2012).

In human cells, the FACT complex cooperates with the ATP-dependent chromatin remodelling factor CHD1 for the maintenance of the level of CENP-A containing nucleosomes at centromeres (Okada et al., 2009). The fission yeast homologue of CHD1, Hrp1, is involved in disassembling histone H3 nucleosomes at the transcriptional start sites of several genes. Moreover, Hrp1 is required for maintaining the normal level of centromeric CENP-A\textsuperscript{Cnp1} at centromeres (Walfridsson et al., 2007; 2005). Interestingly, it seems that CENP-A tends to assemble at euchromatic promoters where histone H3 is removed by Hrp1, this suggests a role for transcription in driving the exchange of histone H3 with CENP-A containing nucleosomes (Choi et al., 2011). Therefore, it is possible that nucleosome disassembly associated with transcription facilitates the replacement of canonical histone H3 by histone H3 variants, such as CENP-A.

### 1.7 Aims of the project

In many organisms, it is apparent that the DNA sequence does not specify centromere location and instead epigenetic mechanisms play a major role. However, CENP-A, and thus kinetochore assembly, in most organisms generally occurs on certain preferred sequences. The analysis presented in this thesis focuses on the role of the DNA sequence in the selection of the locus where CENP-A chromatin is formed and in particular whether there are any particular DNA features that promote CENP-A\textsuperscript{Cnp1} assembly at fission yeast centromeres.
**Aim1: Dissecting minimal central core DNA**

It is known that CENP-A$_{\text{Cnp1}}$ is specifically enriched over the central core domain of fission yeast centromeres. The central core sequence of centromere 2 is 8.5 kb in length. Previous studies have shown that not all of the cc2 sequence is required for centromere function since it shows redundancy in the function (Baum et al., 1994). Therefore, studying a smaller fragment of this sequence may facilitate the detection of what is strictly required for centromere specification. Therefore, in order to identify what features enable centromeric sequences to assemble CENP-A chromatin, I set out to define a minimal central core sequence capable of allowing CENP-A$_{\text{Cnp1}}$ nucleosomes assembly in fission yeast.

**Aim2: Investigating sequence specific requirements**

Once the minimal sequence has been identified, the possible characteristics that distinguish this minimal sequence will be studied. Centromere sequences may possess many different features that promote kinetochore assembly and centromere function. For instance, centromere sequences are highly bendable and this property may facilitate a preferential high-order structure required for centromere function and that the sequence itself may influence CENP-A$_{\text{Cnp1}}$ nucleosome deposition. The A/T nucleotide composition, which influences the structural property of the sequence, will be tested for its ability to allow CENP-A$_{\text{Cnp1}}$ nucleosome assembly. In addition, it is possible that specific DNA sequences in the central domain are required to directly bind specific centromere proteins that mediate CENP-A deposition. Since several centromere associated proteins contain known or predicted DNA binding motifs, their ability to recognise central domain sequence *in vivo*, prior to CENP-A assembly, is tested. Furthermore, transcription may be involved in CENP-A$_{\text{Cnp1}}$ deposition and maintenance of the centromeric structure. Therefore, I aim to determine whether the minimal sequence possesses a transcription unit and whether transcription correlates with CENP-A$_{\text{Cnp1}}$ deposition.
2 MATERIALS AND METHODS

2.1 Fission Yeast Protocols

2.1.1 Mating and random spore analysis
Crosses were carried out on ME medium in order to nitrogen starve the cells and induce sporulation. A similar amount of cells from two strains of opposite mating types (h+/-h-) were mixed together and incubated for 2 days at 32°C (or 25°C for temperature sensitive strains). The presence of ascii containing four spores was assessed by light microscopy. Cells were then resuspended in 300µl of 1:10 diluted glusulase (NEN) and incubated overnight at 37°C (25°C for 2 day for temperature sensitive). Glusulase digests the asc wall and vegetative cells so that only the spores remain alive. 10 ml of water was then added and 1µl and 10µl of the diluted spores plated on selective media and grown at appropriate temperature.

2.1.2 S. pombe transformations through electroporation
A 50ml culture of cells in log phase (5x10^6 to 1x10^7 cells/ml) was harvested at 3500rpm for 2 minutes. The pellet was washed three times in 10ml of 1.2M ice-cold sorbitol and then resuspended in 500-1000 µl of 1.2M ice-cold sorbitol. 200µl of cells were mixed with between 100ng (plasmids) and 10µg (linear fragments) of DNA in an ice-cold cuvette. Cells were pulsed using a Bio-Rad Gene Pulser II at a setting of 2.25kV, 200Ω and 25µF. Immediately following pulsing, 400µl of 1.2M ice-cold sorbitol was added. Cells were spread onto selective media using sterile glass beads and incubated at appropriate temperature.

2.1.3 Spotting assay
To assay the growth of different S. pombe strains on different media cells were dissolved in 200 µl of water and serial dilutions of 1:4 were made in sterile microtitre plates in dH2O starting with ~5x10^6/ml cells and 5µl of each plated on the appropriate media. Cells were then incubated at the desired temperature until colonies were formed.
2.1.4 **Centromere silencing assay**

Wild type cells which have the *ade6+* gene inserted into centromeric outer repeats are red when grown under restricted adenine conditions. This is due to transcriptional repression which causes the accumulation of amino-imidazole ribonucleotide (AIR). Mutants which alleviate silencing at the outer repeats are white due to alleviation of silencing. This assay can also be carried out using the insertion of the *ura4+* gene inserted at the *imr* or *arg3* inserted at the TM region of *cc1*. In wild type cells, *ura4+* expression is repressed and cells are able to grow well on counter-selective media containing FOA. Mutant cells grow well on media lacking uracil but are unable to grow on media containing FOA.

2.1.5 **Minichromosome transformation**

Minichromosomes were transformed by electroporation. Transformants were selected by growth on PMG –ura –ade at 32°C.

For the experiments on H3K9R strain, cells were incubated with pre-transformation solution for 15min at 32°C prior of the transformation protocol. pMC1 based minichromosomes were transformed in H3K9R strain and wild-type strains and 5 independent colonies from each transformation were analysed for the presence of CENP-A*^Cnp1* by chromatin immunoprecipitation (ChIP).

**Pre-transformation solution:** 20mM HEPES, 0.6M sorbitol, 25mM DTT

2.1.6 **Establishment assay**

For establishment assay, cells were transformed with pMC21-based minichromosomes (containing 2.1 kb of *otr*). The untransformed wild type with the nonsense mutation on the *ade6* gene (*ade6*-704) plated on 1/10<sup>th</sup> adenine medium forms red colonies while an *ade6*<sup>+</sup> strain forms white colonies. The minichromosome utilised here possesses a *sup3e* tRNA, a nonsense suppressor able to complement the mutation *ade6*-704. Therefore, a plasmid that assembles a functional centromere is retained in cells and forms white colonies. Cells were transformed by electroporation with 200 µg of miniprep DNA and plated on –ade-ura plates to select for the minichromosome and incubated at 32°C for 5 days. The colonies where then replicated into rich medium with limiting amount of adenine. Presence of white colonies indicate that a functional centromere is established and the minichromosome in non selective medium. The percentage of white colonies over the total amount of transformants represent the efficiency of the plasmid to establish functional centromere.
2.1.7 Minichromosome loss assay
Cell that have a functional centromere established in the minichromosome, were grown –ade-ura plates o/n at 32°C and consequently plated on rich medium with limiting amount of adenine. The percentage of red sectors within the white colonies represents an estimate of mitotic stability of the plasmid. In contrast, a mitotically unstable plasmid is rapidly lost in non-selective rich medium and the colonies appear red coloured.

2.1.8 Media and Growth
*S. pombe* cultures and colonies were incubated at temperatures between 18°C and 36°C for between overnight and 3 days as indicated for each experiment. Haploid strains will 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>3 hours</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>2 hours 10 minutes</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>2 hours</td>
</tr>
<tr>
<td>PMG minimal</td>
<td>25</td>
<td>4 hours</td>
</tr>
<tr>
<td></td>
<td>32</td>
<td>2 hours 30 minutes</td>
</tr>
<tr>
<td></td>
<td>36</td>
<td>2 hours 20 minutes</td>
</tr>
</tbody>
</table>

*Growth Media*
All solutions were made up to a final volume in dH₂O and autoclaved unless otherwise stated.

**PMG agar in 900ml:**
- Pthallic acid: 3g
- di-sodium orthophosphate: 2.2g
- glutamic acid: 3.75g
- D-glucose anhydrous: 20g
- vitamins 1000x: 1ml
- minerals 10,000x: 0.1ml
- salts 50x: 20ml
- agar (OXOID): 20g
Chapter 2: Materials and Methods

PMG liquid in 900ml:

- Phthallic acid 3g
- di-sodium orthophosphate 2.2g
- glutamic acid 3.75g
- D-glucose anhydrous 20g
- vitamins 1000x 1ml
- minerals 10,000x 0.1ml
- salts 50x 20ml

YES agar (-ade):

- Yeast extract (DIFCO) 5g
- D-glucose anhydrous 30g
- Arginine (Sigma) 0.2g
- Lysine (Sigma) 0.2g
- Histidine (Sigma) 0.2g
- Uracil (Sigma) 0.2g
- Leucine (Sigma) 0.2g
- Agar (OXOID) 20g

YES liquid:

- Yeast extract (DIFCO) 5g
- D-glucose anhydrous 30g
- Arginine (Sigma) 0.2g
- Lysine (Sigma) 0.2g
- Histidine (Sigma) 0.2g
- Uracil (Sigma) 0.2g
- Leucine (Sigma) 0.2g

ME plates (1L):

- Malt extract (OXOID) 30g/L
- Adenine (Sigma) 250g/L
- Arginine (Sigma) 250g/L
- Histidine (Sigma) 250g/L
- Uracil (Sigma) 250g/L
- Leucine (Sigma) 250g/L

Vitamins 1000x (100ml):

- Pantothenic acid 0.5g
- Nicotinic acid 1g
- Inositol 1g
Chapter 2: Materials and Methods

Biotin 1mg
Filter sterilised

**Minerals 10,000x (100ml):**

- Boric acid 5g
- MnSO₄ 4g
- ZnSO₄ 4g
- FeCl₂·6H₂O 2g
- Molybdc acid 1.6g
- CuSO₄·5H₂O 0.4g
- Citric acid 10g
Filter sterilised

**Salts 50x:**

- Magnesium chloride 53.5g
- Calcium chloride 1g
- Potassium chloride 50g
- di-sodium sulphate 2g

**Supplement stocks:**

- Adenine 50x (Sigma) 5g/L
- Arginine 100x (Sigma) 10g/L
- Histidine 100x (Sigma) 10g/L
- Uracil 20x (Sigma) 2g/L
- Leucine 100x (Sigma) 10g/L

**Additional supplements:**

- Fluoroorotic acid (FOA) (Melford Laboratories): final concentration of 0.5g/500ml (1x)
- Thiabendazole (TBZ) (Sigma): final concentrations of 10µg/ml or 20µg/ml in DMSO.
- Nourseothricin (cloNAT) (Werner BioAgents): final concentration of 2000x
- Geneticin (G418) (Gibco): final concentration of 0.1mg/ml.

### 2.2 DNA protocols

#### 2.2.1 Genomic DNA Isolation from fission yeast
A 5ml stationary phase culture was harvested at 3000rpm for 5 minutes. The pellet was resuspended in 250µl SP1 buffer containing 0.4mg/ml zymolyase 100-T (MP Biomedicals) and incubated for 30 to 60 minutes at 37°C. The cells were then pelleted at 13000rpm in a microfuge for 15 seconds and the pelleted resuspended in 0.5ml TE, 50µl 10% SDS and vortexed. 165µl 5M potassium acetate was then added and the samples incubated on ice for 30 minutes. After centrifugation at 13000rpm at 4°C for 10 minutes, the supernatant was added to 0.75ml isopropanol, incubated on dry ice for 5 minutes and then centrifuged as before. The pellet was resuspended in 0.3ml TE containing 10μg/ml of RNase A (Roche). After 30 minutes at 37°C the sample was then extracted with phenol/chloroform and precipitated by addition of 2-3 volumes of ethanol and 1/10 volume of 3M sodium acetate. The pellet was resuspended in 20µl TE and stored at -20°C.

**SP1: 1.2M sorbitol, 50mM sodium citrate, 50mM sodium phosphate, 40mM EDTA, pH 5.6**

2.2.2  *Yeast genomic DNA isolation for colony-PCR*

A small amount of a single colony of *S. pombe* was suspended in 10µl SPZ buffer and incubated at 37°C for 10 minutes. 1 µl of the genomic DNA was used as template for the PCR reaction.

**SPZ: 1.2M sorbitol, 100mM sodium phosphate, 2.5mg/ml zymolyase-100T (MP Biomedical)**

2.2.3  *Agarose gel electrophoresis*

Agarose gel electrophoresis was used to analyse the size of DNA fragments. Agarose was dissolved in 1 x TBE buffer by heating in a microwave. Once cooled, ethidium bromide (Sigma) was added to a concentration of 0.03µg/ml. DNA samples were loaded in Orange G loading buffer and visualised under a UV transilluminator.

**10xTBE:** 108g Trizma base, 55g boric acid, 9.3g EDTA

**6x Loading buffer:** 30% Glycerol, 0.25% Orange G

2.2.4  *Polymerase Chain Reaction (PCR)*

PCR reactions were carried out as follows in 0.2µl thin walled PCR tubes: template DNA, 10pM primer, 2.5mM dNTPs, 10 x PCR buffer, 0.5U Taq (Roche), dH2O.
When precise amplification was required, in the case of cloning, Platinum Pfx taq polymerase from Invitrogen was used as per manufacturers instructions. The following programs were used as indicated.

**Ura program**: 94°C for 4 minutes, (94°C for 30 seconds, 55°C for 30 seconds, 72°C for 1 minute), 29 cycles, 72°C for 5 minute.

**Bahlong**: 96°C for 5 minutes, (94°C for 1 minute, 55°C for 1 minute, 72°C for 4 minutes), 34 cycles, 72°C for 10 minutes.

**Bahvlong**: 96°C for 5 minutes, (94°C for 1 minute, 55°C for 1 minute, 68°C for 6 minutes), 34 cycles, 68°C for 10 minutes.

### 2.2.5 Fusion PCR

Standard PCR reaction was set up for each fragment (program PFX): template DNA, 10mM primer 3 μl/each, Pfx buffer 12μl, 50mM MgSO₄ 1μl, 10mM dNTP mix 1.8μl, Pfx polymerase 0.3μl, H₂O to 50μl. The PCR products were then run on agarose gel and purified with Gel-purification Kit (Qiagen) according to manufacturer instructions. The gel-purified PCR fragments were mixed and used as template for the second reaction (Fusion-2): 200ng of template (1:3:1 mixture), 5μl Pfx buffer, 2μl dNTP mix (2.5mM), 0.5μl MgSO₄ (50mM), 0.2μl Pfx polymerase, H₂O to 25μl. The final PCR amplify fusion PCR with nested primers in large scale. 2 μl fusion PCR products (obtained from Fusion-2), 1μl/each nested PCR primer (100μM), 40 μl Pfx buffer, 4μl MgSO₄ (50mM), 6μl dNTP mix (2.5mM), 2μl Pfx polymerase, H₂O to 100μl. This final PCR product was gel purified using gel-purification kit (Qiagen) and eluted in 30μl of water. 10μl were used for transformation of fission yeast by electroporation.

**PFX**: 94°C-5min, (94°C-30sec, 55°C-30sec, 68°C-min/kb) 35 cycles, 68°C-10min  
**Fusion-2**: 94°C-2 min , (94°C-30sec, 58°C-10 min, 68°C-5min) 15 cycles, 68°C-10min  
**Fusion-3**: 94°C-5min , (94°C-1 min, 55°C-1 min, 68°C-6min) 35 cycles, 68°C-10min

### 2.2.6 Minichromosome construction

In this study, minichromosomes were created by PCR amplification of the fragment of interest from central core 2 sequence using oligos with 5’ overhanging BamHI and 3’ BglII site. The sequence was inserted into BamHI/BglII digested pMC1 or pMC21 vector backbone. The minichromosomes contains a *URA* marker that complement
the ura-auxotrophy and a sup3e tRNA, a non sense suppressor able to complement the genomic mutation ade6-704. It is also present a Kanamycin resistance gene so that the minichromosomes can be selected in G418 plates.

Plasmids without heterochromatin were based on the pMC1 plasmid and were used for the analysis of the minimal centromere sequence in H3K9R strain and analysis of transcripts in wild type strain. For the analysis of the stability in wild type strain, the fragments were cloned on pMC21 vector. pMC21 possesses 2.1 kb region of outer repeats that is sufficient to establish an heterochromatic domain in wild type strain.

2.2.7 Sequencing
Reactions were set up as follows: 2µl ABI Prism BigDye Terminator Cycle Sequencing Ready Reaction Kit v 3.0 (Applied Biosystems), 3.2pmol/µl primer, template DNA as recommend by manufacturers (1-1000ng for PCR products, 200-500ng for dsDNA) and dH2O up to 20µl. Reactions were run on the following program in 0.2µl thin walled PCR tubes in a PTC-225 thermal cycler (MJ Research): 95°C for 5 minutes, (95°C for 30 seconds, ramp 1°C per second to 55°C, 55°C for 15 seconds, ramp 1°C per second to 64°C, 64°C for 4 minutes), perform 25 cycles. Samples were then sent to the central sequencing service for analysis.

2.2.8 Southern analysis
Genomic DNA samples were run on agarose gel. The gel was incubated in depurinating solution for 10 min, rinsed twice with dH2O and then incubated in
denaturing solution for 15 min. The denaturing solution was changed and the gel incubated for further 15 min.

Gels were blotted by capillary transfer for at least 16 hours in denaturing solution using Hybond-NX (Amersham). The membrane was incubated in 50mM phosphate pH7.2 at least for 10min and then crosslinked twice at 1200 joules in a UV crosslinker.

Hybridisation and probe preparation

Probes were created using PCR products as substrate. 50-100ng of DNA were diluted to 13μl with dH2O and boiled for 10 min. The DNA was placed on ice for few minutes, 3 μl of 32PdCTP and 4 μl of High Prime solution (Roche) added. The random labelling was performed at 37°C for 30 min. To stop the reaction, 80 μl of 25mM EDTA were added.

The membrane was pre-hybridised with warm Church buffer at 65°C for 1hour. When ready, the probe was boiled for 5 min and added to fresh Church buffer. The membrane was hybridised with the probe at 65°C overnight and subsequently washed twice for 10 min with washing buffer.

The signals were visualised after between 4 hours and overnight using a phosphoscreen using a Storm phosphoimager with ImageQuant TL v 2005 (Amersham).

Depurinating solution: 12.5ml HCl in 500ml dH2O
Denaturing solution: 0.5M NaOH, 1.5M NaCl
1M Na phosphate: 89g Na2HPO4 pH to 7.2 with orthophosphoric acid
Church Buffer: 0.5M Na phosphate pH7.2, 7% SDS, 1mM EDTA, 1% BSA
Washing: 40mM Na phosphate pH7.2, 1mM EDTA, 1% SDS

2.3 RNA protocols

2.3.1 Total RNA isolation

RNA was extracted using a Qiagen RNeasy Miniprep or Midiprep kit as per manufacturer’s instructions. RNA was quantified using a Nanodrop spectrophotometer.
2.3.2 Poly(A) RNA isolation

Poly(A) RNA was isolated from 500µg total RNA prepared using a Qiagen Midiprep Kit. Poly(A) RNA was isolated using PolyATtract mRNA Isolation System IV (Promega). RNA was quantified using a Nanodrop spectrophotometer.

2.3.3 Northern Blotting

GEL preparation

Cells transformed with circular minichromosome were grown on selective media and total RNA was isolated by RNaseasy Midi kit (Qiagen). To 10 µg of total RNA were added 3 volumes of loading buffer and then incubated 10min at 65°C. The samples were cooled on ice 5-10min before running on a 1% agarose gel in 1X HEPES buffer and blotted on a nylon membrane. Gels were blotted by capillary transfer for at least 16 hours in 20 x SSC using Hybond-NX (Amersham). The membrane was washed in 6XSSC for 10min and then crosslinked twice at 1200 joules in a UV crosslinker (Stratagene)

10x HEPES: (1l) 47.66g HEPES, 3.2g EDTA, 9.5g KOH.

Gel: (100ml) 1g agarose, 73.8 ml dH₂O, 10 ml 10xHEPES, 16.2 ml 37%Formaldehyde, formamidine

Loading buffer: (1ml) 100µl 10X HEPES, 500 µl deionised formamide, 60 µl BromoPh blue, 13 µl Ethidium Bromide 10mg/ml, 160 µl 37%Formaldehyde, 165 µl dH₂O.

Hybridisation and probe preparation

Membranes were pre-hybridised for 1 hour at 68°C in pre-warmed Church buffer. In the meantime, RNA probes were in vitro transcribed from 0.5 µg of purified PCR product using the MaxiScript T7 Kit (Ambion, Inc) and labelled with α³²P-UTP according to the manufacturer instruction. The reaction was incubated for 1h at 37 °C and unincorporated radionucleotide was removed using a NucAway Spin column (Ambion). Membranes were hybridized overnight at 68°C and washed 2x30min with 1%SDS-2X SSC followed by 2x15min washes with 1%SDS-0.5X SSC at 68°C. A phosphoscreen was exposed to the membrane overnight and then analysed on a Storm phosphorimager with ImageQuant TL v 2005 (Amersham).

Church buffer: 0.5M sodium phosphate pH 7.2, 1mM EDTA, 7% SDS
2.3.4 RT-PCR

10µg of total RNA was incubated with 1µl Turbo DNase (Ambion) in a final volume of 50µl for 1 h at 37°C. The DNase treated RNA was then purified using Qiagen RNeasy Miniprep (clean-up protocol) and quantified. RT-PCR was performed using Superscript III Reverse Transcriptase (Invitrogen) according to the manufacturer instruction. 1µg of total RNA or 100 ng poly(A) purified RNA was incubated with 100ng random hexamers, 1µl 10mM dNTP Mix (10mM each) in a final volume of 13µl and incubated at 65°C for 10 minutes and then placed on ice. For first strand synthesis, the following reagents were added; 4µl 1st strand buffer, 1 µl 0.1M DTT and 1µl Superscript III. The enzyme was not added to the –RT control. Samples were then incubated at 25°C for 5 minutes and 50°C for 60 minutes and then 70°C for 15 minutes to inactivate the enzyme. 10µl of water was added and 1µl of this diluted cDNA was used as a template in a 20µl PCR reaction.

5'RACE-PCR

In order to detect the TSSs, 5'RACE PCR was performed using FirstChoice RLM-RACE kit (Ambion) according to the manufacturer protocol with the following modifications. 1 µg Poly(A) purified RNA was treated with Calf-Alchaline Phosphatase (CIP) and then purified using Qiagen RNeasy Miniprep (clean-up protocol). 70 ng of RNA were treated with Tobacco Acid Pyrophosphatase (TAP) to remove the cap structure from mRNA leaving a 5’-monophosphate. A 45 bp RNA adapter was ligated to the RNA population using T4 RNA ligase. The adapter cannot ligate molecules that do not contain the 5’-monophosphate. RT reaction was performed as previously described using Superscript III Reverse Transcriptase (Invitrogen). A nested PCR reaction was performed using combinations of primers corresponding to the Adapter sequence and to the LM region and using standard Taq polymerase (Roche). The oligonucleotides employed were spaced approximately every 100 bp to cover the entire LM sequence. The PCR products were run on a agarose gel and the different bands extracted from gel. The fragment were then cloned into pGMT-easy vector system (promega) and sequenced (T7 primer fw) to detect the TSS.

2.4 Protein Protocols

2.4.1 S. pombe protein extraction
A 50ml culture was grown to log phase and harvested at 3000rpm at 4°C. The sample was then frozen in liquid nitrogen. Cells were then resuspended in 500µl sample buffer/1mM PMSF and 500µl acid-washed glass beads (Sigma). Cells were lysed by bead-beating for 3 minutes. Samples were boiled for 5 minutes at 95°C then spun in a microfuge for 30 seconds to remove beads and cell debris.

2x Sample buffer: 2% SDS, 50mM Tris-HCl pH6.8, 2mM EDTA, 10% glycerol, 0.03% bromophenol blue, 2% β-mercaptoethanol

2.4.2 Western analysis
Samples were run on NuPAGE Bis-Tris Mini Gels (life technologies) in 1xNuPAGE MOPS buffer (life technologies) at 200V constant. Proteins were transferred onto Protran nitrocellulose membrane (Schleicher and Schuell) using XCell Blot Module (Invitrogen). The blotting apparatus was assembled in transfer buffer with 2 blotting pads, 2 pieces of 3MM paper, the gel, the membrane, 2 pieces of 3MM paper and 2 blotting pads. Transfer was carried out for 1-2 hours at 30V constant. After transfer, the membrane was washed briefly in deH₂O and stained with Ponceau S (Sigma). The membrane was washed in PBS and placed in blocking buffer (5% milk in PBS + 0.1% Tween-20) for 1 hour at room temperature. Primary antibody was then added at appropriate concentration in 3% milk in PBS + 0.1% Tween-20 and incubated overnight at 4°C. Membranes were then washed twice in PBS + 0.1% Tween 20 for 15 minutes and then the secondary HRP-conjugated antibody was added in blocking buffer. The secondary antibody was incubated for less than 1 hour at room temperature. The membrane was washed again in twice in PBS + 0.1% Tween-20 for 15 min and then rinsed briefly in PBS. Proteins were detected using an Enhanced Chemi-Luminescence kit (Amersham) as per manufacturer's instructions. The blot was exposed to Bio-Max Light film (Kodak).

Transfer buffer: 1X NuPAGE Transfer Buffer (life technologies), 10% MetOH (1 gel) or 20% (2 gels).

2.4.3 Chromatin Immunoprecipitation (ChIP)
Cells were grown on selective medium at 32°C to log phase and fixed in 1% PFA for 15 min at room temperature. The pellet was washed twice with ice-cold PBS and then froze in dry ice.
When minichromosomes were assayed, possible integration of the plasmids was assessed by plating part of the culture before fixation in rich medium and limiting adenine concentration. All the strains used in this study contain a mutation on the ade6 gene (ade6-704) and show red coloured colonies on 1/10th adenine plates. The sup3e present in the minichromosome is able to complement the mutation. Therefore, the episomal plasmid is lost in rich medium giving red colonies. If integration has occurred, the colonies will show a white colour.

Frozen pellets were resuspended in 350 µl ice-cold lysis buffer containing 3.5µl protease inhibitors (1:100) and 1mM PMSF. Cells were mechanically lysed using acid-washed glass beads (Sigma) and the chromatin sheared by sonication (3 x 5min, 30 s on/off cycle) using a BioRuptor water bath. Samples were then pelleted in a microfuge at 13,000rpm for a total of 20 mins at 4°C and the supernatant transferred to a fresh tube. Protein A or G were washed 3 times with cold lysis buffer at 2000rpm. The lysate was pre-cleared with 25 µl of washed Protein A or Protein G for 1hour at 4°C. The final samples was then transferred into a clean tube and 10 µl retained as the crude control. The appropriate antibodies was then incubated overnight at 4°C together with 25 µl of Protein A or G. Beads were centrifuged at 2000rpm for 2minutes in a microfuge and washed in lysis buffer (rinse), lysis buffer + 0.5 M NaCl (10min), wash buffer (10min), TE pH8 (rinse).

A 10% slurry of Chelex-100 resin (Bio-Rad) was freshly prepared in sterile water and 100µl of 10% Chelex added to each IP sample (beads) and Input sample (10 µl WCE). Samples were boiled at 100°C for 12 min and cooled by quick spin. 2.5 µl of 10mg/ml Proteinase K was added to each sample and incubate at 55°C for 30min. Proteinase K was inactivated by boiling the samples at 100 °C for 10 min. The samples were then transferred to a fresh tube.

Samples were then analysed by gel or real time PCR.

**Lysis buffer:** 50mM HEPES-KOH pH7.5, 140mM NaCl, 1mM EDTA, 1% Triton-X100, 0.1% sodium deoxycholate

**Wash buffer:** 10mM Tris-HCl pH8, 0.25M lithium chloride, 0.5% NP-40, 1mM EDTA, 0.5% sodium deoxycholate

**TE:** 10mM Tris-HCl pH8, 1mM EDTA

**TES:** 50mM Tris-HCl pH8, 10mM EDTA, 1%SDS

**Gel quantification**

The quantitative analysis on a gel was performed on sample run on 1.5% TBE-agarose gel using Image J software. Each sample was diluted 1:30 in TE and 2 ul
used in duplex PCR. The CENP-A ChIP was monitored by assessing enrichment of endogenous cc1/3 sequences. The enrichment of cc1/3 or minichromosome cc2 bands was calculated relative to the control fbp1 band. Specific oligonucleotides were used to distinguish endogenous cc2 from the one on the minichromosome. The strains possess a silent his3+ gene inserted at the SphI site in the endogenous cc2 so that specific primers were used to amplify the region on the minichromosome. For each experiment 5 independent transformants were analysed by ChIP.

**Real-time PCR (qPCR)**

The quantitative PCR reactions were carried out in 10 µl volume, with 5µl Light Cycler 480 SybrGreen Master Mix (Roche), 0.05µl/each primers and 3µl ChIP or total template. The ChIP samples were diluted in water 1:20 and Total/input sample 1:60. The data were analysed using Light Cycler 480 Software 1.5 (Roche)

The quality of CENP-A ChIP was monitored by assessing enrichment of endogenous cc1/3 sequences and on the negative region act1+.

**Program:** 95°C for 2 minutes, (95°C for 20 seconds, 55°C for 20 seconds, 72°C for 20 seconds), 55 cycles, Melting curve.

### 2.5 ENZYMATIC ASSAY

#### 2.5.1 Spotting assay for β-galactosidase

Serial dilution of yeast were spotted on YES medium covered by Protran nitrocellulose membrane (Schleicher and Schuell) and incubated at 32°C. When colonies are formed, remove the membrane with the yeast from the plate and freeze it on liquid nitrogen. Lay the membrane on two pieces of 3MM paper previously soaked in incubation buffer and put on an empty petri dish (remove the excess of buffer before placing the membrane). Take care not move too much the membrane with the yeast lysate. Close the petri dish with parafilm and incubate for 3h at 37°C.

**Z buffer** 0.06M Na2HPO4, 0.04M NaH2PO4, 0.01M KCl, 0.001M MgSO4.

**Incubation buffer** 5ml Z-buffer, 78µl 2% X-Gal, 50µl 1M DTT.

#### 2.5.2 Liquid assay for β-galactosidase

Yeast containing the vector expressing the LacZ gene, were grown on selective
medium till the OD600 reached 0.5. The OD600 was recorded. Cells were pelleted for 2 min at 3500g and then washed resuspended in 1ml of cold Z buffer. To permeabilise the cells, 1-2 drops of 0.1% SDS and 1-2 drops of chloroform were added with a Pasteur pipette and the cells were vortexed vigorously for 15-30sec. The extract was equilibrated for 5 min at 30°C and 200µl of ONPG was added to each tube and the time of addition recorded. The reaction run until the solution had turned yellow. High activity constructs (e.g. nmt41/81 promoters) changed colour rapidly. The reaction was stopped with 0.5 ml of 1M Na2CO3. The time of the reaction was recorded and the elapsed time calculated in minutes. Cell debris were spun out and the OD420 of the supernatant red.

"Miller units" were calculated as follow

\[
\text{Units} = 1000 \times \frac{\text{OD}_{420}}{\text{Volume} \times \text{Time} \times \text{OD}_{595}}
\]

\[
\text{Volume} = \text{volume cells assayed (1 ml)}
\]

\[
\text{Time} = \text{time of reaction (min)}
\]

\[Z\text{ buffer} \quad 0.06 \text{M Na2HPO4, 0.04M NaH2PO4, 0.01M KCl, 0.001M MgSO4. Before use, add fresh 0.03M beta-mercaptoethanol (0.27ml BME / 100ml Z buffer)}
\]

\[1\text{M Na2CO3 store at room temperature indefinitely}
\]

\[\text{ONPG (2-Nitrophenyl-β-D-galactopyranoside), 4mg / ml in water. Filter, store at 4°C}
\]

2.6 Bacterial Protocols

2.6.1 Competent cell

One single E. coli colony from a freshly streaked plate was inoculated into 5 ml of LB and grown O/N at 37°C. The culture was diluted 1:200 into pre-warmed 100 ml LB + 20 mM MgSO4. Bacteria were grown at 37°C with shaking until OD600 reached 0.48 (approximately 3-4 h) and then transferred to chilled centrifuge bottles and incubated on ice for 10 min. Cells were pelleted by centrifugation for 5 min at 5K rpm, 4°C and resuspended by gently pipetting up and down in 40 ml cold TFB1 buffer per 100 ml culture. Bacteria were incubated on ice for 5 min and then pelleted for 10 min at 3K rpm, 4°C. Cells were resuspended in 4 ml cold TFB2 per 100 ml culture and incubated on ice for 15 min.

The competent cells were dispensed into 100 µl aliquots into pre-chilled eppendorf tubes and stored at -80°C.
Reagents required

LB and 1 M MgSO₄

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Final concentration</th>
<th>Autoclaved stock</th>
<th>for 500 ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 mM KAc</td>
<td>0.3 M</td>
<td>50 ml</td>
<td></td>
</tr>
<tr>
<td>100 mM RuCl₂</td>
<td>1 M</td>
<td>50 ml</td>
<td></td>
</tr>
<tr>
<td>10 mM CaCl₂</td>
<td>1 M</td>
<td>5 ml</td>
<td></td>
</tr>
<tr>
<td>50 mM MnCl₂</td>
<td>1 M</td>
<td>25 ml</td>
<td></td>
</tr>
<tr>
<td>15% glycerol</td>
<td>50%</td>
<td>150 ml</td>
<td></td>
</tr>
</tbody>
</table>

pH to 5.8 with dilute HAc. Filter sterilise and store at 4°C.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Final concentration</th>
<th>Autoclaved stock</th>
<th>for 100 ml</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 mM MOPS</td>
<td>1 M, pH 6.5</td>
<td>1 ml</td>
<td></td>
</tr>
<tr>
<td>75 mM CaCl₂</td>
<td>1 M</td>
<td>7.5 ml</td>
<td></td>
</tr>
<tr>
<td>10 mM RuCl₂</td>
<td>1 M</td>
<td>1 ml</td>
<td></td>
</tr>
<tr>
<td>15% glycerol</td>
<td>50%</td>
<td>30 ml</td>
<td></td>
</tr>
</tbody>
</table>

Readjust pH to 6.5 with KOH. Filter sterilise and store at 4°C.

Note: Filter sterilise MOPS stock buffer.

2.6.2 Transformation

For most of the cloning performed, GT116 cells were used in order to decrease the probability of recombination within the bacteria. In brief, 40 µl of cells were mixed with DNA and incubated on ice for 30min. Heatshock was performed at 42°C for 45sec followed by incubation of the cells on ice for 1min and addition of 700 µl of LB. Cells were then grown at 37°C for 45 min and plated onto appropriate selective media.

2.6.3 Plasmid miniprep

Single bacterial colonies were grown in 5ml LB plus appropriate supplement overnight at 37°C. Cells were harvested and plasmid prep performed using a QIAGEN miniprep kit according to manufacturer’s instructions.
2.6.4 *Bacteria Media*

**LB medium per litre:**
- Bacto tryptone: 10g
- Bacto yeast extract: 5g
- Sodium chloride: 10g

**LB agar per litre:**
- Bacto tryptone: 10g
- Bacto yeast extract: 5g
- Sodium chloride: 10g
- Bacto agar: 15g

**Antibiotics:**
- **Ampicillin** 100mg/ml
- **Carbenicillin** 50mg/ml

2.7 Antibodies

2.7.1 ChIP
- sheep anti-Cnp1 1:30
- sheep anti-Cnp3 1:30
- rabbit anti-H3 (abcam ab1791) 1:150
- rabbit anti-GFP (Invitrogen- A11122)- 1:200

2.7.2 Western Analysis
- mouse anti-CTD (cell signalling- 4H8) 1:2000
- mouse anti-myc (9E10) 1:1500
2.8 Primers used in this thesis

### Cloning of cc2 sequence (Chapter 3):

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5'-3'</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>cc2-F1 (J)</td>
<td>tactacgatccccccatggaaatat</td>
<td></td>
</tr>
<tr>
<td>cc2-R1 (J)</td>
<td>tactacgtcatctcggagatctcagtgctgttctgtga</td>
<td></td>
</tr>
<tr>
<td>cc2-F2 (K)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-R2 (K)</td>
<td>tactacgtcatctcggagatctcagtgctgttctgtga</td>
<td></td>
</tr>
<tr>
<td>cc2-F3 (L)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-R3 (L)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-F4 (M)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-R4 (M)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-F5 (N)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-R5 (N)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-F6 (O)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-R6 (O)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-F7 (P)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-R7 (P)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-F8 (Q)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-R8 (Q)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-F9 (R)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
<tr>
<td>cc2-R9 (R)</td>
<td>tactacgatccgagactgcttcatcatcttca</td>
<td></td>
</tr>
</tbody>
</table>

### Oligonucleotides used in the analysis of the CENP-A levels on the minimal sequence (Chapter 3)

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5'-3'</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>cc1/3-F</td>
<td>aacaataacagcaatgcctc</td>
<td></td>
</tr>
<tr>
<td>cc1/3-R</td>
<td>atagtacatcgcgattgtct</td>
<td></td>
</tr>
<tr>
<td>fbp-A</td>
<td>aatgacaattcccccactta</td>
<td></td>
</tr>
<tr>
<td>fbp-B</td>
<td>acttcacgtagctcagac</td>
<td></td>
</tr>
<tr>
<td>WA293</td>
<td>caaacctactaatttccttcc</td>
<td></td>
</tr>
<tr>
<td>WA295</td>
<td>ccaactctacttactactacctcc</td>
<td></td>
</tr>
<tr>
<td>WA292</td>
<td>cgtctacatgataactttcc</td>
<td></td>
</tr>
<tr>
<td>WA530</td>
<td>aggccagttcgagaaacag</td>
<td></td>
</tr>
<tr>
<td>M13F</td>
<td>cggccaggtttcccagcttcg</td>
<td></td>
</tr>
<tr>
<td>WA171</td>
<td>aggacaaaatagctgtagg</td>
<td></td>
</tr>
<tr>
<td>cc2-ChIP-4f</td>
<td>tctgccgatttagtgattgtgaa</td>
<td></td>
</tr>
<tr>
<td>cc2-ChIP-6r</td>
<td>gccgtaattatttatgttgt</td>
<td></td>
</tr>
<tr>
<td>cc2-ChIP-7r</td>
<td>ccacactatctcgtcga</td>
<td></td>
</tr>
<tr>
<td>cc2-Probe-3F</td>
<td>gtctttgttttcgacaag</td>
<td></td>
</tr>
<tr>
<td>cc2-RT-6 r</td>
<td>aatacttttactataaaacggt</td>
<td></td>
</tr>
<tr>
<td>WA607</td>
<td>tagacaagcactcataagttc</td>
<td></td>
</tr>
<tr>
<td>WA608</td>
<td>tagacaagcactcataagttc</td>
<td></td>
</tr>
</tbody>
</table>
cc2 deletion (Chapter 4)

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5'-3'</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>ura4 insertion in cc2(cc2:ura4):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCR to fuse ura4 sequence and homology regions of cc2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p1: 50bp cc2+ura4 amplif- Fw</td>
<td>atttctggtaaataagtatatcaggggttttgt</td>
<td>ura4 amplification</td>
</tr>
<tr>
<td>p1: 50bp cc2+ura amplif- Rev</td>
<td>tccacrla tat tgt tta att tgt taa tcc aca gga tat tgt aag ttg aag aac c</td>
<td></td>
</tr>
<tr>
<td>5' p2: L - cc2 fusion URA - For</td>
<td>gttcgtatgcgaatttgagag</td>
<td>left homology to cc2</td>
</tr>
<tr>
<td>3' p2: L - cc2 fusion URA- Rev</td>
<td>acactgcgccaaatcataa</td>
<td></td>
</tr>
<tr>
<td>5' p3: R- cc2 fusion URA - For</td>
<td>cacaatatactgtgatta</td>
<td>right side homology to cc2</td>
</tr>
<tr>
<td>3' p3: R - cc2 fusion URA- Rev</td>
<td>gccaattacctaaacgtag</td>
<td></td>
</tr>
<tr>
<td>5' p5: Final cc2-fusion URA For</td>
<td>gttgaaaggtagttaacctc</td>
<td>final fusion/nested PCR</td>
</tr>
<tr>
<td>3' p5: Final cc2-fusion URA Rev</td>
<td>atgacgggttgttaatag</td>
<td></td>
</tr>
<tr>
<td>cloning construct cc2l-cc1-cc2r (figure 4-2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5' XhoI - cc2 L3 Fw</td>
<td>tactacctcgagcacaaatatcctgtcgatta</td>
<td>cc2L</td>
</tr>
<tr>
<td>3' KpnI - cc2 L3 Rev</td>
<td>tccactacaatcttttaatttttcccagtcggaatag</td>
<td></td>
</tr>
<tr>
<td>5'-NotI-cc2-b</td>
<td>tat tat ggc gcc gcc gtt tgt cca tta ctt tgt 999 g</td>
<td></td>
</tr>
<tr>
<td>3'- BamHI-cc2-b</td>
<td>aat aat gga tcc ggc caa aat cat aaa acc ccc</td>
<td></td>
</tr>
<tr>
<td>5' check cnt1 at cc2 -Fw</td>
<td>cacttcacaaacatcccaag</td>
<td>check cc1:cc2</td>
</tr>
<tr>
<td>3' check cnt1 at cc2 -Rev</td>
<td>caaagaactccaagggag</td>
<td></td>
</tr>
<tr>
<td>partial deletion of cc2 with ura4 genes to create cc2Δ::cc1+ura4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p1: cnt1 in cc2:del cc2-URA Fw</td>
<td>tataaatactacacttttttttttaaatggaa</td>
<td>ura4 amplification</td>
</tr>
<tr>
<td>p1: cnt1 in cc2:del cc2-URA R</td>
<td>gtttctgtagatgtatgcttattac</td>
<td></td>
</tr>
<tr>
<td>p2:cnt1 in cc2:del cc2-URA Fw</td>
<td>gattttgtcacagccgggtt</td>
<td></td>
</tr>
<tr>
<td>p2:cnt1 in cc2:del cc2-URA R</td>
<td>gaaattgatcgttctcc</td>
<td></td>
</tr>
<tr>
<td>p3:cnt1 in cc2:del cc2-URA Fw</td>
<td>aaagcacaaccagtaacc</td>
<td></td>
</tr>
<tr>
<td>p3:cnt1 in cc2:del cc2-URA rev</td>
<td>atacatatagtctgc</td>
<td></td>
</tr>
<tr>
<td>p5:cnt1 in cc2:del cc2-URA Fw</td>
<td>ataaccgactacatcactgc</td>
<td>final fusion/nested PCR</td>
</tr>
<tr>
<td>p5:3' cnt1 at cc2: deletion cc2 URA Rev</td>
<td>catatcctaaaatggttcgtaatgcttccacc</td>
<td></td>
</tr>
<tr>
<td>cnt1 in cc2:Check delcc2-URA R</td>
<td></td>
<td>primers to check cc2Δ::cc1+ura</td>
</tr>
<tr>
<td>5' cnt1 at cc2:check deletion cc2 URA Fw</td>
<td>cagttcgaatacttctac</td>
<td></td>
</tr>
<tr>
<td>deletion of ura from cc2Δ::cc1+ura</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p1: cc2del:deletion ura F</td>
<td>p2:5' cnt1 at cc2: deletion cc2 ura</td>
<td></td>
</tr>
<tr>
<td>p1: cc2del:deletion ura R</td>
<td>p2:3' cnt1 at cc2: deletion cc2 ura rev</td>
<td></td>
</tr>
<tr>
<td>p2:cc2del: deletion ura F</td>
<td>tataaaactacaccttttttttttaaatggaa</td>
<td></td>
</tr>
<tr>
<td>p2:cc2del: deletion ura R</td>
<td>gctatagcctacatcacttcagttcttc</td>
<td></td>
</tr>
<tr>
<td>final PCR</td>
<td>p5:cnt1 in cc2:del cc2-ura Fw</td>
<td>final</td>
</tr>
</tbody>
</table>
### Chapter 2: Materials and Methods

<table>
<thead>
<tr>
<th>p2:cc2del: deletion ura r</th>
<th>fusion/nested PCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>check deletion ura</td>
<td>p2:5' cnt1 at cc2: deletion cc2 ura fw</td>
</tr>
<tr>
<td>check deletion ura</td>
<td>CCGGTAATCGTTGCAAAGTGC (probe 1f)</td>
</tr>
<tr>
<td>to check cc2 deletion use</td>
<td>gcaaaaatacttgatgcgca</td>
</tr>
<tr>
<td>imr1-r13</td>
<td>ccggtaatcggtaaagtgc</td>
</tr>
<tr>
<td>5'cnt2- 1R</td>
<td></td>
</tr>
</tbody>
</table>

#### Southern analysis of cc2Δ::cc1 (Chapter 4)

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-3’</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>cc2-ChIP 2F</td>
<td>gctttggtatcgtgtaagccg</td>
<td>probe on cc2 (Fig 4.3)</td>
</tr>
<tr>
<td>3’-cnt2-2F</td>
<td>ctaaaacctccctgatatc (probe 2R)</td>
<td></td>
</tr>
<tr>
<td>F3-cnt1</td>
<td>cggttagttacatatatcg</td>
<td>probe on cc1(Fig 4.3)</td>
</tr>
<tr>
<td>R4-cnt1</td>
<td>gattaatcataaatagcc</td>
<td></td>
</tr>
<tr>
<td>5’-cnt2-6R</td>
<td>cggtgcacacctttgaaag (probe 6f)</td>
<td>probe on cc2 (Fig 4.4)</td>
</tr>
<tr>
<td>cc2- ChIP 6rev</td>
<td>tctcgctagttttgaaag</td>
<td></td>
</tr>
<tr>
<td>F16-cnt1</td>
<td>ccatttgctaagttcgactc</td>
<td></td>
</tr>
<tr>
<td>R17-cnt1</td>
<td>cagttttgtatgttagtgg</td>
<td></td>
</tr>
</tbody>
</table>

#### qPCR analysis of plasmid Chapter 4)

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-3’</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>q_act1 Fw</td>
<td>cccaaatccaaccgtgagaagtg</td>
<td>act1</td>
</tr>
<tr>
<td>q_act1 rev</td>
<td>cccagttccagctgatacctagt</td>
<td>act1</td>
</tr>
<tr>
<td>qcnt1 Fw</td>
<td>cagacaatgcatggtactatc</td>
<td>cc1/3</td>
</tr>
<tr>
<td>qcnt1 rev</td>
<td>aggtgaaagcgtgaagtg</td>
<td>cc1/3</td>
</tr>
<tr>
<td>qPCR LM 9Fw</td>
<td>gcatctatgtgctcttc</td>
<td>L</td>
</tr>
<tr>
<td>qPCR LM 9Rev</td>
<td>gaaagttcatgatgcagtt</td>
<td>L</td>
</tr>
<tr>
<td>qPCR 15 fw</td>
<td>gtaatgcctttttggtcg</td>
<td>M</td>
</tr>
<tr>
<td>qPCR 15 rev</td>
<td>atgacatgctggaagagtc</td>
<td>M</td>
</tr>
<tr>
<td>q_cnt2 F</td>
<td>ctaaaacatcaacacggtccac</td>
<td>N</td>
</tr>
<tr>
<td>q_cnt2 R</td>
<td>taagcggcagatccttgag</td>
<td>N</td>
</tr>
<tr>
<td>RT-cc2-19 for</td>
<td>gactataactagaccacttag</td>
<td>O</td>
</tr>
<tr>
<td>3’-cnt2-6F (probe 6R)</td>
<td>ctgtagtaaatctcaagagaac</td>
<td>O</td>
</tr>
<tr>
<td>RT-cc2-21 rev</td>
<td>ctgcatatctgcatcttgag</td>
<td>P</td>
</tr>
<tr>
<td>cc2-chip-7f</td>
<td>agccgttcctgcaacagg</td>
<td>P</td>
</tr>
</tbody>
</table>

#### Synthetic sequence (Chapter 5)

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-3’</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ random AT BamHI-Fw</td>
<td>tactagatcgcataataatccagttatcttc</td>
<td>SynAT cloning</td>
</tr>
<tr>
<td>3’ random AT BglIII-R</td>
<td>tactagatcgcataataatccagttatcttc</td>
<td></td>
</tr>
<tr>
<td>WA607</td>
<td>aatacttttatctacaagcctg</td>
<td>SynAT ChIP (fig 5-1)</td>
</tr>
<tr>
<td>WA608</td>
<td>tagacagcactcataagtctc</td>
<td></td>
</tr>
<tr>
<td>5’ BamHI-SynR34 F1</td>
<td>tactagatcgcataataatccagttatcttc</td>
<td>SynRLM cloning</td>
</tr>
<tr>
<td>3’- BglIII SynR34 - R2</td>
<td>tactagatcgcataataatccagttatcttc</td>
<td></td>
</tr>
<tr>
<td>5’ R34-a-ChIP- 1Fw</td>
<td>gactgtctctagattatatta</td>
<td>SynRLM ChIP</td>
</tr>
</tbody>
</table>
### Creation of orc1-BAHΔ (Chapter 6)

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-3’</th>
<th>Note</th>
</tr>
</thead>
</table>
| cloning bah del into pDUAL then integration in 2025 (orc1-4ts)  
p1: orc1 bahdel fusion fw | gaagtataatgctagccga | 5’ region of orc1+promoter |
| p1: orc1 bahdel fusion rev | agacgtgaataaagaagttc |
| p2: orc1 bahdel fusion fw | ggagtgaagagcccttcacaaaaagagggcttattactgtcctcttatattagctc |
| p2: orc1 bahdel fusion rev | ggccgtttttgtagtaagc |
| nested orc1 bahdel fusion fw- pStI | tactacctgcaggttgtatatattggtgcc |
| nested orc1 bahdel fusion rev- BamHI | tactacggtatctgtgaatatgaaatgg |
| orc1 rev1 | gatttaactttggaactgc |
| orc1 fw1 | gctgccggtgacggtagtt |
| **orc1 fw2** | gatgtggctttattagctc |
| orc1 rev2 | gaaatccaggaaagtgtgc |
| orc1 fw3 | gcagaacctctgaatttc |

### Deletion of orc1-4ts from 2025 integration pDUAL -BAHΔ  

| p1: 5’ orc1 deletion fw | p1: orc1 bahdel fusion fw ctataagggctactgtcatg |
| p1: 5’ orc1 deletion rev | ctatagaggctactgtcatg |
| p2: orc1 deletion bahler fw | ttgtgtatatatttggctcttgtacatagcatgcagctgct attaggtattttagcttgctc |
| p2: orc1 deletion bahler rev | taglaacggcaaaaattcgtctgctttcataaagtcgatagtgaataggg |
| p3: 3’ orc1 deletion fw | cccttttcatattcagac |
| p3: 3’ orc1 deletion rev | aggagtcttgatctac |

### RT-PCR on p-3xLM (Chapter 7)

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-3’</th>
<th>Figure 7-4</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM 1F</td>
<td>accgaggtgtaaactgtgacg</td>
<td>1</td>
</tr>
<tr>
<td>LM 2F</td>
<td>tcaattagggccatgttaag</td>
<td></td>
</tr>
<tr>
<td>RT-cc2 7Fw</td>
<td>ccccctagcaagttgagcaccc</td>
<td></td>
</tr>
<tr>
<td>qPCR LM 16 rev</td>
<td>gttgcgtagcaacattatgg</td>
<td></td>
</tr>
<tr>
<td>qPCR LM 17 fw</td>
<td>gatgcgttaaaccagctatat</td>
<td></td>
</tr>
<tr>
<td>qPCR LM 5 rev</td>
<td>gaacaaacccctcttgactac</td>
<td></td>
</tr>
<tr>
<td>LM 18 fw</td>
<td>gcataagtaaactaaagtc</td>
<td></td>
</tr>
<tr>
<td>LM 18 rev</td>
<td>cttgtagcataactaatgg</td>
<td></td>
</tr>
<tr>
<td>LM 19 fw</td>
<td>tcattgataacactaccc</td>
<td></td>
</tr>
<tr>
<td>LM 7 rev</td>
<td>tatactgtcgcagctac</td>
<td></td>
</tr>
<tr>
<td>LM 8 fw</td>
<td>cgtgaattatatatagcagc</td>
<td></td>
</tr>
<tr>
<td>RT-cc2-9 rev</td>
<td>gcagcaactcatctagtgc</td>
<td></td>
</tr>
<tr>
<td>LM 21 fw</td>
<td>catatccatctctgtaag</td>
<td></td>
</tr>
<tr>
<td>LM 21 rev</td>
<td>tctgttccaagagatataaag</td>
<td></td>
</tr>
<tr>
<td>RT-cc2-11 rev</td>
<td>cttgtagaactctgtgcttg</td>
<td></td>
</tr>
<tr>
<td>LM 22 rev</td>
<td>gaaatgcataacatttccag</td>
<td></td>
</tr>
<tr>
<td>LM 23 fw</td>
<td>cttgtaataaaactaatgtgc</td>
<td></td>
</tr>
<tr>
<td>LM 14 rev</td>
<td>cctattttaactaagccag</td>
<td></td>
</tr>
<tr>
<td>qPCR LM 15 fw</td>
<td>gtaattgcaaacccttgtgccc</td>
<td></td>
</tr>
<tr>
<td>3-cc2-M1kb-rev</td>
<td>tcgagttatgtgtatctac</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 2: Materials and Methods

Oligonucleotides used the 5’-RACE PCR (Chapter 7)

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-3’</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>ChIP 4f</td>
<td>ctagaccttaagttctaccc</td>
<td>TSS-2 outer</td>
</tr>
<tr>
<td>RT-10rev</td>
<td>ctaacagttagtttgttg</td>
<td>TSS-2 inner</td>
</tr>
<tr>
<td>Cc2-R4</td>
<td>tactacctgctgacctggacatatctgtgtg</td>
<td>TSS-3 outer</td>
</tr>
<tr>
<td>LM 15 rev</td>
<td>atgcacatgcgtgaaatgc</td>
<td>TSS-3 inner</td>
</tr>
<tr>
<td>RT-cc2 11 fw</td>
<td>ctgtagaaactgctgtgg</td>
<td>TSS-5 outer</td>
</tr>
<tr>
<td>ChIP 4f w</td>
<td>ctagaccttaagttctaccc</td>
<td>TSS-5 inner</td>
</tr>
<tr>
<td>5’ont2-3R (probe 3F)</td>
<td>cacaatattcgtgacatg</td>
<td>TSS-7 outer</td>
</tr>
<tr>
<td>LM 7rev</td>
<td>tatctctgctgacatc</td>
<td>TSS-7 inner</td>
</tr>
<tr>
<td>LM 9rev</td>
<td>Gaaggttggtactgacgt</td>
<td>TSS-8 outer</td>
</tr>
<tr>
<td>RT-cc2 9 rev</td>
<td>gcagcaatctctggtgc</td>
<td>TSS-8 inner</td>
</tr>
<tr>
<td>LM 19 fw</td>
<td>ctgtagaaactgctgtg</td>
<td>TSS-24 outer</td>
</tr>
<tr>
<td>RT-cc2 9 fw</td>
<td>actacatctctggtgc</td>
<td>TSS-24 inner</td>
</tr>
<tr>
<td>ChIP 4 fw</td>
<td>ctagaccttaagttctaccc</td>
<td>TSS-28 outer</td>
</tr>
<tr>
<td>qPCR LM 13 fw</td>
<td>ctgaccttttgcctttaa</td>
<td>TSS-28 inner</td>
</tr>
<tr>
<td>LM 23 fw</td>
<td>gctgaaatataactataatgtgc</td>
<td>TSS-29 outer</td>
</tr>
<tr>
<td>RT-cc2 12 fw</td>
<td>ctgacattatatataatcc</td>
<td>TSS-29 inner</td>
</tr>
<tr>
<td>Act rev</td>
<td>gatctaccacagcataccag</td>
<td>Actin outer</td>
</tr>
<tr>
<td>Act rev-small</td>
<td>aggagctccagttacaaacg</td>
<td>Actin inner</td>
</tr>
<tr>
<td>5’RACE inner primer</td>
<td>cgccggtcgaacactgctggttgctggttgatg</td>
<td>5’RACE inner</td>
</tr>
<tr>
<td>5’RACE outer primer</td>
<td>gcgtgggctgaggtgacactg</td>
<td>5’RACE outer</td>
</tr>
</tbody>
</table>

LacZ assay (Chapter 5):

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-3’</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>prom-transcr28A/29A Fw</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p200b</td>
</tr>
<tr>
<td>prom-transcr28A/29A rev</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p200b</td>
</tr>
<tr>
<td>prom-transcr3-5 fw</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p200a</td>
</tr>
<tr>
<td>prom-transcr3-5 rev</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p200a</td>
</tr>
<tr>
<td>promoter 29 FW</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p50a</td>
</tr>
<tr>
<td>promoter 29 Rev</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p50a</td>
</tr>
<tr>
<td>promoter28A fw</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p50b</td>
</tr>
<tr>
<td>promoter28A rev</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p50b</td>
</tr>
<tr>
<td>prom28A/29A flip fw</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p200b-inv</td>
</tr>
<tr>
<td>prom28A/29A flip rev</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p200b-inv</td>
</tr>
<tr>
<td>promoter 50bp-1 rev</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p50c</td>
</tr>
<tr>
<td>promoter 50bp-1 fw</td>
<td>tactacctgtcgagttctactgtagtgtgtctctac</td>
<td>p50c</td>
</tr>
</tbody>
</table>
Construction of Rpb1-tcs-CTD (Chapter 7)
## Chapter 2: Materials and Methods

### 62

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-3’</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>p1:Rpb1TEV II Fw</td>
<td>ggagaaagagatggatggcaaaatgggactagtaaaaccagagggatatcctggcaagaggtttttcaaccagaacaa</td>
<td></td>
</tr>
<tr>
<td>p1:Rpb1TEV II rev</td>
<td>catagcagacggagagatgggactagtaaaaccagagggatatcctgcaagaggtttttcaaccagaacaa</td>
<td></td>
</tr>
<tr>
<td>Rpb1 4 fw</td>
<td>tgtaccccttataggatgtctcccgaggtttttcaaccagaacaa</td>
<td></td>
</tr>
<tr>
<td>HA rev</td>
<td>gtaaaaaggtacacaggctatatccggtttttcaaccagaacaa</td>
<td></td>
</tr>
<tr>
<td>Bahler rev</td>
<td>cggatccccgggttaatttaaaatgggactagtaaaaccagagggatatcctgtttttcaaccagaacaa</td>
<td></td>
</tr>
<tr>
<td>p3:Rpb1TEV II rev-Bahler</td>
<td>attgaatagttaacgtacaattttagatgtatgactcatgtttttcaaccagaacaa</td>
<td></td>
</tr>
<tr>
<td>nested PCR</td>
<td>p1:rbp1tev ii fw ctacaagtatattgactctggataggtttttcaaccagaacaa</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence 5’-3’</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>cnt1 probes (Choi et al 2011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T7_cnt1-f4</td>
<td>taatacgactcactatagggagacattttagtattaatacagc</td>
<td>Figure 7-1</td>
</tr>
<tr>
<td>cnt1-r6</td>
<td>agaatgtgtgcagacttttg</td>
<td>Figure 7-1</td>
</tr>
<tr>
<td>T7_cnt1-f6</td>
<td>taatacgactcactatagggagacaaagtagctcgacatcccttg</td>
<td>Figure 7-1</td>
</tr>
<tr>
<td>cnt1-r8</td>
<td>ccagacactaatctcccttg</td>
<td>Figure 7-1</td>
</tr>
<tr>
<td>T7_cnt1-r6</td>
<td>taatacgactcactatagggagagaaatgtgcagacttttg</td>
<td>Figure 7-1</td>
</tr>
<tr>
<td>cnt1-f4</td>
<td>cattatatagcataataacag</td>
<td>Figure 7-1</td>
</tr>
<tr>
<td>T7_cnt1-r8</td>
<td>taatacgactcactatagggagacagcactaatataccttcttg</td>
<td>Figure 7-1</td>
</tr>
<tr>
<td>cnt1-r6 5'-t7_cnt2-4f</td>
<td>taatacgactcactatagggagacagcactatatctgtatcttcttg</td>
<td>Figure 7-3</td>
</tr>
<tr>
<td>cnt1-r3 3'-t7_cnt2-3f</td>
<td>taatacgactcactatagggagacagcactatatctgtatcttcttg</td>
<td>Figure 7-3</td>
</tr>
<tr>
<td>cnt2-3f</td>
<td>gcatcgatataatttacag</td>
<td>Figure 7-3</td>
</tr>
<tr>
<td>cnt2-3r 3'-t7_cnt2-3f</td>
<td>taatacgactcactatagggagacagcactatatctgtatcttcttg</td>
<td>Figure 7-3</td>
</tr>
<tr>
<td>cnt2-4f</td>
<td>tgcgagttagttgatatcc</td>
<td>Figure 7-3</td>
</tr>
<tr>
<td>cnt2-4r 3'-t7_cnt2-3f</td>
<td>taatacgactcactatagggagacagcactatatctgtatcttcttg</td>
<td>Figure 7-3</td>
</tr>
<tr>
<td>cnt2-4r</td>
<td>atgacagctgcgtgagaagct</td>
<td>Figure 7-3</td>
</tr>
<tr>
<td>cnt2-4r 3'-t7_cnt2-4r</td>
<td>taatacgactcactatagggagacagcactatatctgtatcttcttg</td>
<td>Figure 7-3</td>
</tr>
</tbody>
</table>
### 2.9 Strains used in this thesis

<table>
<thead>
<tr>
<th>Strain Number</th>
<th>Genotype</th>
<th>Chapter</th>
<th>note</th>
</tr>
</thead>
<tbody>
<tr>
<td>9924</td>
<td>h90 ade6-704 H3.2K9R-HYGR H3.1/H4.1::his3+ H3.3/H3.4::arg3+ his3-D1 arg3-D4 leu1-32 ura4-D4/D18 cc2::his3</td>
<td>3</td>
<td>K9R</td>
</tr>
<tr>
<td>A5110</td>
<td>h- ade6-704 H3.2-wt H3.1/H4.1::his3+ H3.3/H3.4::arg3+ his3-D1 arg3-D4 leu1-32 ura4-D4/D18 cc2::his3</td>
<td>3</td>
<td>wt*</td>
</tr>
<tr>
<td>A2213</td>
<td>h+ ade6-704-hygMX6 ura4-D4/D18 leu1-32 his3-D1 cc2::his3 arg3-D4</td>
<td>3, 4</td>
<td></td>
</tr>
<tr>
<td>A7213</td>
<td>h? clr4Δ::NAT ade6-210 arg3-D4 his3-D1 leu1-32 ura4-D4/D18 cc2::ura4* (2 Kb from Sphi)</td>
<td>4</td>
<td>cc2::ura4+</td>
</tr>
<tr>
<td>A7214</td>
<td>h? clr4Δ::NAT ade6-210 arg3-D4 his3-D1 leu1-32 ura4-D4/D18 cc2::cc1(TM)</td>
<td>4</td>
<td>cc2::cc1</td>
</tr>
<tr>
<td>A7216</td>
<td>h? clr4Δ::NAT ade6-210 arg3-D4 his3-D1 leu1-32 ura4-D4/D18 cc2::cc1(TM)+ura4*</td>
<td>4</td>
<td>cc2Δ::cc1+ ura4</td>
</tr>
<tr>
<td>A7255</td>
<td>h? clr4Δ::NAT ade6-210 his3-D1 leu1-32 ura4-D18 arg3-D4 cc2Δ6kb::cc1</td>
<td>4</td>
<td>cc2Δ::cc1</td>
</tr>
<tr>
<td>A7373</td>
<td>h- ade6-704-HYGMX6 his3-D1 leu1-32 ura4-D4/D18? arg3? cc2Δ6kb::cc1</td>
<td>4</td>
<td>cc2Δ::cc1</td>
</tr>
<tr>
<td>A7919</td>
<td>h90 ade6-704 H3.2K9R-HYGR H3.1/H4.1::his3+ H3.3/H3.4::arg3+ his3-D1 arg3-D4 leu1-32 ura4-D4/D18 cc2Δ6kb::cc1::</td>
<td>4</td>
<td>K9R</td>
</tr>
<tr>
<td>A7776</td>
<td>h- ade6-704 H3.2- wf H3.1/H4.1::his3+ H3.3/H3.4::arg3+ his3-D1 arg3-D4 leu1-32 ura4-D4/D18 cc2Δ6kb::cc1</td>
<td>4</td>
<td>wt*</td>
</tr>
<tr>
<td>A7555</td>
<td>h? cc2Δ6kb::cc1 mis18-GFP-Leu2+ ade6-704-HYGMX6 leu1-32 his3-D1? ura4-D4/D18? arg3+</td>
<td>6</td>
<td>Mis18-GFP</td>
</tr>
<tr>
<td>A7552</td>
<td>h? cc2Δ6kb::cc1 sim1-GFP- his3+ ade6-704-HYGMX6 leu1-32 his3-D1</td>
<td>6</td>
<td>Scm3-GFP</td>
</tr>
<tr>
<td>A7550</td>
<td>h? cc2Δ6kb::cc1 cnp20-GFP-KAN ade6-704-HYGMX6 leu1-32 his3-D1</td>
<td>6</td>
<td>Cnp20-GFP</td>
</tr>
<tr>
<td>2025</td>
<td>h- orp1-4ts ade6-210 leu1-32 ura4-D18</td>
<td>6</td>
<td>orp1-4</td>
</tr>
<tr>
<td>A8025</td>
<td>h- orp1-ts leu1+: pDUAL-orp1 ade6-210 leu1-32 ura4-D18</td>
<td>6</td>
<td>Orp1-4</td>
</tr>
<tr>
<td>A8027</td>
<td>h- orp1-ts leu1+: pDUAL-orp1-BAHΔ ade6-210 leu1-32 ura4-D18</td>
<td>6</td>
<td>Orp1-4 + orp1 Δ</td>
</tr>
<tr>
<td>A8112</td>
<td>h- orp1-4 ts:: NAT leu1+: pDUAL-orp1 ade6-210 leu1-32 ura4-D18</td>
<td>6</td>
<td>BAHΔ</td>
</tr>
<tr>
<td>A8137</td>
<td>h- orp1-4 ts:: NAT leu1+: pDUAL-orp1-BAHΔ ade6-210 leu1-32 ura4-D18</td>
<td>6</td>
<td>orp1 Δ + orp1 Δ</td>
</tr>
<tr>
<td>1180</td>
<td>h+ ade6-210 leu1-32 ura4-D18 or1R(Sph1):ade6</td>
<td>6</td>
<td>or1R:ade6</td>
</tr>
<tr>
<td>A7723</td>
<td>h+ dcr1Δ::NAT ade6-210 leu1-32 ura4-D18 or1R(Sph1):ade6</td>
<td>6</td>
<td>dcr1Δ</td>
</tr>
</tbody>
</table>

63
A8149  h? orp1-ts:: NAT leu1+: pDUAL-orp1 ortr1R(SphI):ade6 ade6-210 leu1-32 ura4-D18
A8151  h? orp1-ts:: NAT leu1+: pDUAL-orp1-BAHΔ ortr1R(SphI):ade6 ade6-210 leu1-32 ura4-D18
498   h+ ade6-210 leu1-32 ura4-DS/E imr1R(Ncol)::ura4
8877  h sir2Δ::Nat ade6-210 leu1-32 arg3D4 his3Δ+ imr1(Ncol)::ura4
A8164  h? orp1-4ts:: NAT leu1+: pDUAL-orp1 imr1R(Ncol)::ura4
A8166  h? orp1-4ts:: NAT leu1+: pDUAL-orp1-BAHΔ imr1R(Ncol)::ura4
A7635  h+ ade6-32 uRA4-D18 ade6-M210/D1? arg3Δ::ura4+ TM1::arg3
A868   TM1::arg3 hrp1D::LEU2 ade6-210 ura4-DS/E or D18 leu1-32 his3-D1 arg3-D
A8170  h? orp1-4ts:: NAT leu1+: pDUAL-orp1 TM1::arg3 arg3Δ::ura4
A8174  h? orp1-4ts:: NAT leu1+: pDUAL-orp1-BAHΔ TM1::arg3
A592   h- mis18-262
2919  h- mis6-302 leu1-32
6960  h- leu1-32 ura4- cnp1::ura4+ lys1::::cnp1-1
A7528  h? cc2Δ6kb::cc1 dis3-54 ade6-704-HYGMX6 his3-D1? leu1-32 ura4-DSE/D18?
A8249  h- ade6-704-HYGMX6 his3-D1 leu1-32 ura4-DSE/D18?
c2Δ6kb::cc1 Rpb1-tcs-CTD-KAN
A8141  h- ade6-704-HYGMX6 his3-D1 leu1-32 ura4-DSE/D18?
c2Δ6kb::cc1 Z::padh21-Cnp3C-CFP-9xmyc-TEV-NAT
A8254  h- ade6-704-HYGMX6 his3-D1 leu1-32 ura4-DSE/D18?
c2Δ6kb::cc cc1 Rpb1-tcs-CTD-KAN 1 Z::padh21-Cnp3C-CFP-9xmyc-TEV-NAT
A8594  h+ cc2D6kb::cc1 Leu1+::padh21-TetR-CFP-9xmyc-TEV(enzyme) ade6-704-HYGMX6 ura4-D18/D3E arg3? his3-D1? leu1-32?
A8592  h+ cc2D6kb::cc1 Rpb1-tcs-CTD-KAN Leu1+::padh21-TetR-CFP-9xmyc-TEV(enzyme) ade6-704-HYGMX6 ura4-D18/D3E arg3? his3-D1? leu1-32?
CHAPTER 3

3 INVESTIGATION OF THE MINIMAL SEQUENCE REQUIREMENT FOR CENP-A<sup>Cnp1</sup> CHROMATIN ASSEMBLY

3.1 INTRODUCTION

In fission yeast, although the three centromeres possess different sequences, they are characterised by the same organisation where two outer repeat regions (otr) surround the central core domain (cc/cnt) that is flanked by two innermost repeat regions (imr) (Steiner et al., 1993; Takahashi et al., 1992). Two distinct chromatin domains can be described at centromeres: a heterochromatic domain over the outer repeats and part of the innermost repeats and a CENP-A<sup>Cnp1</sup> chromatin region found exclusively within the central core region (reviewed in (Pidoux and Allshire, 2004). However, previous studies suggest that in fission yeast there is not a single critical sequence and centromere sequence requirements are more complex than the centromeres in <i>S. cerevisiae</i> (Clarke and Carbon, 1985).

Functional analyses have been performed in order to identify the minimal elements required for centromere formation. On a minichromosome, the minimal sequence requirement for the formation of a functional centromere consists of both outer repeats and central core domain (Baum et al., 1994; Hahnenberger and Carbon, 1991; Takahashi et al., 1992). The outer repeats allow the assembly of methylated H3K9 dependent heterochromatin, which somehow promotes the assembly of CENP-A<sup>Cnp1</sup> on neighbouring central domain (Folco et al., 2008a). However, the establishment of CENP-A<sup>Cnp1</sup> chromatin and the acquisition of mitotic stability can be slow and the segregation function of such minichromosomes in wild-type cells is relatively low and stochastic (Steiner and Clarke, 1994).

On minichromosomes, CENP-A<sup>Cnp1</sup> is deposited exclusively over the central core domain and not on other sequences of the plasmid, suggesting the existence of sequence preferences in centromere formation (Folco et al., 2008a). However, the cc2 region shows functional redundancy since deletion of part of the sequence does not affect centromere function, while complete deletion of the sequence abolishes centromere formation (Baum et al., 1994). The functional redundancy together with the lack of a conserved sequence between cc1, cc2 and cc3 suggest that if a critical
sequence exists, it is either very small or may encode for some general biological process/feature.

In this chapter, I will describe the analysis of sequence requirements for centromere establishment and the identification of a minimal central core sequence that would facilitate the discovery of the features required for CENP-A\(^{\text{Cnp1}}\) chromatin formation.

### 3.2 RESULTS

Manipulation of the endogenous centromere is difficult since the disruption of regulatory elements may affect the functionality of the centromere itself leading to cell death. Moreover, the requirements for maintenance of CENP-A\(^{\text{Cnp1}}\) chromatin at a pre-existing endogenous centromere may differ from those required to establish CENP-A\(^{\text{Cnp1}}\) chromatin on naïve centromeric DNA templates. Thus, plasmid based circular minichromosomes have been utilised to investigate the DNA sequence requirements for the establishment of CENP-A\(^{\text{Cnp1}}\) chromatin.

Previous studies have utilised a functional minichromosome that possesses 8.5 kb of central domain 2 sequence (NcoI fragment) that includes 1kb of each \textit{imr2L} and \textit{imr2R} (Folco et al., 2008a). Centromere 1 (cen1) and cen3 share a 99% identical sequence (TM sequence) while cen2 possesses an element in its central core that is 48% identical to cen1/3 sequence (Pidoux and Allshire, 2004; Wood et al., 2002). This difference in sequence between cc1/3 and cc2 makes cc2 the ideal centromere sequence for this type of study since it is easily distinguished from the other centromeric sequences using PCR and other sequence-based assays.

Recent observations (Folco et al. unpublished–Allshire lab) indicate that the need for flanking heterochromatin can be bypassed in cells expressing a mutant histone H3 with lysine 9 replaced by arginine (H3K9R). In contrast to the canonical heterochromatin containing plasmid whose establishment of CENP-A\(^{\text{Cnp1}}\) chromatin in wild type cells is stochastic and not efficient, a plasmid containing only the 8.5 kb of the central domain of cen2 (pcc2) consistently assembles CENP-A\(^{\text{Cnp1}}\) chromatin when introduced into cells expressing only mutant histone H3K9R and it does not assemble in cells expressing only wild type H3 or mutant H3K9A (data not shown).

The strains utilised possess only one copy of histone gene H3 (H3.2K9R or wt) and histone H4 (H4.3) while the other two copies (H3.1/H4.1 and H3.3/H4.3) are deleted and replaced with marker genes \textit{his3}\(^{+}\) and \textit{arg3}\(^{+}\) respectively (H3.1/H4.1::\textit{his3}\(^{+}\) H3.3/H4.3::\textit{arg3}\(^{+}\)) (Mellone et al., 2003). The strains utilised in this thesis are referred as “wt*” for the strain possessing the only copy of histone H3 with its original sequence and H3K9R for the strain expressing the mutant version of H3.
The use of the pcc2 plasmid combined with the H3K9R strain provides the basis for a robust and simple assay for the detection of those elements from the central core sequence that are required for CENP-A\textsuperscript{Cnp1} assembly. Therefore, in order to determine if specific central core 2 (cc2) sequences are critical for CENP-A\textsuperscript{Cnp1} deposition and whether they encode for specific features, fragments of the cc2 were analysed for their ability to establish CENP-A\textsuperscript{Cnp1} chromatin.

3.2.1 The central core 2 sequence shows functional redundancy

In fission yeast, the central core sequences are surrounded by inverted elements (imr). Whether the direction of these repeats plays a role in the ability to assemble CENP-A\textsuperscript{Cnp1} had not been investigated yet. The strains used in this study (H3K9R and wt*) possess a silent his3\textsuperscript{+} gene inserted at the Sph\textsuperscript{I} site in the endogenous cc2 so that specific primers can be used to distinguish endogenous centromere and cc2 sequence in the vectors (Figure 3-1A).

In order to analyse a possible role of directionality of sequence elements in CENP-A\textsuperscript{Cnp1} deposition, a minichromosome was constructed that contains the second half (fragment N-R) inverted relative to the orientation of the native sequence (Figure 3-1C). In addition, the fragment "N" (1 kb in length) was deleted from the construct. The plasmid was transformed into H3K9R and wild type strains and ChIP for CENP-A\textsuperscript{Cnp1} was performed. CENP-A\textsuperscript{Cnp1} was found in the plasmid in all colonies analysed, indicating that the reversed orientation and the deletion of the N fragment do not affect the recruitment of CENP-A\textsuperscript{Cnp1} (Figure 3-1C).

In order to find a functional minimal sequence, the capability of two “halves” of cc2 (4 kb and 4.5 kb) to assemble CENP-A\textsuperscript{Cnp1} was tested in H3K9R mutant. For both constructs low levels of CENP-A\textsuperscript{Cnp1} were detected in one colony of the 5 analysed compared to the cc2 full-length where all 5 have CENP-A\textsuperscript{Cnp1} assembled (Figure 3-1D-E). However, for this analysis oligonucleotides annealing on the vector backbone and on centromere sequences were used to distinguish between the endogenous centromere and minichromosome. On a minichromosome CENP-A\textsuperscript{Cnp1} nucleosomes assemble just on centromere sequence without spreading on other sequence of the vector (Folco et al., 2008a). Therefore, the probability that the fragment between vector and central core is present in the immunoprecipitation is lower if compared to one in the internal region of the centromere. Therefore, the low level of CENP-A\textsuperscript{Cnp1} detected could be influenced by the region analysed.
Figure 3-1. Central core 2 shows functional redundancy. Schematic representation of the constructs created and ChIP results. The minichromosomes were transformed in wild type strain (wt*) and H3K9R mutant and anti-CENP-A<sup>Cnp1</sup> ChIP was performed on 5 colonies. Quantification of CENP-A<sup>Cnp1</sup> enrichment is normalised on <i>fbp1</i> and the average of the 5 colonies indicated. cc1/3: positive control at endogenous cnt1/3; cc2: plasmid containing the cc2 fragment. When only 1 colony showed CENP-A<sup>Cnp1</sup> deposition, the value indicated is derived from this one single colony. A) The endogenous centromere 2 (left) is represented with the insertion at its SphI site of the <i>his3</i><sup>+</sup> gene that permits to distinguish cnt2 from the minichromosome. Schematic representation of the plasmid used in this study (right). cc2 is divided into fragments (from J to R) of approximately 1 kb in length. B) Scheme of the construct containing the full length 8.5 kb of cc2 sequence and ChIP data related. C) Plasmid containing the first half of cc2 (4 kb) and the second half (3.5 kb) inverted in the orientation relative to the native sequence; D-E) Minichromosomes containing the “halves” of cc2 sequence; F-H) Plasmids with deletions of the central region of the cc2 sequence.
The constructs containing subfragments of cc2 described above showed a reduction in the efficiency of CENP-A\textsuperscript{Cnp1} recruitment compared to that containing the full length cc2. One of the main differences between those constructs is that they contain 900 bp of either \textit{imr2R} or \textit{imr2L}. To determine whether inverted \textit{imr} repeats at each side of cc2 promote CENP-A\textsuperscript{Cnp1} deposition, cc2-containing plasmids with several deletions in central region of cc2 were constructed and assayed (Figure 3-1F-H). In these plasmids it was possible to analyse the middle region of the central core sequence cloned in the plasmid since the different deletions create novel and specific primer pair combinations.

All the constructs containing 7.5 kb and 5.5 kb of cc2 sequence were found capable of assembling CENP-A\textsuperscript{Cnp1} on their sequence in all the colonies analysed (5/5), suggesting that the deleted central sequences are not necessary for CENP-A\textsuperscript{Cnp1} recruitment. In contrast, the efficiency of CENP-A\textsuperscript{Cnp1} recruitment drops considerably with the plasmid containing 4 kb (Figure 3-1D-E-H). The constructs shown in Figure 3-1G and Figure 3-1H differ by just 1 kb. It is possible that this fragment may be responsible for the decrease in frequency of CENP-A\textsuperscript{Cnp1} establishment. However, the comparison of all the different constructs analysed suggests that the central core is composed of redundant sites interspersed within its sequence, as previously suggested by Baum et al. (1994).

3.2.2 A 2 kb-repeated sequence from cc2 is sufficient to assemble CENP-A\textsuperscript{Cnp1}

If the central core 2 sequence contains redundant elements that attract CENP-A\textsuperscript{Cnp1}, any combination of central core sequences would be able to recruit CENP-A\textsuperscript{Cnp1}. To address this possibility, a 2kb region from the middle of cc2 (called here LM) was selected and cloned into the vector without the \textit{otr} repeats. To assess if the total length of the sequence could affect CENP-A\textsuperscript{Cnp1} assembly, the LM fragment was multimerized to create 2xLM (4Kb) and 3xLM (6kb) arranged as direct repeats (Figure 3-2A). It is important to note that the LM fragment is not critical for CENP-A\textsuperscript{Cnp1} establishment since the construct shown in Figure 3-1G lacks this sequence and retains the ability to efficiently recruit CENP-A\textsuperscript{Cnp1}.

The constructs were transformed into H3K9R and wild type strains and anti-CENP-A\textsuperscript{Cnp1} ChIP was performed on 5 independent colonies.

The vector containing one copy of the LM sequence was not able to recruit CENP-A\textsuperscript{Cnp1} in any of the colonies analysed (Figure 3-2). However, the limitation of not having plasmid specific primer sites in the central position of the construct may influence the interpretation of this analysis. In fact, an ambiguous result was
Figure 3-2. A 2 kb repeated sequence from cc2 is sufficient to assemble CENP-A<sup>Cnp1</sup>. Plasmid containing the LM fragment, 2xLM and 3xLM were transformed into wild type and H3K9R cells anti-CENP-A<sup>Cnp1</sup> ChIP was performed in 5 independent colonies. A) Multiplex PCR on endogenous centromere 1/3 (cc1/3), <i>fbp1</i> + (control) and LM sequence. The plasmid is analysed with primer pairs in the between LM and the vector (primer pair 1- black) and between two repeats of the LM sequence (primer pair 2- red). B) Frequency of colonies with CENP-A<sup>Cnp1</sup> deposited on the different plasmids.
observed for the construct 2xLM. The analysis of the junction between the LM sequence and the vector backbone suggests that CENP-A<sup>Cnp1</sup> is not assembled at 2xLM while if primer pairs specific for the M-L junction (a non native sequence) were used, CENP-A<sup>Cnp1</sup> was detected in 2/5 of the colonies analysed. The p3xLM plasmid allowed a more efficient CENP-A<sup>Cnp1</sup> assembly since 5/5 colonies analysed clearly recruited CENP-A<sup>Cnp1</sup>.

### 3.2.3 The 3xLM sequence is able to assemble a functional centromere in wild type cells

Generally, when a plasmid containing full length cc2 flanked by outer repeats is transformed into wild-type cells, heterochromatin assembles on the outer-repeats and this triggers CENP-A<sup>Cnp1</sup> establishment on cc2 (Folco et al., 2008a). The H3K9R strain used in the assay presented above is an efficient tool for the rapid analysis of the minimal centromeric sequence requirement to recruit CENP-A<sup>Cnp1</sup>. However, even though H3K9R strain is able to assemble CENP-A<sup>Cnp1</sup>, it is devoid of heterochromatin since it is not able to methylate lysine 9 of histone H3 resulting in weak sister chromatin cohesion and minichromosome instability. In addition, it is not clear if CENP-A chromatin or a functional kinetochore is assembled in H3K9R mutants. To assess the ability of the LM fragment to assemble functional centromeres in wild type cells, the LM sequence was cloned into a plasmid containing 2.1 kb of heterochromatic elements (pMC21). The LM fragment was multimerized in 2xLM (4 kb) and 3xLM (6 kb) tandem repeats and the constructs transformed into a wild type strain. The establishment of a functional centromere was assessed by using the colony colour sectoring assay. The untransformed wild type with the nonsense mutation on the ade6 gene (ade6-704) plated on 1/10<sup>th</sup> adenine medium forms red colonies while an ade6<sup>+</sup> strain forms white colonies. The minichromosome utilised here possesses a sup3e tRNA, a nonsense suppressor able to complement the mutation ade6-704. Therefore, a plasmid that assembles a functional centromere is retained in cells and forms white colonies. The presence of red sectors within the white colonies indicates the level of mitotic stability of the plasmid. In contrast, a mitotically unstable plasmid is rapidly lost in non-selective rich medium and the colonies appear red coloured (Figure 3-3).
Figure 3-3. Schematic representation of the establishment assay.

The untransformed wild type (with the nonsense mutation ade6-704) plated on 1/10th adenine plates always forms red colonies while an ade+ strain forms white colonies. In our assay, the minichromosome possesses a sup3e tRNA, a nonsense suppressor able to complement the mutation ade6-704. Therefore, a plasmid with a functional centromere is retained in cells and forms white colonies and the presence of red sectors within the white colonies indicates the level of mitotic stability of the plasmid. In contrast, a mitotically unstable minichromosome is rapidly lost in non-selective rich medium and the colonies appear red coloured.
Chapter 3: Investigation of the minimal sequence requirements

The establishment of CENP-A\textsuperscript{Cnp1} chromatin and the acquisition of mitotic stability can be slow and the segregation function of the minichromosomes in wild-type cells is relatively low (Steiner and Clarke, 1994). 11% of cells transformed with a plasmid containing 5kb of outer repeats and 8.5kb of cc2 (pcc2-K") were able to establish a functional centromere immediately after transformation as indicated by the presence of colonies that exhibit a white-red sectored colour. In contrast, the plasmid containing just 2.1 kb of K-repeats and 8.5 kb of cc2 (pMC22) established a functional kinetochore after being grown for about 60 generations on selective medium (Figure 3-4B). However, 3% of cells containing the tandem 3xLM sequence showed ability of assembly a functional centromere while the 1xLM and 2xLM were not able to form a kinetochore just after transformation (Figure 3-4A-B).

To measure the stability of these minichromosomes with established centromeres, cells were grown on selective medium and then plated on non-selective rich medium with limited amount of adenine. A plasmid with functional and stable centromere is generally retained through multiple cell divisions generating white coloured colonies. Moreover, if during cell division the minichromosome is lost, these cells will create red sectors within the white colonies. The level of mitotic stability could be therefore measured by counting the number of colonies with half sectors. In a minichromosome loss assay, the pH-3XLM vector containing a functional kinetochore is mitotically stable and comparable to pMC22 (Figure 3-4B). In fact, 27% cells transformed with pH-3xLM showed minichromosome loss during cell cycle, a number that is comparable to the loss of the plasmid containing full length of cc2 (22% minichromosome loss).

In those colonies in which functional centromere appeared to be established in the minichromosome, ChIP was performed to determine the presence of kinetochore proteins. To distinguish endogenous centromere from minichromosomes, the strain used for the establishment assay possessed a silent his3\textsuperscript{+} gene inserted at the Sphl site in the endogenous cc2. CENP-A\textsuperscript{Cnp1} and the kinetochore protein CENP-C\textsuperscript{Cnp3} were equally present on the control plasmids (pcc2-K" and pMC22) as well as in the plasmid containing the 3XLM repeats indicating that the minimal central core region is able to assemble a functional kinetochore in wild type cells when flanked by heterochromatic repeats (Figure 3-4C).
Figure 3-4. pH-3xLM is able to establish a functional centromere in wild type.

A) Establishment assay in wild type cells of a minichromosome containing 5 kb of heterochromatic repeats (yellow box) and cc2 full length (8.5 kb) on the left and and 2.1 kb of heterochromatic repeats and 3xLM (6 kb) on the right. White colonies with red sectors indicate formation of functional centromeres. B) The minichromosomes were transformed into wild type cells and the numbers of colonies with established centromere just after transformation were counted (left). To test the stability of the minichromosomes with established centromeres, cells were grown on selective medium and then plated on non-selective medium containing low adenine. The numbers of half sectored colonies were counted (right). C) ChIP for CENP-A<sup>Cnp1</sup> and CENP-C<sup>Cnp3</sup> on minichromosomes containing 5 kb of heterochromatic repeats (pcc2-K<sup>"</sup>) or 2.1kb (pMC22 and 3xLM) and full length cc2 (pcc2-K<sup>"</sup> and pMC22) and 3xLM. Primer pairs specific for centromere 1/3 (endogenous), actin (act) and plasmid were utilised.
3.2.4 A second 2 kb-repeat sequence from cc2 is able to confer centromeric function when inserted into a plasmid.

Previous studies suggested that the central core 2 sequence shows functional redundancy (Baum et al., 1994). In the experiments above, this property was confirmed by dissecting cc2 sequence using minichromosome assays. In addition, a 2 kb region from cc2 (LM) was able to assemble a functional centromere in wild type cells when multimerised (3xLM) and flanked by heterochromatin, indicating that the LM sequence possesses all the properties required for centromere establishment. To test if the ability to form a functional centromere is specific of the LM sequence, a second 2 kb region from the cc2 (OP) was multimerized as direct repeats to create 3xOP (6kb) (Figure 3-5A). The 3xOP DNA was cloned into a plasmid containing heterochromatic repeats and subsequently transformed into wild type cells. The establishment of a functional centromere was assessed by using the colony colour sectoring assay. The 3xOP sequence was able to establish a functional centromere as shown by the appearance of red sectors formed (Figure 3-5B). However, the 3xOP sequence is less efficient compared to the 3xLM plasmid since only 0.5% of cells transformed with this plasmid were able to establish a functional centromere (Figure 3-5C). This result may be explained by a reduced number of sequence-encoded features present in the OP sequence relative to LM. Therefore, because of its higher efficiency, the LM sequence rather than OP was chosen for further studies on the role of the DNA sequence in influencing CENP-A\textsuperscript{Cnp1} assembly.

3.3 DISCUSSION

In this chapter, I have described the identification of a minimal central core sequence able to recruit CENP-A\textsuperscript{Cnp1}. The analysis presented lead to the conclusion that there is no specific region of central core 2 that is absolutely necessary for establishing CENP-A\textsuperscript{Cnp1} chromatin. Instead, the central core sequence seems to be modular and consists of elements many of which are capable of promoting CENP-A\textsuperscript{Cnp1}. This confirms what has been previously proposed regard the functional redundancy of the cc2 sequence (Baum et al., 1994). Here, deletion of the characterised minimal LM sequence was found to not affect CENP-A\textsuperscript{Cnp1} establishment nevertheless it contains the required features to induce CENP-A\textsuperscript{Cnp1} deposition. In addition, the efficiency of CENP-A\textsuperscript{Cnp1} recruitment seems to correlate with the length of the central core sequence present on the minichromosome (Figure 3-1D-E, H). It is possible that shorter sequences are unable to induce the assembly
Figure 3-5. pH-3xOP is able to establish a functional centromere in wild type. A) Schematic representation of \textit{cc}2 full length (up) and plasmid containing heterochromatic repeats (yellow box-Het) and 3 repeats of the OP fragment (down). The position of the OP sequence relative to the entire \textit{cc}2 is indicated by the black arrow in the upper panel. B) Establishment assay of a plasmid containing 3xOP transformed into wild type White colonies with red sector indicate functional centromere formation. C) percentage of cells that assemble a functional kinetochore on a minichromosomes just after transformation.
of CENP-A<sup>Cnp1</sup> nucleosomes because a certain threshold density of some feature is required and therefore the length of the central core sequence may be a critical parameter for CENP-A<sup>Cnp1</sup> assembly. Alternatively, these observations may be influenced by the centromeric DNA ability to stochastically acquire centromeric function (Steiner and Clarke, 1994). In fact, when the pH-3xLM plasmid was transformed and tested for its ability to establish a functional centromere, it was shown that the LM region contains all the features required to promote kinetochore formation when placed on a episomal plasmid (Figure 3-4). However, when the plasmid with 1xLM or 2xLM were analysed, stable centromere formation could not be detected just after transformation of the plasmid. In previous studies, it was shown that even the acquisition of a functional centromere by minichromosomes containing a large part of cen2 (>8.5 kb) can be slow and such plasmids adopt either a “stable” or “unstable” segregation states when transformed into cells. This supports the involvement of epigenetic mechanism in the establishment of a functional centromere (Steiner and Clarke, 1994). In the analysis presented here, it was possible to detect a functional centromere in 11% of cells transformed with the plasmid containing 8.5 kb of the cc2 sequence and 3% of cells transformed with pH-3xLM plasmid (Figure 3-4). Longer periods may be required to detect the formation of functional centromere in a population of cells with the pH-1xLM and pH-2xLM constructs. However, the frequency of such process may be low and thus barely detectable. If functional centromeres can not be found after many generations, it would support that the density of cis-acting element is below the threshold required to trigger CENP-A<sup>Cnp1</sup> deposition.

In conclusion, I have identified a 2 kb minimal sequence that can assemble a functional centromere on a plasmid-based minichromosome. The characteristics and the properties of this sequence are investigated further in the following chapters.
CHAPTER 4

4  GENERATING FISSION YEAST WITH NO CENTRAL CORE DOMAIN AT CENTROMERE 2 TO ALLOW ANALYSES OF SEQUENCE REQUIREMENTS FOR CENP-A ESTABLISHMENT

4.1  INTRODUCTION

Although they possess different sequences, all three fission yeast centromeres share the same overall arrangement of DNA elements. Each centromere has a central core (cc) flanked by a large inverted repeat (~18 kb with >98% identity at cen1) composed of inner most repeats (imr) and outer repeats (otr); containing the dg/dh aka K/L elements) (Pidoux and Allshire, 2004; Steiner et al., 1993; Takahashi et al., 1992). Moreover, the three centromeres share a similar organisation of chromatin domains: the entire cc and part of the imr form the central domain, which assemble nucleosomes containing the histone H3 variant CENP-A$^{Cnp1}$. In contrast, heterochromatin is formed over the otr repeats where convergent non-coding transcripts result in RNAi-directed H3K9 methylation by the methyl-transferase Clr4 (reviewed in Lejeune et al., 2011). The central core regions of cen1 and cen3 are 99% identical over ~ 4 kb (known as the TM element). The equivalent central region of cen2 is more divergent with only 48% identical to TM (Pidoux and Allshire, 2004; Wood et al., 2002). Since the cc2 sequence is essentially unique and can be easily distinguished from cc1/3 it provides the most convenient kinetochore region for dissecting its sequence requirements.

To identify key sequences required to direct functional kinetochore assembly, several fragments from cc2 were analysed in Chapter 3 for their ability to recruit CENP-A$^{Cnp1}$ on a plasmid-based minichromosome. However, the analyses presented were hindered by the presence of endogenous central domain sequence at cen2. For this reason, the analyses performed were limited to the use of primers pairs in ChIP-PCR assays that allow the centromere sequences of the endogenous and minichromosome to be distinguished. However, these primer pairs do not allow more comprehensive analyses of the cc2 sequences on minichromosomes. To specifically dissect the features associated with the cc2 sequences on the plasmid-based minichromosome, I therefore devised strategies to remove the central domain from endogenous cen2.
In this chapter, the creation of *S. pombe* cells that lack cc2 sequences at endogenous cen2 are described along with the analyses of minichromosomes introduced onto this background.

### 4.2 RESULTS

#### 4.2.1 Removing the central core 2 from endogenous cen2

To construct strains that lack most of the cc2 region, two different strategies were devised. Initially, traditional homologous recombination was used to directly replace the cc2 region of cen2 with the equivalent cc1 region from cen1 (see Diagram in Figure 4-1). 1 kb PCR fragments from imr2L and cc2 were cloned on the left and right sides, respectively, of a 5.5 Kb fragment containing the cc1 region from cen1 in a plasmid (Figure 4-1A). The resulting imr2L-cc1-cc2 fragment was then released by digestion with specific restriction enzymes and transformed into a strain bearing an ura4+ marker gene inserted in the middle of endogenous cc2 (SphI site). Centromeric chromatin is silent and it is known that loss of heterochromatin can allow increased frequency of homologous recombination, even within the central domain. Therefore, to facilitate recombination between the imr2L-cc1-cc2 fragment and the central domain of cen2 centromeric chromatin cells lackingClr4 methyltransferase were transformed (clr4Δ). Following transformation, cells were grown at 36°C to favour expression of the ura4+ marker at the centromere and plated in on medium containing the counter-selective drug FOA (5-fluoro-orotic acid) to select for transformants that had lost the ura4+ gene from cen2 (Figure 4-1B). In the presence of FOA, the expression of ura4+ is toxic for the cells and as a result, only cells that have deleted the ura4+ gene, and also the cc2 sequence, should form colonies. However, after several attempts, no FOA resistant transformants were obtained in which cc2 was replaced by cc1. The failure of this strategy may be explained by the fact that cc2 was replaced with naked cc1 DNA provided to cells in the transformation buffer. Since this DNA lacks any trace of CENP-A<sup>Cnp1</sup>, and perhaps the signals that normally ensure CENP-A<sup>Cnp1</sup> assembly, transformants with the correct insertion even may be immediately lethal.

To work around this problem a second strategy was implemented. As described below, this approach includes an intermediate step where the cc1 sequences are first inserted into cc2 at endogenous cen2 and thereby give the opportunity for CENP-A<sup>Cnp1</sup> chromatin to assemble on the resident cc1 DNA (Figure 4-2 - cc2:cc1).
Chapter 4: Deletion of endogenous central core 2

Figure 4-1. Schematic representation of replacement of cc2 with cc1 by traditional homologous recombination. A) Construct used for homologous recombination created by cloning 1 kb of the *imr2L* and 1 kb of cc2 sequence on the left and right hand side respectively of the cc1 sequence (TM region). B) Cells containing a ura4 marker gene at the cc2 region were transformed with the construct described in (A). Cells were grown at 36°C to favour expression of the *ura4*+ marker at the centromere and plated in on medium containing the counter-selective drug FOA (5-fluoro-orotic acid) to select for transformants that had lost the *ura4*+ gene and therefore cc2.
Figure 4-2. Schematic representation of replacement of cc2 by insertion of cc1 sequence followed by deletion of cc2. A) Construct used for insertion of cc1 and created by cloning 1 kb of cc2 sequence surrounding cc1 (cc2L-cc1-cc2R). B) The process of cc2 replacement involved a first step of integration of cc1 sequence (cc2:cc1) followed by replacement of cc2 with a construct containing a *ura4* marker gene (cc2Δ::cc1+ura4). In the last step, the *ura4* gene is deleted (cc2Δ::cc1).
Subsequently additional steps allow the deletion of cc2 sequences from endogenous cen2 and so that it is replaced with cc1 (Figure 4-2). Below each step is described in detail.

4.2.2  Step1: Insertion of the central core sequence from cen1 into cen2

To allow incoming naive DNA to assemble CENP-A$^{Cnp1}$, cc1 was first inserted into the central core of cen2. A plasmid in which the 5.5 kb from cc1 was flanked by adjacent 1 kb regions from cc2 was created (Figure 4-2A). The cc2L-cc1-cc2R fragment was released from the plasmids and transformed into cells lacking heterochromatin ($clr4\Delta$) and with $ura4^{+}$ inserted 2 kb to the left of the central SphI site in cc2 (cc2:$ura4$). Successful homologous recombination should result in the loss of cc2:$ura4^{+}$ and insertion of 5.5kb of cc1 DNA into cc2 in its place (Figure 4-2B). FOA resistant transformants were therefore selected and the insertion of cc1 into cc2 (cc2:cc1) was confirmed by Southern analyses (see details in Figure 4-3). Probes homologous to cc1 and cc2 detect bands of the predicted size for the correct insertion of cc1 into cc2 of cen1. These analyses confirm that in the resulting FOA-resistant transformants diagnostic cc2:$ura4^{+}$ had been deleted and that 5.5 kb of cc1 had been inserted at the required position in cc2. This allowed the next step to be attempted.

4.2.3  Step2 & 3: Removal of central core 2 sequences from cen2

The cc2:cc1 (Figure 4-2C) strain contains 5.5kb from cc1 inserted on the left side of cc2 at cen2. Therefore, transformation of a DNA fragment with homology to part of cc1 and the right hand extremity of cc2 should result in the deletion of the intervening DNA by homologous recombination. To achieve this, fusion PCR was used to create a DNA fragment with homology to cc1 on the left side of the ura4 gene and a region of homology to cc2 on its right side (construct: cc1-$ura4^{+}$-cc2). Cells containing cc2:cc1 were transformed with the linear cc1-$ura4^{+}$-cc2 DNA fragment and selected on plates lacking uracil (Figure 4-2D- cc2$\Delta$:cc1-$ura4^{+}$).

The final step (Step3) was to remove this inserted $ura4^{+}$ gene along with remaining region of cc2 sequence and part of imr2R. To accomplish this, a DNA fragment with 0.5 kb homology to the cc1 DNA on the left side of the inserted $ura4$ gene was fused to a 0.5 kb region homologous to imr2R (construct: cc1-imr2R). Following transformation of cc2$\Delta$:cc1-$ura4^{+}$ cells with cc1-imr2R, FOA resistant transformants
Figure 4-3. Southern analysis of strain created in STEP1. A) Genomic DNA was extracted from wild type strain (wt), a strain containing the *ura4* gene inserted within *cc2* (*cc2:ura4*) and the strain containing the *cc1* sequence inserted at *cc2* sequence (*cc2:cc1*). The DNA was digested with *NcoI* and southern was performed using probe on *cc1* (blue) or *cc2* (red). B) The probe used in the southern recognises the TM region which is shared between *cc1* and *cc3* (Pidoux and Allshire, 2004). The expected bands after probing for the *cc1* region are at 6 kb for *cc1* and >10 kb for *cc3*. A third band at 3.6 kb is detected in *cc2:cc1* strain indicating the insertion of the *cc1* sequence in *cc2*. C) Southern using probe on *cc2*. Expected band for wild type cells is 8.5 kb. In the presence of the *ura4* gene inserted within *cc2* (*cc2:ura4*) the band is 10 kb. A 3.6 kb band was observed in the *cc2:cc1* cells confirming the proper insertion of *cc1*. 
were selected and tested by PCR for deletion of *ura4* and the *cc2* DNA between *ura4* and *imr2R*. Cells with the correct configuration were designated *cc2Δ::cc1* (Figure 4-2E). For the strains generated in Steps 2 and 3 the presence of the expected manipulated DNA configuration within cen2 was confirmed by Southern analyses with bands of the predicted sizes being detected with specific probes (see details in Figure 4-4).

### 4.2.4 Replacement of cc2 with cc1 DNA does not affect cell viability

The removal of a large part of the normal *cc2* region from cen2 and its replacement with *cc1* DNA may affect cell viability. To test this serial dilutions of wild type cells (*cc2+*) and cells lacking *cc2* (*cc2Δ::cc1*) were spotted on rich medium in presence of Phloxin B and incubated at different temperatures. Phloxin B is a red vital dye that is absorbed by cells. Metabolically active cells are able to pump it out and remain colorless while the red dye persists in dead cells. No obvious differences in cell growth or cell viability were apparent upon comparison of *cc2Δ::cc1* cells with wild-type (Figure 4-5 – Phloxin).

It is possible that the duplication of *cc1* sequences and the loss of *cc2* DNA from cen2 might affect chromosome segregation. Mutant cells with defects in centromere function are frequently sensitive to microtubule destabilising drugs such as TBZ (Thiabendazole), particularly at lower temperatures. Serial dilutions of cells with a normal or manipulated cen2 were spotted onto plates containing increasing amounts of TBZ. Cells that lack centromeric heterochromatin (e.g. *clr4Δ*) are known to have elevated frequencies of chromosome missegregation and display sensitivity to TBZ. Cells with wild-type (*cc2+*) or manipulated cen2 (*cc2Δ::cc1*) did not display any obvious increase in TBZ sensitive. As a control cells with *clr4Δ* cells with a wild type, *cc2Δ::cc1* or *cc2:cc1* cen2 were also compared. All *clr4Δ* strains exhibited increased sensitivity to TBZ but those with the *cc2Δ::cc1* or *cc2:cc1* configurations did not display increased sensitivity than *cc2+* cells. This indicates that the manipulated centromere in these strains has little or no impact on mitotic chromosome segregation, as monitored in this assay.

### 4.2.5 Assessing the ability to establish a functional centromere on a plasmid

To determine whether functional centromeres can be established in cells with *cc2* replaced by *cc1* at cen2 (*cc2Δ::cc1*), a plasmid containing outer repeat (*otr*)
Figure 4-4. Southern analysis of the deletion of central core 2 sequence. A) Genomic DNA was extracted from cells containing the cc1 sequence inserted at cc2 sequence (cc2:cc1), cells containing cc1 inserted at cc2 locus but where cc2 sequence was replaced by a ura4 marker (cc2Δ::cc1+ura4) and a strain where the cc2 was deleted (cc2Δ::cc1). The DNA was digested with BglII and SpeI and probed for cc1/3 (blue) or digested SphI and SpeI and probed for cc2 (red). B) Southern analysis using probe on cc1/3. All the strains tested showed a band at 3.8kb corresponding to the endogenous cc1/3. In the cc2 deleted strain (cc2Δ::cc1), an additional 2.8 kb band was observed and corresponding to the cc1 insertion and concurrent deletion of part of cc2 and ura4 marker inserted in the previous steps C) Southern using a probe for cc2 showed a signal of 2.8 kb in the control strains and the absence of signal in the cc2Δ::cc1 confirming the deletion of centromere 2 from its original locus, and its replacement with cc1 sequence.
Figure 4-5. Deletion of centromere 2 is not affecting centromere structure.

Spotting assay of wild type or clr4Δ cells containing centromere 2 sequence at the endogenous centromere (cc2+) or with the cc1 inserted at the cc2 (cc2:cc1) or with the cc2 sequence replaced by the cc1 (cc2Δ). Cells were grown at different temperature to assess temperature sensitivity and viability checked by growth on Phloxin B (Phloxin). Metabolically active cells are able to pump it out and remain colorless while dead cells are stained with the red dye. To assess the stability of chromosomes during cell division, the cells were spotted on plates containing increasing amount of the microtubule-destabilising drug TBZ (10-15 µg/ml). Cells with impaired centromere function (e.g. clr4Δ) show sensitivity to the drug, especially at lower temperature (25°C). Wild type cc2+ and cc2Δ are able to grow in the presence of TBZ suggesting that cc2Δ does not affect centromeres stability.
heterochromatin adjacent to 6 Kb of cen2 central core DNA (3xLM) was transformed into the cc2Δ::cc1 strain and the matching wild-type with a normal cen2 (cc2+). The frequency of establishment of a functional kinetochore on this plasmid was comparable in both the cc2Δ::cc1 and cc2+ strains. This indicates that in these establishment assays the cc2Δ::cc1 strain performs equivalently to the cc2+ strain and can therefore be used for more detailed analyses of cc2 sequences in establishing CENP-A chromatin (Figure 4-6).

4.2.6 Assessing the performance of cc2 fragments in CENP-A<sup>Cnp1</sup> establishment assays in cells lacking most of cc2 at cen2

In Chapter 3, the minimal sequence requirement for centromere formation was dissected using a minichromosome-based assay in the H3K9R mutant background to assess the establishment of CENP-A chromatin on constructs containing various parts of the central domain from cen2. However, the analysis was limited by the presence of central domain DNA at endogenous cen2. The creation of the cc2Δ::cc1 strain that lacks most cen2 central domain sequences removes this hindrance to further analyses. To comprehensively evaluate the efficiency of CENP-A<sup>Cnp1</sup> recruitment by the various cc2 containing minichromosomes analysed previously, the same assays were performed in cells that lack endogenous cc2 (cc2Δ::cc1). First, the cc2Δ::cc1 was placed on a background in which only a single wild type or H3K9R mutant (lysine 9 mutated to arginine) histone H3 gene is retained (H3* and H3K9R; described in Chapter 3). The plasmids of interest were then transformed into these cc2Δ::cc1 H3* and H3K9R strains. Although some sequences from the central domain are still present at endogenous cen2 (left side imr1L and cc2; fragments J, K, R Figure 4-7A), the deletion of most of the natural central domain from cen2 (fragments L-Q) allows more detailed analysis of specific sequences across cc2 in the context of the minichromosome. Primer pairs were designed to detect the PCR products from the L, M, N, O & P regions of cc2 (Figure 4-7A). Consistent with previous analyses (Chapter 3, Figure 3-1), when the pcc2 minichromosome that contains a complete cen2 central domain (8.5 kb; fragments J-R) and lacks flanking outer heterochromatic repeats, was transformed into H3+ wild-type cells, CENP-A<sup>Cnp1</sup> was not detected on any of the regions of plasmid borne central core 2 (Figure 4-7C) but histone H3 remains in place (Figure 4-7D). In contrast, when pcc2 was transformed into the H3K9R mutant, CENP-A<sup>Cnp1</sup> was detectable across all regions (L-P) at levels that were comparable to those at
Figure 4-6. Central core 2-deleted strain can establish a functional kinetochore on a minichromosome. Minichromosome establishment assay in wild type (cc2+) and cc2-deleted strain (cc2Δ:cc1). A plasmid containing part of the outer repeat (2.1 kb) and the 3xLM from the cc2 sequence was transformed into yeast. The presence of white colonies with red sectors indicates the establishment of a functional kinetochore in the plasmid. The cc2Δ:cc1 strain is able to establish functional centromeres.
Figure 4-7. Analysis of full-length cc2 on a plasmid in H3K9R-cc2Δ cells. A plasmid containing full length cc2 was transformed into cells lacking part of endogenous cc2 (cc2Δ::cc1) and expressing only one histone H3 gene wild type (wt*) or mutated H3K9R as previously described in chapter 3. A) Although part of the endogenous central domain 2 is still present at endogenous cen2 (fragments J, K, R), the absence of most of the central domain sequence (fragments L-Q) allows a more detailed analysis across most of cc2 when inserted in the plasmid. ChIP analysis with anti-H3 and anti-CENP-A<sup>Cnp1</sup> antibodies was performed on wt* (grey) and H3K9R (green) cells. B) Levels of H3 and CENP-A<sup>Cnp1</sup> on cc1/3 sequences at endogenous cen1/3. C) Levels of CENP-A<sup>Cnp1</sup> on plasmid borne central domain 2 on regions L M N O P. D) Levels of H3 the same plasmid (n=3).
endogenous centromeres (Figure 4-7B). The enrichment level suggests that CENP-A<sup>Cnp1</sup> is distributed uniformly across the central domain of pcc2 (Figure 4-7C). Furthermore, the levels of histone H3 associated will all regions tested on the plasmids (L-P) are consistently lower compared to that detected on pcc2 in the H3* wild type. This suggests that H3 nucleosomes have been replaced with CENP-A<sup>Cnp1</sup> nucleosomes in the H3K9R* strain (Figure 4-7D).

The same assay was performed using the plasmids containing different regions from the central domain of cen2 following their transformation into H3* and H3K9R strains that lack most of the endogenous central domain at cen2 (cc2Δ::cc1). As expected, H3* cells expressing wild-type histone H3 were unable to establish CENP-A<sup>Cnp1</sup> on any of the plasmids utilised. In contrast, in H3K9R cells, CENP-A<sup>Cnp1</sup> was clearly enriched on the different plasmid borne fragments of cc2, however, the level of enrichment relative to the act1 locus was dependent on the centromeric DNA contained in the plasmid (Figure 4-8). These data therefore confirm and extend previous analyses and indicates that there is no particular ‘magic’ sequence within the central domain that is necessary and sufficient for attracting CENP-A<sup>Cnp1</sup>. Interestingly, deletion of 1 kb (fragment N) from the full-length central domain appears to decrease the level of CENP-A<sup>Cnp1</sup> deposition on the remaining central domain sequences (L M O P) on the introduced plasmid (compare Figure 4-8A with Figure 4-7C). This suggests that the total size of the plasmid-borne central domain DNA may influence the efficiency of CENP-A<sup>Cnp1</sup> deposition. However, it can not just be the total length since by simply inverting the right half of the central domain sequence the level of CENP-A<sup>Cnp1</sup> deposited on the same sequences appears to increase (compare LKLMOPQR with JKLMRQPO; Figure 4-8A & B). However, direct comparison between plasmid bearing a single copy of the LM sequence (p1xLM) to plasmids carrying two or three copies of LM arranged as head to tail tandem repeats (p2xLM and p3xLM) again shows that the level of CENP-A<sup>Cnp1</sup> deposited correlates with the total length of the central domain derived DNA (Figure 4-8F). The original limited analyses of plasmids suggested that a single LM element was unable to attract any CENP-A<sup>Cnp1</sup>, however, by using this cc2Δ::cc1 strain it is clear that low levels of CENP-A<sup>Cnp1</sup> can be detected on the internal part of the LM fragment. It is evident from the direct comparison of 1xLM, 2xLM and 3xLM plasmids that increasing the number of LM elements results in a stepwise increase in the deposition of CENP-A<sup>Cnp1</sup> on the three plasmids (Figure 4-8F). This suggests that increasing the size of the total region of central domain sequence increases the probability and efficiency of CENP-A<sup>Cnp1</sup> deposition.
Figure 4-8. Analysis of the level of CENP-A\textsuperscript{Cnp1} on plasmids containing fragments of cc2 sequence. Plasmids containing different regions of cc2 were transformed into a strain lacking part of endogenous central core 2 (cc2Δ::cc1) and expressing one copy of H3 wild type (wt\textsuperscript{*}) or mutated H3K9R. Quantification for CENP-A enrichment in wild type (grey)(n=3) and H3K9R (green)(n=3) was calculated relative to the actin gene (act1\textsuperscript{+}). A) Plasmid with deletion of 1 kb of cc2 from its internal region (N); B) Plasmids containing the first half of cc2 (J to M- 4 kb) and the second half deleted of the N fragment (O to R- 3.5 kb) inverted in the orientation relative to the native sequence; C-D) Plasmids containing the “halves” of cc2 sequence (J-M and N-R); E) Plasmid with deletion of part of the central region (fragment L to N); F) Analysis of plasmids containing a monomer of LM or two repeat in tandem (2xLM) or three LM repeats (3xLM). Quantification of the CENP-A levels at the endogenous centromeres can be found in Appendix 1.
If during the assembly of CENP-A\textsuperscript{Cnp1} chromatin on central domain DNA each canonical histone H3 is replaced by CENP-A\textsuperscript{Cnp1} with a 1:1 ratio, then following the transformation of these plasmid into H3K9R less H3 should be detected on the centromeric DNA relative to the same plasmid transformed into H3* wild-type cells which are unable to deposit CENP-A\textsuperscript{Cnp1}. Indeed, plasmids containing the full-length central domain (pcc2; Figure 4-7) exhibit significantly reduced levels of histone H3 associated with centromeric DNA in H3K9R cells compared with H3* cells. This indicates that when CENP-A\textsuperscript{Cnp1} chromatin is assembled on the full length central domain plasmid H3 nucleosomes are replaced with CENP-A\textsuperscript{Cnp1} (Figure 4-7D). However, when plasmids bearing smaller centromeric DNA fragments were analysed by ChIP essentially no reduction in H3 levels are observed in H3K9R relative to H3* cells, despite the assembly of CENP-A\textsuperscript{Cnp1} nucleosomes on these sequences (Figure 4-9).

4.3 DISCUSSION

In chapter 3, in order to identify crucial sequences for centromere function, I described the analysis of several fragments from central core 2 and their ability to recruit CENP-A\textsuperscript{Cnp1} when inserted into plasmids. However, this study was complicated by the presence of endogenous cc2 at centromere 2. For this reason, most of the results were produced using specific combinations of primers that could distinguish between endogenous centromere and minichromosome. However these primers were often not optimal for a full characterisation of cc2 sequence. Moreover, the large region of homology between plasmid and centromeric sequence led to homologous integration of the plasmid, adding a further layer of complication to the analyses. To allow a more detailed dissection of the features involved in CENP-A\textsuperscript{Cnp1} establishment across cc2, the endogenous cc2 region was partially replaced by the cc1 sequence. The resultant cc2Δ::cc1 strain possesses a copy of the 5.5 kb region (TM element) at each of the three different centromeres. However, the cells are viable at different temperatures and centromere structure seems not to be affected by the microtubule depolymerising drug TBZ (Figure 4-5). Moreover, it is possible to obtain viable spores after crossing wild type cells cc2+ with the cc2Δ strain, suggesting that centromeric sequence might have a marginal role in
Figure 4-9. Levels of H3 on plasmids containing fragments of cc2 sequence.
Plasmids containing different regions of cc2 were transformed into a strain lacking part of endogenous central core 2 (cc2Δ::cc1) and expressing one copy of H3 wild type (wt*) or mutated H3K9R. Quantification for H3 enrichment in wt* (grey)(n=3) and H3K9R (green)(n=3) was calculated relative to the actin gene (act1*). A) Plasmids with deletion of 1 kb of cc2 from its internal region (N); B) Plasmids containing the first half of cc2 (J to M- 4 kb) and the second half deleted of the N fragment (O to R- 3.5 kb) inverted in the orientation relative to the native sequence; C-D) Plasmids containing the “halves” of cc2 sequence (J-M and N-R); E) Plasmid with deletion of part of the central region (fragment L to N); F) Analysis of plasmids containing a monomer of LM or two repeat in tandem (2xLM) or three LM repeats (3xLM). Quantification of the H3 levels on the endogenous centromeres can be found in Appendix 2.
Chapter 4: Deletion of endogenous central core 2

chromosomes paring during meiosis. However, statistical analyses on spore viability are needed to completely exclude defects that may occur during meiosis.

Deletion of the central core 2 region has improved the analysis of plasmids containing different fragments of cc2. In this strain, it was possible to directly compare the same regions from different constructs using same primer pairs. Thus, a plasmid containing full-length cc2 is able to efficiently recruit CENP-A\textsuperscript{Cnp1} nucleosomes and it is comparable to the endogenous centromere (Figure 4-7C). As a consequence, the H3 levels on this plasmid are decreased when CENP-A\textsuperscript{Cnp1} is deposited, indicating that when CENP-A\textsuperscript{Cnp1} chromatin is assembled on the full length central domain plasmid H3 nucleosomes are replaced with CENP-A\textsuperscript{Cnp1} (Figure 4-7D). However, although the plasmids bearing a sub-fragment of cc2 were able to recruit a moderate level of CENP-A\textsuperscript{Cnp1}, the enrichment for H3 observed in H3K9R was similar to that H3* cells (Figure 4-9). The reason for the observed difference between full-length cc2 and the cc2 sub-fragment with respect to H3 levels is unknown. But a possible explanation could be that the amount of CENP-A\textsuperscript{Cnp1} incorporated in the sub-fragment is not actually very high, and may perhaps vary within the cell population or the plasmid population. In full-length cc2, H3 levels are similar to a control locus in H3* cells, and very low in H3K9R cells due to high CENP-A\textsuperscript{Cnp1}. But, because population-wide CENP-A\textsuperscript{Cnp1} is only moderate for the cc2 sub-fragment, there is only a small (and therefore hard to detect) reduction in H3 from control levels in H3* compared to H3K9R.

Alternatively, in H3* cells (and wt cells), cc2 is packaged in H3 nucleosomes but could also contain some nucleosome-free regions (NFRs). In an H3K9R mutant, it may be that these NFRs are the places where CENP-A\textsuperscript{Cnp1} is incorporated initially, followed by local expansion of the CENP-A\textsuperscript{Cnp1} domain to replace H3. If the cc2 sub-fragment used happened to consist of mostly H3 occupied sites and little/no NFR, there would proportionately be less opportunity for Cnp1 incorporation and expansion in H3K9R cells, and therefore the appearance of no reduction in H3 levels in H3K9R vs H3* for the cc2 subfragments. The small amount of CENP-A\textsuperscript{Cnp1} incorporation (zero compared to moderate) would be still represent a significant difference in H3* vs H3K9R.

Thus, the cc2△::cc1 strain provides a more sensitive assay for CENP-A\textsuperscript{Cnp1} establishment than cells in which the analyses is limited by the presence of a normal cen2 (Chapter 3). This strain together with cc2 containing minichromosomes will give the opportunity to analyse in more details the transition from H3 to CENP-A\textsuperscript{Cnp1} chromatin state.
CHAPTER 5

5 A ROLE FOR NUCLEOTIDE COMPOSITION IN SPECIFYING REGIONAL CENTROMERE

5.1 INTRODUCTION

The position of nucleosomes on DNA is determined by the action of chromatin remodelers during biological processes in addition to the primary DNA sequence itself since certain sequence can more easily wrap around and accommodate a nucleosome (Kaplan et al., 2009; reviewed in Segal and Widom, 2009b; Wippo et al., 2011). However, the relative contribution of active process versus primary DNA sequence to the final locations of nucleosomes remains a matter of debate. Several studies have indicated that the underlying DNA sequence patterns direct nucleosome positioning in vivo (Kaplan et al., 2009; Segal et al., 2006). Indeed, it is known that the composition and mechanical properties of a sequence, such as its ability to bend and twist, can determine if a particular sequence will more easily wrap around a nucleosome (Widom, 1992). Comparative analyses of the positions occupied by nucleosomes in vivo versus those taken up following in vitro assembly on DNA, have demonstrated that nucleosomes can wrap DNA molecules with different affinities and, apart from promoters and highly transcribed regions, the approximately up to 20% of in vivo nucleosome locations correlate with those observed/predicted in vitro (Kaplan et al., 2009; Lantermann et al., 2010). Such analyses suggest that the underlying DNA sequence plays an important role in positioning nucleosomes in vivo (Kaplan et al., 2009). Analyses of the positions adopted by nucleosomes on DNA has allowed computational algorithms to be developed that predict the probability that a nucleosome will be positioned on a particular DNA sequence (Kaplan et al., 2009).

One of the major characteristics shared between centromeres of different organisms is the particular elevated A/T nucleotide content (Kanizay and Dawe, 2009). In many organisms, centromeres are assembled in regions that contain long stretches of poly(dA:dT) often arranged in tandem repeats. In fission yeast centromeres, the central domain consists of a region enriched in A/T nucleotides (72% A/T) compared to the overall A/T content of the genome (64% AT) (Wood et al., 2002). Interestingly, poly(dA:dT) elements are known to possess nucleosome-excluding properties and in S. cerevisiae such sequences are found at promoters where they
generate nucleosome depleted regions (NDRs) and aid promoter definition (reviewed in Jiang and Pugh, 2009; Yuan et al., 2005). These NDRs are required at promoter regions to generate a zone of nucleosome depletion, and suggests that nucleotide composition might promote nucleosome instability and exchange (Segal and Widom, 2009a; Segal et al., 2006). However, in S. pombe although NDRs are found at promoters they are not correlated with such dA:dT tracts (Lantermann et al., 2009). Since the sequences associated with centromeres generally have a high AT content it is possible that H3 nucleosomes have a lower affinity for such sequences and this may favour the replacement of H3 nucleosome with CENP-A containing nucleosomes which may in turn have evolved a greater affinity for such centromere DNA sequences.

In this chapter, I explore how nucleotide composition affects the establishment of CENP-A chromatin on centromeric DNA sequences in vivo.

5.2 RESULTS

5.2.1 A random A/T rich sequence can not establish CENP-A$^{\text{Cnp1}}$ in a H3K9R mutant.

To test if the property of A/T-richness itself is sufficient to assemble a 2 Kb random synthetic AT-rich sequence (SynAT) was designed in silico to have an A/T content comparable to that of the endogenous LM region from the central domain of cen2 (70% A/T for SynAT versus 72% for the LM fragment). This SynAT sequence was commercially synthesised and cloned into a plasmid (pMC1) that lacks heterochromatic repeats but that allows selection in S. pombe cells. As with the native LM centromeric fragment, plasmids containing the 2 kb monomer and direct repeats of two (2xSynAT- 4kb) and three (3xSynAT- 6kb) copies of the SynAT sequence were transformed into the H3K9R strain to test for the ability to assemble CENP-A$^{\text{Cnp1}}$ chromatin in the absence of heterochromatin. Anti-CENP-A$^{\text{Cnp1}}$ ChIP was then performed on five transformants from each plasmid. Unlike the native LM plasmids, which always attract CENP-ACnp1 when transformed into H3K9R cells, CENP-A$^{\text{Cnp1}}$ was not enriched on the SynAT constructs. This demonstrates that simply having a similar level of AT content as the central CENP-A domain sequences from fission yeast centromeres is not sufficient to recruit CENP-A$^{\text{Cnp1}}$ (Figure 5-1).
Figure 5-1 SynAT can not establish CENP-A<sup>Cnp1</sup> in H3K9R cells. To test if the nucleotide content is sufficient to recruit CENP-A<sup>Cnp1</sup>, a 2 Kb random synthetic AT-rich sequence (SynAT) was designed in silico to have an A/T content comparable to that of the endogenous LM region from the central domain of cen2 (70% A/T for SynAT versus 72% for the LM fragment). SynAT sequence was cloned in the vector that lacks heterochromatin repeats in single 2 kb monomer (SynAT) and direct repeats 2xSynAT (4kb) and 3xSynAT (6kb). The constructs were tested in the H3K9R strain and CENP-A<sup>Cnp1</sup> ChIP performed on 5 colonies. Endogenous (cc1/cc3) was used as positive control and normalise on the housekeeping gene <i>fbp1</i>. In none of the three constructs tested was CENP-A<sup>Cnp1</sup> assembly detected, indicating that SynAT sequences are unable to recruit CENP-A<sup>Cnp1</sup>.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Endogenous ChIP</th>
<th>Plasmid ChIP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 x SynAT</td>
<td>cc1/cc3 T/AT</td>
<td>cc1/cc3 T/AT</td>
</tr>
<tr>
<td>2 x SynAT</td>
<td>cc1/cc3 T/AT</td>
<td>cc1/cc3 T/AT</td>
</tr>
<tr>
<td>3 x SynAT</td>
<td>cc1/cc3 T/AT</td>
<td>cc1/cc3 T/AT</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Constructs</th>
<th>n° of colonies with CENP-A deposited on the plasmid</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 x SynAT</td>
<td>0/5</td>
</tr>
<tr>
<td>2 x SynAT</td>
<td>0/5</td>
</tr>
<tr>
<td>3 x SynAT</td>
<td>0/5</td>
</tr>
</tbody>
</table>
However, since this SynAT sequence is completely random it is possible that it lacks features such as the basic underlying nucleotide patterns that are known to favour the wrapping of DNA around nucleosomes (Jiang and Pugh, 2009). The nucleosome-DNA interaction model developed by Kaplan et al. (2009) is based on *in vitro* measurement of preferences that nucleosomes have for particular DNA sequences and significantly correlates with *in vivo* nucleosome occupancy. Both *in vitro* and *in vivo*, genomic DNA is characterised by 10-bp periodicities of dinucleotides along the DNA-nucleosome interface that is translated into defined and sharp peaks of nucleosome occupancy (Kaplan et al., 2009). The Kaplan’s algorithm predicts nucleosome peaks within the CENP-A<sup>Cnp1</sup> domain that match those of mapped *in vivo* by ChIP-seq (Figure 5-2A, Lando et al. 2012). This indicates that similar fundamental rules must govern the sequences that can wrap around H3 or CENP-A<sup>Cnp1</sup> nucleosomes. cen1 and cen2 represent a good example of nucleosome positioning (Figure 5-2B, -C).

To determine the predicted nucleosome positioning on the 2 kb SynAT sequence compared to 2 kb native LM sequence, the algorithm developed by Kaplan et al. (2009) was utilised (Figure 5-3). The resulting comparison shows that SynAT has a very different predicted nucleosome pattern to that of the endogenous LM sequence. This raises the possibility that SynAT may be unable to assemble CENP-A<sup>Cnp1</sup> chromatin because it lacks the ability to assemble nucleosomes in a pattern required for its recognition *in vivo*.

5.2.2  *A synthetic AT-rich sequence with the same predicted nucleosomes position as endogenous centromeric DNA can not establish CENP-A<sup>Cnp1</sup> chromatin.*

The fact that predicted nucleosome positions in the SynAT sequence is distinct from that of the LM sequence suggests one explanation why SynAT may not be a substrate for CENP-A<sup>Cnp1</sup> assembly. To determine if an A/T-rich sequence with a similar predicted nucleosome-positioning pattern to endogenous centromeric central domain DNA could direct CENP-A<sup>Cnp1</sup> assembly, a new synthetic sequence was designed. Moreover, because it had been shown that the GC content at fission yeast centromeres correlates with position of CENP-A nucleosomes (Song et al., 2008), this design takes into consideration the periodicity of nucleosome positions and both AT/GC content and distribution. Several synthetic sequences were designed by randomising the endogenous LM sequence in 5 bp windows. The predicted nucleosome position on these synthetic LM sequences was then compared to the original LM and modified in order to have a predicted nucleosome
Figure 5-2. Predicted nucleosome positions on centromeric DNA. A) Comparison between nucleosome positioning prediction (grey) created using the algorithm developed by Kaplan et al. (2009) and ChIP-seq analysis for CENP-A^{Cnp1} (Lando et al. 2012). The algorithm predicts nucleosome peaks within the central domain that match those of mapped in vivo by ChIP-seq. B) DNA sequence from centromere 1 was analysed using the prediction program. C) DNA sequence from centromere 2 was analysed using the prediction program developed by Kaplan et al. (2009).
Figure 5-3. Predicted nucleosome positions on centromeric DNA and SynAT is distinct. Comparison between the predicted nucleosomes positioning pattern between the LM sequence from cc2 (red) and SynAT (blue). SynAT has a very different predicted nucleosome pattern to that of the endogenous LM sequence and lacks the typical pattern of “biological” sequences.
distribution close to that of the endogenous LM sequence (Figure 5-4A). The synthetic sequence that with the best match (SynRLM) was then commercially synthesised and cloned as single copy and, 2xSynRLM and 3xSynRLM tandem repeats which were all then tested for CENP-A\(^{Cnp1}\) recruitment in H3K9R and in wild type cells. Overall the SynRLM sequences is 65.5% similar to with the native LM sequence however, despite this degree of similarity the SynRLM was unable to establish CENP-A\(^{Cnp1}\) in H3K9R or wild type cells (Figure 5-4B). This analysis demonstrates that nucleotide composition and predicted nucleosome positioning are not the key or only features of centromeric DNA that determine where CENP-A\(^{Cnp1}\) nucleosomes assemble in fission yeast. Furthermore, these analyses reveal that underlying properties of specific sequences play a role in directing CENP-A assembly in vivo.

### 5.2.3 Native centromeric DNA sequence can drive CENP-A\(^{Cnp1}\) deposition on adjacent synthetic DNA sequence.

Although SynRLM is unable to recruit CENP-A\(^{Cnp1}\), it is possible that native centromeric DNA sequences can attract CENP-A\(^{Cnp1}\) and drive its assembly onto adjacent synthetic DNA. To test this, 1 kb from the synthetic SynRLM sequence (SynL or SynM) were combined with 1kb of the native LM centromeric DNA (L or M) (Figure 5-5A). These new monomers (M+SynL and L+SynM) were cloned in three copies as a tandem repeats (Figure 5-5). These plasmids were again introduced into H3K9R and wild-type cells and tested for enrichment of CENP-A\(^{Cnp1}\) by ChIP. Interestingly, the 3xM+SynL, but not 3xL+SynM construct consistently attracted CENP-A\(^{Cnp1}\). This demonstrates that the synthetic sequence itself is not an obstacle to forming CENP-A\(^{Cnp1}\) chromatin. In addition, these analyses suggest that the 1 kb fragment from the M region, but not the L region, contains sequence information that for some reason are capable of promoting CENP-A\(^{Cnp1}\) deposition (Figure 5-4B).

### 5.3 DISCUSSION

Several studies have demonstrated that DNA sequence patterns affect nucleosome positioning in vivo since particular nucleotide tracts influence the mechanical properties of a particular DNA sequence and its ability to wrap around a nucleosome (Kaplan et al., 2009; Segal et al., 2006; Widom, 1992). In this chapter, the influence of nucleotide composition on the ability to recruit CENP-A\(^{Cnp1}\) was analysed using
Figure 5-4. SynRLM is not able to recruit CENP-A^{Cnp1}. SynRLM was designed taking into consideration both periodicity of nucleosome positions and AT/GC content distribution by randomising the LM sequence in a 5 bp window. A) The predicted nucleosome positioning of SynR-LM (blue) compared to LM sequence (red) shows a similar pattern. B) ChIP anti-CENP-A^{Cnp1} was performed on SynRLM, 2xSynRLM and 3xSynRLM transformed into H3K9R and in wild type.
A combination of endogenous and synthetic sequence can assemble CENP-A<sup>Cnp1</sup>. A) Two new monomers were created by combining 1 kb of synthetic sequence with 1 kb of the centromeric DNA from the LM region. The monomers (M+SynL and L+SynM) were multimerised in three tandem repeats and tested for the presence of CENP-A<sup>Cnp1</sup> in H3K9R strain. B) anti CENP-A<sup>Cnp1</sup> ChIP performed on 5 colonies of wild type and H3K9R cells transformed with the new constructs. The M sequence is able to drive CENP-A<sup>Cnp1</sup> deposition on the synthetic sequence.
different synthesised sequences. One of the major characteristics shared between centromeric DNA from different organisms is the fact that they tend to have a relative high A/T nucleotide content compared to most of the genome (Kanizay and Dawe, 2009). This raised the possibility that this elevated AT content is itself sufficient to selectively assembly CENP-A rather than H3 nucleosomes on centromeric DNA elements. However, the findings presented demonstrate that AT rich content alone is not sufficient to recruit CENP-A<sup>Cnp1</sup>. To take into consideration the sequence constraints that limit those that can efficiently wrap around nucleosomes, the algorithm developed by Kaplan et al. (2009) allowed a synthetic sequence that has a similar pattern of predicted nucleosomes positions to be designed by randomising the LM region from native cen2 in 5 bp windows. However, despite the maintenance of sequence similarity and predicted nucleosome positions, this synthetic SynRLM sequence was unable to assemble CENP-A<sup>Cnp1</sup> in vivo. One explanation is that the randomisation process scrambled some particularly important features encoded in the primary DNA sequence of the LM fragment and thereby eliminated the ability of this DNA to direct the assembly of CENP-A<sup>Cnp1</sup> chromatin. However, a chimeric centromeric sequence composed of synthetic sequence fused with native centromeric sequence was found to be able to assemble CENP-A<sup>Cnp1</sup> chromatin. Thus just 1 kb of endogenous centromeric sequence is sufficient to attract CENP-A<sup>Cnp1</sup> and drive its deposition on a flanking non-centromeric sequences. This findings presented reveal that centromeric sequences must possess some intrinsic characteristic that is strictly dependent on the primary DNA sequence so that alterations in the composition of that DNA (SynRLM) affect its ability to recruit CENP-A<sup>Cnp1</sup>.
CHAPTER 6

6 INVESTIGATING CANDIDATE DNA BINDING PROTEINS FOR A ROLE IN DEFINING CENTROMERE DNA.

6.1 INTRODUCTION

CENP-A containing nucleosomes play a key role in centromere function as they create the interface between centromere DNA and kinetochore proteins which mediate attachment to spindle microtubules and ensure proper chromosome segregation during cell division (Earnshaw and Rothfield, 1985; Earnshaw et al., 1989; Palmer et al., 1991; 1987). Although centromeres are assembled on rapidly evolving DNA sequences that do not share obvious homology between divergent species and several studies point to the conclusion that centromeres are epigenetically determined (Allshire and Karpen, 2008), the fact remains that CENP-A assembly in all organisms generally occurs on particular preferred sequences (e.g. alpha satellite in human cells, central domain DNA in fission yeast). Therefore, it is possible that within one species or even between different species, centromeres and neocentromeres share particular sequence features that result in their selection as sites of CENP-A deposition and kinetochore assembly.

One possibility is that centromeric DNA sequences could act as binding sites for specific centromeric proteins. At most human centromeres, canonical core alpha-satellite monomers contain a 17-bp motif recognised by CENP-B, a protein required for CENP-A recruitment during centromere establishment (Ando et al., 2002). Furthermore, alpha satellite arrays containing CENP-B boxes have been shown to promote efficient centromere and artificial chromosome formation indicating that CENP-B influences de novo centromere establishment (Ohzeki et al., 2002; Okada et al., 2007). It is therefore possible that specific DNA sequences do allow the binding of specific factors to direct the establishment of centromeric chromatin.

Once CENP-A chromatin and the kinetochore is established at that site epigenetic mechanisms could ensure its maintenance and propagation at that site irrespective of specific binding factors or the surrounding chromatin context.

A large number of proteins are involved in assembling kinetochores and several possess the ability to bind DNA or contain predicted DNA binding motifs. Such DNA
binding domains could be involved in the initial recognition of specific sequence in centromeric DNA and contribute to the establishment of CENP-A chromatin at centromeres. Indeed, it has been shown that the artificial tethering of HJURP or CENP-A\textsuperscript{CID} to DNA can lead to kinetochore assembly (Barnhart et al., 2011; Mendiburo et al., 2011).

Several observations in chapters 4 and 5, suggest that centromeric central domain DNA might contain sequence specific features that are required for promoting CENP-A\textsuperscript{Cnp1} assembly. A simple explanation for this sequence dependence might be that specific DNA binding sites in the central domain are required to recruit specific centromere proteins that mediate CENP-A\textsuperscript{Cnp1} deposition. Since several centromere associated proteins contain known or predicted DNA binding motifs, their ability to recognise central domain sequence \textit{in vivo}, prior to CENP-A assembly, was tested.

6.2 RESULTS

6.2.1 Analysis of \textit{in vivo} binding of candidate proteins with potential nucleotide binding motifs

When a plasmid containing the central domain of \textit{cen2} without adjacent heterochromatic repeats is transformed into wild type cells CENP-A\textsuperscript{Cnp1} is not deposited on the central domain sequence. This state must represent a central domain coated in H3 chromatin prior to replacement with CENP-A\textsuperscript{Cnp1} nucleosomes. Any protein that binds specifically to the centromeric DNA before the deposition of CENP-A\textsuperscript{Cnp1} should be detected on this plasmid borne central domain DNA even in wild-type cells. In Chapter 3, the 3xLM sequence was shown to assemble functional centromeres when flanked by heterochromatin. This indicates that the LM sequence possesses the required features required to attract CENP-A\textsuperscript{Cnp1} and assemble a functional kinetochore. In addition, the short length of the LM region, 2 kb compared to the 8.5 kb full-length central domain, facilitates analysis of the sequences involved. Therefore, in order to assess the presence of kinetochore proteins that might prime the central domain for the recruitment of CENP-A\textsuperscript{Cnp1}, a plasmid containing the 3xLM sequence (lacking heterochromatin repeats) transformed into wild type cells was tested for a limited set of candidate proteins analysed by ChIP. The strain utilised for these assays had endogenous \textit{cc2} replaced with \textit{cc1} to allow more specific and detailed analysis of different regions with the LM sequence on the plasmids \textit{in vivo}.
**CENP-A^{Cnp1} chaperone Scm3**

*S. pombe* Scm3 is the homologue of vertebrate HJURP which acts as a chaperone to direct CENP-A assembly. In fission yeast Scm3 is essential for the localization of CENP-A^{Cnp1} and it interacts with both CENP-A^{Cnp1} and Mis16/Mis18 (Pidoux et al., 2009; Williams et al., 2009). Scm3 exhibits similar dynamics to Mis16 and Mis18 being release from centromeres in early mitosis and reassociating in late anaphase. Interestingly, Scm3 remains associated with centromeres in mutants that disrupt CENP-A localisation (Pidoux et al., 2009). Recently Scm3 from *S. cerevisiae* has been shown to contain a DNA binding that has an a preference for AT-rich DNA, (Xiao et al., 2011). To determine if Scm3 can bind naive centromeric DNA in vivo the presence of GFP-tagged Scm3 was assessed by ChIP on the 3xLM plasmid. Although Scm3-GFP is 10x enriched on endogenous centromere relative to the act1+ gene, it is not recruited to plasmid borne 3xLM that lacks CENP-A (Figure 6-1B).

**CENP-C^{Cnp3}**

Although CENP-C is not absolutely necessary for kinetochore assembly, it is associated with functional centromeres and is known to bind CENP-A (Earnshaw et al., 1989; Fukagawa et al., 1999; Tanaka et al., 2009). Studies in several organisms have lead to the conclusion that CENP-C functions as scaffold for recruitment of other kinetochore proteins in the CCAN complex, such as the Sim4 complex in fission yeast and Mis18BP in human. In addition, it has also been proposed to mediate the interaction between the kinetochore and microtubules via the Mis12 complex (Dambacher et al., 2012; Przewloka et al., 2011; Screpanti et al., 2011; Tanaka et al., 2009). Moreover, CENP-C has also been shown to binds DNA in vitro though an AT-hook domain that preferentially binds AT-rich sequences (Du et al., 2010; Sugimoto et al., 1994; Yang et al., 1996). Thus CENP-C is a good candidate protein to prime centromeric sequences for CENP-A deposition. However, ChIP analyses did not detect CENP-C on the plasmid borne 3xLM repeats even though it is highly enriched at endogenous centromeres. Thus CENP-C does not associate with centromeric central domain DNA when they are not associated with a functional centromere (Figure 6-1C).
Figure 6-1. ChIP to test in vivo binding of centromeric proteins to central domain DNA. The capability of some kinetochore proteins to bind the LM sequence in absence of an established kinetochore was tested by ChIP. A) A plasmid containing 3xLM sequence was transformed into wild type for CENP-Cnp3 analysis and in cells expressing GFP-tagged-Scm3 or GFP-CENP-TCnp20 (note the plasmid does not contain heterochromatic repeats). The presence of the kinetochore proteins on the LM sequence was assessed by ChIP using anti-GFP or anti-CENP-Cnp3 antibody. The endogenous and the plasmid IP (L and M) were normalised relative to the act1+ gene.
Chapter 6: Candidate DNA binding proteins and centromere determination

**CENP-T\(^{\text{Cnp20}}\)**

Recently the CENP-T and W proteins were identified as a sub-complex of the kinetochore which acts in parallel with CENP-C (Hori et al., 2008). The conserved function of CENP-T proteins is to directly interact with the Ndc80 complex (Schleiffer et al., 2012). CENP-T and CENP-W are conserved between species and interact though their histone-fold domains to form a stable complex (Hori et al., 2008; Nishino et al., 2012). Together with the CENP-S and X proteins CENP-T and W form a very stable heterotetramer that appears to possess a nucleosome-like structure that can bind DNA and assemble chromatin-like structures (Nishino et al., 2012). Thus the CENP-T-W-S-X complex may form non-canonical nucleosome-like structure that could associate with centromeric sequences in advance of CENP-A.

The fission yeast homolog of CENP-T, CENP-T\(^{\text{Cnp20}}\) has not been characterised in detail, but GFP-tagged CENP-T\(^{\text{Cnp20}}\) is enriched in the central kinetochore domain of endogenous centromeres. However, CENP-T\(^{\text{Cnp20}}\)-GFP is not enriched on the 3xLM DNA carried on a plasmid in these cells indicating that it does not recognise this DNA prior to CENP-A/kinetochore assembly (Figure 6-1D).

### 6.2.2 Determining if the BAH domain of the origin recognition complex 1 protein plays a role in defining centromeres

Replication of eukaryotic DNA initiates at multiple loci, called replication origins (ARS), which are distributed throughout the genome. Individual origins initiate DNA replication at different times in S-phase, allowing DNA to be synthesised in a restricted period of time. *S. pombe* replication origins utilize long AT-rich DNA sequences (500-1500bp) and it is known that they do not conform to any particular consensus sequence (Okuno et al., 1999). One reason for this association with elevated AT elements is that the Origin of Replication Complex (ORC) binds replication origin sequences exclusively through the protein Orc4 via its long N-terminal AT-hook domain (Kong and DePamphilis, 2001).

In addition to its role during replication, in budding yeasts ORC is known to be involved in forming repressive chromatin. In *S. cerevisiae*, silencing at the silent mating-type cassettes (HML and HMR) is initiated at specific Silencer DNA elements that contain binding sites for ORC and other silencing proteins. The binding of these complexes induces the recruitment of the Sir2, Sir3 and Sir4 proteins that spread over the region to form a repressed chromatin domain. Sir1 is recruited to the mating type silencer by Orc1, the largest subunit of ORC and serves as a scaffold
for the other Sir proteins (Figure 6-2A) (Gardner et al., 1999). In addition, Sir2-dependent deacetylation of H3 nucleosome, increases the affinity for the histone binding proteins Sir3 and Sir4, which in turn recruit additional Sir2 reinforcing silencing over the region (Figure 6-2A) (Hecht et al., 1995; Hoppe et al., 2002; Rusche et al., 2002). The Sir1, Sir3 and Sir4 proteins are not widely conserved and are restricted to this form of repressive chromatin formation in budding yeasts. The histone deacetylase Sir2 is conserved between all eukaryotes and is involved in forming repressive chromatin. Interestingly, Sir3 is a paralog of Orc1 that arose from a gene duplication event in S. cerevisiae (Hickman and Rusche, 2010). Both Sir3 and Orc1 retain a conserved protein-protein interaction BAH (bromo-adjacent homology) domain that can bind histones and mediates the interaction between Orc1 and Sir1 and is required for the spreading of Sir3 through the interaction of the Sir3 BAH domain with deacetylated nucleosomes (Figure 6-2B) (Gardner et al., 1999; Hecht et al., 1995). Despite their common ancestry, Orc1 can not substitute Sir3 in forming silent chromatin and its only function in silencing is to recruit Sir1, and thus the Sir complex, through its BAH domain. The divergence of Sir3 from Orc1 suggests that prior to this gene duplication event Orc1 may have cooperated with Sir2 in forming repressive chromatin, and following this most silencing functions were restricted to Sir3, while Orc1 specialised in promoting replication (reviewed in Hickman et al., 2011). In support of this view, the related budding yeast Kluyveromyces lactis does not possess a homologue of Sir3 and the BAH domain of Orc1 is not required for replication but contributes to silent chromatin formation by binding nucleosomes in a Sir3-like manner (Hickman and Rusche, 2010). Furthermore, although deletion of the Orc1 BAH domain does not affect the viability of K. lactis, it is required to recruit the Sir2 and Sir4 proteins and allow their distribution over silenced regions (Hickman and Rusche, 2010).

In S. pombe, Sir2 is involved in the integrity of silent chromatin at the mating type locus, telomeres and centromeric repeats where it mainly deacetylates lysine 9 (K9) of H3 and thereby promotes the methylation of H3K9 and heterochromatin formation (Shankaranarayana et al., 2003). In addition to its role in forming heterochromatin, Sir2 has been found to influence the establishment of CENP-A<sup>Cnp1</sup> chromatin on minichromosomes (Allshire lab- unpublished). However, it is not known how Sir2 is recruited to the heterochromatic loci in fission yeast or how it influences CENP-A chromatin assembly over the central kinetochore domain.
Figure 6-2. Function of Sir proteins in *S. cerevisiae*. A) Orc1 together with other silencer binding proteins (Rap1-(R) and Abf1 (A)) recruit the Sir proteins (1-4) to the silencer. Sir2-dependent deacetylation of H3 nucleosomes generates binding sites for Sir3 and Sir4 which in turn recruit additional Sir2 (figure from Hickman et al., 2011). B) Conserved domains between Sir3 and Orc1. BAH: BAH domain, AT-hook: DNA binding domain; AAA+: (ATPases associated with diverse activities) domain; WH: Winged helix- predicted DNA binding domain.
Fission yeast does not possess a homologue of budding yeast Sir3, however, genome wide ChIP analysis have shown that both Orc1 and Orc4 are strongly associated with the CENP-A\textsuperscript{Cnp1} domain of centromere but replication does not appear to be initiated in these regions, suggesting a replication-independent role for ORC in these regions of centromeres (Hayano et al., 2012; Hayashi et al., 2007; Matsumoto et al., 2011). The fact that \textit{S. pombe} Orc1 contains a conserved BAH domain which is known to recruit Sir2 in \textit{K. lactis}, may suggest that the centromeric pool of ORC might act to recruit Sir2 via the BAH domain of Orc1 and promote the assembly of CENP-A\textsuperscript{Cnp1}. To test this hypothesis strains were constructed that lack the BAH of Orc1 and assayed for defects in heterochromatin and CENP-A\textsuperscript{Cnp1} chromatin integrity.

\textit{Deletion of BAH domain from Orc1 does not affect viability of S. pombe}

The \textit{orp1}\textsuperscript{+} gene encoding Orc1 in \textit{S. pombe} is essential for DNA replication and viability, therefore any manipulations that disrupt Orc1 function will cause cell lethality (Grallert and Nurse, 1996). In order to determine if the Orc1 BAH domain is involved in replication an additional copy of the gene encoding Orc1 (\textit{orp1}\textsuperscript{+}) retaining or lacking the BAH domain (\textit{orp1}-BAH\textdagger) was integrated at the \textit{leu1} locus under the control of its own promoter in a strain bearing a temperature-sensitive lethal allele of \textit{orp1} (\textit{orp1}-4; Grallert and Nurse, 1996) If the BAH domain of Orc1 is necessary for replication then its removal would result in loss of Orc1 function of the protein and \textit{orp1}-BAH\textdagger would be unable to complement loss of growth of the \textit{orp1}-4 mutant at 36°C. However, cells expressing Orc1-BAH\textdagger or wild-type Orc1 formed colonies at 36°C, thus cells expressing Orc1 without its BAH domain are completely viable and the BAH domain may have other roles independent of replication (Figure 6-3). This was confirmed by subsequently deleting endogenous gene encoding \textit{orp1}-4, again these \textit{orp1}-BAH\textdagger cells, with no other source of Orc1, were viable. These strains possessing the \textit{orp1} gene deleted from its endogenous location will be used for further experiments.

\textit{Testing Orc1-BAH\textdagger mutant for defects in heterochromatin and CENP-A chromatin integrity}

Marker gene placed within fission yeast centromeres are transcriptionally repressed and this silencing effect is lost when the centromeric chromatin structure is altered.
Figure 6-3. Orc1 depleted of its BAH domain is able to rescue the orp1-4 ts allele. The orp1-4 is a temperature sensitive mutant and does not grow at 36°C. Integration of the orc1 gene or its BAH domain-deleted variant at leu1 locus is able to rescue the temperature sensitivity indicating that the BAH domain is not required for replication and viability.
Chapter 6: Candidate DNA binding proteins and centromere determination

(Allshire et al., 1994; 1995). The centromere contain two distinct chromatin domains, outer repeat heterochromatin and central domain CENP-A chromatin which respond to mutants that affect either heterochromatin or CENP-A/kinetochore chromatin integrity (Partridge 2000; Pidoux et al 2003). To test if the orp1-BAHΔ mutation affects silencing, different marker genes at distinct insertion sites were tested for increased expression alongside known mutations that affect heterochromatin (dcr1Δ or sir2Δ; lack Dicer or Sir2 histone deacetylase) or CENP-A chromatin integrity (hrp1Δ; lack Chd1 chromatin remodelling activity) and silencing at those locations. Cells with arg3+ gene inserted within the central core of cen1 (cen2-cc2-Sph1:arg3+; Pidoux, 2003) where CENP-A chromatin is assembled was used to determine if orp1-BAHΔ silencing at this location (Figure 6-4B). Two locations were tested in heterochromatin: the ade6+ gene inserted in the outer repeat at the Sph1 site on the right side of cen1 (cen1-otr1R-Sph1:ade6+) or the ura4+ inserted in the part of heterochromatin that extends into the imr region (cen1-otr1-Nco1:ura4+; Allshire et al., 1995), which is known to be more sensitive to loss of Sir2 function and might therefore might be affected by the removal of the BAH domain from Orc1 (Figure 6-4C, -D).

cen1-otr1R-Sph1:ade6+: a strain containing this insertion within the outer repeats of cen1 and the orp1-BAHΔ mutation was constructed. Silencing of ade6+ results in red colony colour in wild-type cells, the deletion of dcr1 alleviates repression so that ade6+ is expressed. The deletion of BAH from Orc1 did not appear to affect heterochromatin integrity at this location within the outer repeats since the colonies formed had a level of red colony colour indistinguishable from wild type (Figure 6-4C).
Figure 6-4. Effects of BAHΔ on centromere structure. A) Schematic representation of the marker genes integrated at centromere 1 used in the silencing assay. *orp1+* and *orp1-BAHΔ* had *orp1* gene deleted from the endogenous location and express Orp1 (wt or BAHΔ) from leu1 locus. B) Spotting assay of wild type and mutants with *arg3*′ inserted within the central core 1 (*cc1:arg3*). Cells are plated in rich medium (PMG complete) and in medium lacking arginine (-arg). In wild type silencing of the arg3 marker is maintained and cells are unable to grow on -arg while *hrp1Δ* is known to affect centromeric chromatin and is able to grow on -arg. C) Wild type and mutants containing *ade6*+ inserted at the *otr1R* (*otr:ade6*). Cells are plated in rich medium (YES) and with limiting amount of adenine (1/10th ade). In wild type silencing of the ade6 marker is maintained and cells are red coloured. In *dcr1Δ*, heterochromatin is disrupted and ade6 is expressed giving the cell a white colour. D) Cells possessing *ura4* inserted at the innermost repeats of centromere 1 (*imr1R:ura4*). In wild type cells silencing of the *imr* is maintained and cells are able to grow in medium containing the counter selective drug FOA and grow poorly on medium lacking uracil (-URA). *sir2Δ* shows alleviation of silencing at the *imr* and the *ura* marker is expressed in those mutant allowing the growth on -ura.
Chapter 6: Candidate DNA binding proteins and centromere determination

cen1-imr1R-Nco1:ura4*: cells carrying the ura4* gene in the proximal region of imr and the orp1-BAHΔ mutation were tested for silencing at this site. Silencing of ura4* in wild type cells results poor or no growth on selective plates devoid of uracil and allows growth on counter selective plates containing 5-FOA. In sir2Δ cells heterochromatin is disrupted at this site resulting in increased growth on –URA plates and loss of growth on +FOA plates. The orp1-BAHΔ mutation had no impact on heterochromatin dependent silencing at this location at 25, 32 or 36°C where Sir2 is known to be required (Figure 6-4D).

Establishment of functional centromeres in the orp1-BAHΔ mutant

Plasmid based centromeres composed of a central kinetochore domain adjacent to heterochromatin coated repeat DNA form functional centromeres when transformed into wild-type cells but are unable to do so in cells with defect in establishing heterochromatin or CENP-A chromatin.

If the pool of ORC is involved in promoting CENP-A incorporation through the BAH domain of Orc1 (possibly by recruiting Sir2) then the orp1-BAHΔ mutant would block the establishment of function centromeres on plasmid-based minichromosome. To test this, a minichromosome plasmid containing 5 kb of the heterochromatic repeats and 8.5 Kb of cc2. In wild-type cells the establishment of functional centromeres is indicated by the presence of colonies with white/red sectored colonies, the white sectors contain cells carrying the plasmid and are ade6+ whereas the plasmid is lost form cells in the red sectors. Following transformation into orp1-BAHΔ cells colonies with functional centromeres were clearly formed. This preliminary assay indicates that removal of the BAH domain from Orc1 does not affect the establishment of heterochromatin or CENP-A chromatin on naked centromeric DNA (Figure 6-5).

6.3 DISCUSSION

In this chapter, known centromere proteins containing DNA binding motifs were tested for their ability to bind centromeric DNA carried on a plasmid in a context that CENP-A<sup>Cnp1</sup> deposition has not occurred. All the proteins analysed play a critical role in centromere formation and in particular Scm3 is known act upstream of CENP-A<sup>Cnp1</sup> and mediate its deposition at centromeres. Moreover, in fission yeast Scm3 appears to recognize centromeres independently of CENP-A<sup>Cnp1</sup> since it remains associated with centromeres when CENP-A<sup>Cnp1</sup> nucleosomes are lost.
Figure 6-5. Deletion of BAH domain does not affect centromere establishment on a minichromosome. Cells expressing Orc1 with its BAH domain deleted were transformed with a minichromosome possessing full length cc2 and part of the otr. Cells were growing on selective medium and then plated on rich medium with limiting amount of adenine. The presence of white colonies with red sectors indicates a functional kinetochore is established on the minichromosome.
Candidate DNA binding proteins and centromere determination

(Pidoux et al., 2009; Williams et al., 2009). In wild-type cells the 3xLM plasmid does not attract CENP-A. A protein that directly binds the centromeric LM DNA is involved in the its recognition to establish CENP-A\textsuperscript{Cnp1} assembly would be expect to bind to this DNA in this ground state. However, Scm3, CENP-C\textsuperscript{Cnp3} or CENP-T\textsuperscript{Cnp20} were not found to be enriched on the 3xLM DNA even though this sequence can clearly form functional centromeres when the required contextual cues are provided (Chapter 3 Figure 3-4). Another factor found at the centromeres, Orc1, was analysed for its ability to affect centromere structure. Genome wide ChIP analysis have shown that both Orc1 and Orc4 are associated with the central domain of centromeres but replication appears to be instead initiated in flanking heterochromatic regions and spreading over the CENP-A\textsuperscript{Cnp1} domain (Hayano et al., 2012; Hayashi et al., 2007; Matsumoto et al., 2011). This suggests that ORC may possess a replication-independent role in these regions of centromeres. Indeed, in S. cerevisiae and K. lactis, in addition to its role during replication, ORC is also involved in formation of Sir2-dependent repressive chromatin through its BAH domain (Hickman et al., 2011). In mammalian cells, the BAH domain of Orc1 has recently been shown to define replication origins by binding to histone H4K20me2 chromatin (Kuo et al., 2012). However, neither S. cerevisiae nor S. pombe Orc1 BAH domains were able to bind H3K20me2. Since the orp1-BAH\textsuperscript{Δ} mutant is viable and grows similarly to wild-type cells it can not play a major role in replication (Figure 6-3). In fission yeast, several observations indicate that Sir2 is involved in establishing heterochromatin and CENP-A\textsuperscript{Cnp1} chromatin and functional centromeres on naive plasmids (Allshire lab-unpublished). However, it is not known how Sir2 is recruited at fission yeast centromeres. To test if BAH domain of Orc1 is involved in maintenance of centromeres structure, the integrity of specialised centromeric chromatin in the orp1-BAH\textsuperscript{Δ} mutant was tested using silencing assays as an indirect readout. However, sensitive assays, designed to detect defective heterochromatin or CENP-A chromatin, did not report any role for the BAH domain of Orc1 in the maintenance of these type of chromatin (Figure 6-4). In addition, a preliminary centromere establishment assay also demonstrated that functional centromeres can be established on plasmid-based minichromosome following their transformation into orp1-BAH\textsuperscript{Δ} cells. This indicates that the BAH domain is not involved in the recruitment of factors that mark the centromere for CENP-A\textsuperscript{Cnp1} assembly (Figure 6-5). In K. lactis cells lacking the Orc1 BAH domain are defective in silencing at the mating type locus and telomeres (Hickman and Rusche, 2010). Thus, it remains possible in fission yeast the BAH domain of Orc1 might also contribute to silencing at the mating type locus or telomeres.
There are many other centromere associated proteins which, if such a specific DNA binding protein exists, may act to prime CENP-A<sup>Cnp1</sup> assembly or alternatively, some other novel factor or non-specific activities (e.g., transcription, replication), may initiate CENP-A assembly on these specific sequences. One route to identify specific centromeric binding factors that promote centromere formation would be to employ an unbiased method that does not rely on testing candidate proteins. For example, several laboratories (Akiyoshi et al., 2009; Ohkuni and Kitagawa, 2011) have successfully developed methodology that allows the affinity selection of plasmids from cells and the identification of proteins associated with specific sequences by mass spectrometric analyses. Such an approach could be applied to the plasmid containing centromeric domain DNA in the presence and absence of CENP-A chromatin and allowing the identification of the factors that bind this centromeric DNA before CENP-A and kinetochore assembly.
CHAPTER 7

7 ROLE OF TRANSCRIPTION IN CENTROMERE FUNCTION

7.1 INTRODUCTION

Despite the fact that centromeres are generally located in poorly transcribed regions, increasing evidence suggests that transcription of centromere sequences and centromeric transcripts themselves might contribute centromere function and integrity (Chen et al., 2008; Nakano, 2003; Saffery et al., 2003; Topp et al., 2004). Evidence suggests that the transcripts originating from the centromeric region may play a role in maintaining centromeric structure. Indeed, in maize and marsupials cells, centromeric RNA has been found to co-immunoprecipitate with native CENP-A indicating that RNA is an integral component of centromeric chromatin and is possibly involved in CENP-A recruitment to centromeres (Carone et al., 2009; Topp et al., 2004). Moreover, RNAi mediated knockdown of LINE element transcripts that underlie the neocentromere on the mardel(10) marker chromosome led to reduced levels of CENP-A at the neocentromere and increased missegregation of that chromosome (Chueh et al., 2009; Nakano et al., 2008; Voullaire et al., 1993). However, although centromeric transcripts may play a role in centromere function, maintaining a particular level of centromeric transcription appears important for centromere function. For example, in budding yeast and human centromeres, altering the level of centromeric transcription results in defective centromere function (Chueh et al., 2009; Nakano et al., 2008; Ohkuni and Kitagawa, 2011). Related effects have been observed when the chromatin state on human artificial chromosome by targeting a transcriptional activator or repressor is targeted to its centromere (HAC) (Bergmann et al., 2011; Nakano et al., 2008). Such manipulations result in missegregation and loss of the HAC suggesting that both high transcriptional environments and repression of the centromere transcription are not compatible with kinetochore function (Nakano et al., 2008; Ohkuni and Kitagawa, 2011; Pidoux and Allshire, 2004; Wood et al., 2002).

In fission yeast, the otr repeats surrounding the central domain are transcribed and this process is responsible for the establishment and maintenance of the RNAI-
directed heterochromatin on these repeats (Djupedal et al., 2005; Kato et al., 2005). Moreover, transcription also occurs within the central CENP-$A^{Cnp1}$ chromatin domain of cen1 and transcripts and their transcriptional start sites have been mapped (Choi et al., 2011). However, the functional relevance of these transcripts and whether they play any role in centromere function is not known. In wild-type cells, these central domain transcripts are normally degraded by the exosome so that in cells with a cold-sensitive mutation in the gene encoding Dis3, a major subunit of the exosome, a smear of transcripts is detected by northern analyses (Figure 1-8). In contrast, in mutants impaired in CENP-$A^{Cnp1}$ assembly (cnp1-1 or mis6-302), H3 is deposited at centromeres and discrete sized poly-adenylated transcripts originating from within central domain regions are detected (Figure 1-9).

Transcription through chromatin requires ATP-dependent remodelling machinery, histone chaperones and histone modification enzymes (Li et al., 2007). During transcription nucleosomes are disassembled in front of elongating RNAPII and reassembled behind (Boeger et al., 2003; Rocha and Verreault, 2008; Williams and Tyler, 2007). This transcription-coupled reassembly process allows the deposition of other specialised histone variants, as has been shown for H3.3 in metazoan (Ahmad and Henikoff, 2002). It is possible that transcription of centromeric DNA provides an opportunity for the replacement of canonical H3 nucleosomes with CENP-A nucleosomes.

In this chapter, transcription of central domain of centromere 2 (cc2) DNA sequences is investigated. The transcriptional start sites and upstream promoters from within the minimal cc2 region LM are identified and characterised. Moreover a novel system for specifically inhibiting transcription from these centromere regions is developed; preliminary tests are presented.

7.2 RESULTS

7.2.1 Analysis of transcripts from the central domain of centromere 2

Transcripts from the central domain of all three fission yeast centromeres have been detected by RT-PCR in cells bearing a defective exosome (dis3-54) or with reduced CENP-$A^{Cnp1}$ levels over the central domain (cnp1-1 or mis6-302). Originally, most analysis was carried out on cen1 and cen2 transcripts were detected only by RT-PCR ((Choi et al., 2011). However, the investigation of the minimal sequence requirements for CENP-$A^{Cnp1}$ recruitment (described in chapter 3) was performed using cc2 sequence. Therefore, in order to investigate whether transcription is
important for CENP-A\textsuperscript{Cnp1} assembly on cc2, the nature of transcripts originating from endogenous cc2 was investigated more thoroughly by northern analysis. RNA samples were extracted from cnp1-1, mis6-302 and mis18-262 mutant cells, which have significantly reduced levels of CENP-A\textsuperscript{Cnp1} in the central domain and consequently more H3 chromatin assembled in its place. This RNA was analysed by northern using strand-specific riboprobes homologous to the cc2 sequence. It was previously shown that transcripts from the central domain of centromere 1 give a strong signal (Choi et al. 2011). Although these cc1 transcripts were clearly visible in the same samples, transcripts originating from cc2 were not detected using several probes homologous to, and covering the entire length of, the cc2 sequence (one example is shown on Figure 7-1). The central domains of cc1 and cc3 are highly homologous, therefore it is possible that because cc1 transcripts are detected with probes that also recognise cc3 transcripts they are more abundant and consequently easier to detect than those originating from cc2. It is also possible that although cc2 transcripts were detected by RT-PCR ((Choi et al., 2011), they may more rapidly degraded than those form cc1 and are therefore not detectable by northern analysis.

7.2.2 The detection of transcripts using plasmid borne central domain DNA

When cells bearing the cnp1-1 or mis6-302 mutation are grown at 36°C (non permissive temperature), CENP-A\textsuperscript{Cnp1} is lost from the central domain whereas H3 levels increase (Castillo et al., 2007). In these mutants, the H3 chromatin environment allows the production of stable central domain transcripts so that the RNAs originating from both forward and reverse strand of endogenous cc1 are more easily detected by northern analysis using strand specific probes while they are not detected in wild type cells where CENP-A\textsuperscript{Cnp1} chromatin remains intact (Choi et al., 2011).

It is known that when plasmids, such as pcc2, carrying the central domain from cen2 but lacking heterochromatin, are transformed in wild type cells CENP-A\textsuperscript{Cnp1} is not incorporated and H3 nucleosomes remain on the central core region (Folco et al., 2008a). Thus, in wild-type cells a chromatin state is created on pcc2 that is similar to the endogenous centromeres in cnp1-1 and mis6-302 mutants and thus it might be possible to detect cc2 transcripts more easily from these plasmids. In addition, the use of cc2-containing plasmids increases the copy number of cc2 sequences in cells and consequently the levels of transcripts should be higher and easier to detect.
Figure 7-1. Northern of transcripts originating from endogenous centromeres. A) Schematic representation of endogenous centromere 1 and 2 and position of the probes used in this example (black arrow). B) The figure shows an example of northern analysis using probe on cc1 and cc2 (on K fragment). The total RNA was extracted from kinetochore mutants (mis6-302, cnp1-1, mis18-262) and wild type cells grown at 36°C. The RNA probes specific for cnt1 are described in Choi et al. 2011.
Since transcripts from \textit{cc1} have been already characterised, in order to validate this approach, a plasmid containing 5.5 kb from the central domain of \textit{cen1} (pcc1) was transformed into wild-type cells and extracted RNA was analysed. Choi et al. (2011) showed that both forward and reverse transcripts originating from the endogenous \textit{cc1/3} regions were detected in \textit{mis6-302} mutant cells (Figure 7-2). However, these pcc1 transformants only the forward transcripts were detected. This suggests that factors other than the absence of a functional kinetochore may contribute to transcription from endogenous \textit{cc1} (e.g. environment, transcription factors) (Figure 7-2). The lack of reverse transcripts may also be due to the absence a complete central domain in the 5.5 kb fragment utilised, which may be missing regulatory elements (e.g. promoters or binding sites for transcription factors) that are normally present at the endogenous \textit{cc1} sequence. Alternatively, the reverse transcripts may still be subject to degradation by the exosome in these wild-type cells.

To assess if transcripts originating from the LM sequence can be detected from the full-length \textit{cc2} sequence, northern analysis was performed using RNA extracted from wild type cells transformed with a plasmid containing the 8.5 kb of \textit{cc2} sequence (pcc2) and using riboprobes specific for the LM sequence (Figure 7-3A). However, transcripts were not detected from \textit{cc2} in the condition used for the analysis (Figure 7-3B). One possible explanation is that the level of transcripts may be below the limit of detection by northern. In order to increase the quantity of \textit{cc2} template in the cells, a plasmid containing the 3xLM sequence (p-3xLM) was transformed into wild type cells (Figure 7-3A). Forward transcripts with discrete sizes of \(\sim\)1000 and 3000 nt were detected using riboprobes specific for the LM sequence, no reverse transcripts were detected (Figure 7-3B). More colonies transformed with p-3xLM were analysed (Figure 7-3C). The presence of a band of 3000 nucleotides suggests the presence of a long transcript covering more than one LM repeat. The assay was performed on cells containing the endogenous \textit{cc2} sequence, since no signal detected from cells containing a control plasmid (no \textit{cc2} sequence) the transcripts detected must originate from the p3xLM plasmid (Figure 7-3C). The reason for the difference in the detection of transcripts originating from pcc2 and p-3xLM may be that the LM sequence does not contain promoter elements and the transcripts detected from the p-3xLM sequence originated from start sites created by chance in the linker region between the LM repeats or from within the plasmid backbone itself. Alternatively, since LM is present in three copies on the plasmid the transcripts may just be more abundant and consequently more easily detected than those originating from the single copy carried pcc2.
Figure 7-2. Northern of transcripts originating from a cc1-containing plasmid.
Northern Blot analysis of total RNA extracted from a kinetochore mutant (mis6-302) and wild type cells transformed with pcc1i (pMC101) or a vector without cc sequence (pMC1). The figure shows blot for RNA extracted at restrictive temperature for mis6-302 (36°C) and 32°C for the other samples. The RNA probes used are complementary to cnt1 as indicated in red (forward) and black (reverse) and are described in Choi et al. 2011. *: transcripts detected from mis6-302.
Figure 7-3. Northern of transcripts originating from pcc2 and p-3xLM. A) Schematic representation of plasmids used in the analysis. B) The figure shows blot of total RNA extracted from wild type cells transformed with pcc2 or p-3xLM. C) Blot of total RNA extracted from independent colonies of wild type cells transformed with p3xLM or an empty vector (without cc sequence-pMC1) and PCR-purified DNA of cc2 (cc2-DNA). The two 1 Kb RNA probes are complementary to the L and M fragment of cc2 as indicated in red (forward) and black (reverse). Arrows show the transcripts detected for p3xLM. EtBr: Ethidium Bromide staining of agarose gel.
7.2.3 Analysis of LM transcripts by RT-PCR

To further analyse transcripts from the LM region RT-PCR was performed on poly(A) purified RNA extracted from an exosome mutant (dis3-54) cells containing the p-3xLM plasmid, which does not assemble CENP-A\textsubscript{Cnp1} in these cells. To analyse only transcripts that originate from the plasmid, the strain utilised has most of the endogenous cc2 sequence replaced by cc1 (cc2\textDelta::cc1; as described in Chapter 4). Oligonucleotide probes spaced every 200 bp were used for these analyses and transcripts were detected along the LM region with stronger bands for the fragments 9 and 10 (Figure 7-4A). Interestingly, a secondary band was also detected from fragment 9 that migrates below the primary band. This region of the sequence contains a splicing consensus site (GU-AG) suggesting that the lower abundant band may correspond to a spliced form of the transcript full length. This was confirmed by sequencing of both products from fragment 9; the two transcripts that differ in size by the removal of a 58 bp intron residing between 5' and 3' splice site consensus sequences (Figure 7-4B).

7.2.4 Characterisation of transcriptional start sites and upstream promoter regions from cc2

\textit{In vitro} transcription studies have shown that in \textit{S. pombe} extracts, transcription initiates within a window of 25-40 nucleotides downstream of the TATA sequence (Choi et al., 2002). One of the most studied fission yeast promoters is the promoter from the \textit{nmt1} gene, which has a TATA sequence 25 bp upstream of the transcription start site. Mutations in the TATA element have been shown to affect the level of the expression from the nmt promoter but they do not alter the position of transcription start site (Basi et al., 1993). Apart from \textit{nmt1}, the detailed characterisation of other fission yeast promoters has been minimal, therefore, in order to identify and characterise the promoters within the centromeric LM sequence the transcription start sites (TSS) were first identified by 5'-RACE-PCR. 5'RACE-PCR was performed on poly(A) purified RNA extracted from the exosome mutant (dis3-54) cells lacking endogenous cc2 (cc2\textDelta::cc1)and carrying the p-3xLM plasmid. The 5'-RACE-PCR assay used requires the initial isolation of 5' 7-methylguanosine
Figure 7-4. Detection of transcripts originating from 3xLM. A) RT-PCR and performed on a poly(A) purified RNA extracted from an exosome mutant (dis3-54) depleted of its endogenous cc2 transformed with the 3xLM plasmid. --RT: negative control performed without reverse transcriptase; *: spliced form of transcript 9; act: control of reaction on act1 gene. B) Partial sequence of transcript 9. In red the splicing consensus and grey the intron sequence.
Figure 7-5. **Schematic representation of TSS identified within 3xLM.** 5’RACE-PCR was performed on poly(A) purified RNA extracted from the exosome mutant (dis3-54) lacking endogenous cc2 and transformed with the p-3xLM plasmid. Cells were grown at restrictive temperature for 9 h (18°C). In the figure, a total of 9 TSSs are represented as red arrows. 5 TSSs are originating from forward and 4 from the reverse strands of the LM sequence. p200a and p200b represent the two 200 bp regions upstream the TSSs used for further studies. For more information about TSSs sequence and agarose gels see Appendix 3.
capped RNA and allows the positioning of the TSS at base pair resolution. The
oligonucleotides employed were spaced approximately every 100 bp to cover the
entire LM sequence. These analyses allowed a total of 9 TSSs to be identified, 5
originating from the forward strand and 4 from the reverse strands of LM (Figure 7-5,
Appendix 3). The fact that these transcripts are polyadenylated and possess a 5’-cap
indicates that RNAPII is responsible for their transcription. In addition,
sequencing confirmed the presence of the spliced isoform of the transcript 9
previously identified in the RT-PCR analysis (Figure 7-4A, B).

To identify the promoter regions, 200 bp regions upstream of two of the identified
TSSs were chosen for the analysis (named p200a and p200b- Figure 7-5). These
regions were placed on a plasmid upstream of a LacZ reporter system, which
produces β-galactosidase enzyme if a promoter is present. To validate the system,
repressible promoters (Forsburg, 1993a) that are routinely used to drive different
levels of transcription in S. pombe were also placed upstream of the LacZ reporter.
Following transformation into S. pombe, the level of LacZ transcription was
assessed by lysing colonies grown on nitrocellulose membrane and incubating with
substrate 5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-Gal). In Figure 7-6, the
intensity of the blue colour derived from the expression of LacZ is proportional to the
strength of the promoter as previously described (nmt3>nmt41>nmt81(Forsburg,
1993b). The level of β-galactosidase activity was also assessed in a liquid assay where
protein extracts from equivalent cell numbers was incubated the 2-Nitrophenyl-β-D-
galactopyranoside-ONPG substrate and the product analysed by absorbance at 420
nm in a spectrophotometer. This assay also detected expected differences in the
expression levels from the nmt promoters analysed in their expressed (-thiamine)
and repressed states (+thiamine) (Figure 7-6).

To determine the activity of putative centromeric LM promoters, both the membrane
and the liquid assays were performed on cells carrying p200a-lacZ or p200b-lacZ
constructs (Figure 7-7). Although weak, LacZ/β-galactosidase activity was produced
from both p200a and p200b, indicating that these regions are able to drive
transcription of the LacZ reporter (Figure 7-7A). Importantly, when the p200b
promoter was inverted relative to LacZ (p200b-inv), the activity detected decreased
to background levels; indicating that p200b possesses directionality in its function
(Figure 7-7B, C).
Figure 7-6. LacZ assay on *S. pombe* promoters. *S. pombe* inducible promoters (ntm3, nmt41, nmt81) were cloned upstream of the LacZ gene and transformed into wild type cells. A) Transcription of the LacZ was assessed using lysed colonies on nitrocellulose membrane and incubation with the lacZ substrate 5-bromo-4-chloro-indolyl-β-D-galactopyranoside (X-Gal). In this assay, the intensity of the blue colour is proportional to the strength of the promoter. B) LacZ assay on liquid culture was assessed by incubating protein extract with the substrate (2-Nitrophenyl-β-D-galactopyranoside-ONPG) for the β-Galactosidase enzyme analysable by spectrophotometer (absorbance at 420 nm). The promoters were analysed in their induced (-thiamine) and repressed forms (+thiamine). np: no promoter upstream LacZ.
In chapter 5, the design of a synthetic sequence (SynRLM) was described that, although possessing high sequence similarity with the endogenous LM, was unable to recruit CENP-A<sup>Cnp1</sup> in fission yeast cells. To test the transcriptional capacity of SynRLM compare to the natural LM sequence, the region corresponding to p200b (Synp200b), was also placed upstream of LacZ. No activity was detected for this synthetic sequence indicating that it can not act as a promoter (Figure 7-7C).

To narrow down the regions required to promote transcription, p200b was subdivided into four 50 bp fragments (p50a, b, c & d; figure 7-8A) and their ability to drive LacZ transcription tested (Figure 7-8B). β-galactosidase activity was detected for all four 50 bp fragments analysed, indicating that they can all promote transcription (Figure 7-8B). Interestingly, the p50d alone region displayed much stronger promoter function compared to p200b, where it is adjacent to other centromeric sequences. Since the distance to the LacZ reporter is identical for the 'd' region in the p200b and p50d constructs, this difference can not be due to a difference in the distance between the promoter and the ATG of LacZ. It is possible that sequences in p200b inhibit the p50d promoter when they are in close proximity.

To further dissect the region that promotes transcription, p50b was split into two 25 bp fragments (p25a and p25b; Figure 7-8A) and their promoter activity assayed. p25a does not induce transcription, while p25b clearly possesses promoter activity. The sequence of p200b is very similar to the corresponding region of SynRLM (Synp200b). However, the synthetic Synp200b sequence does not promote transcription of LacZ. Furthermore, although the synthetic version of p50b (Synp50b) shares 75.9 % sequence similarity with p50b, it lacks promoter activity (Figure 7-8C). Therefore, Synp50b can be considered a mutagenized and promoter-less version of p50b. To determine which sequences are required for promoter function, the p50b sequence was altered by changing its nucleotides sequence to those present in the Synp50b version (Figure 7-8D). Since the first 25 nucleotides of p50b do not induce expression of LacZ (Figure 7-8C), the mutations were added to the right half that showed promoter activity (p25b). When these nucleotides were changed to the synthetic version (mut5), the level of β-galactosidase activity was severely reduced (Figure 7-8D).
Figure 7-7. LacZ assay using cc2 sequences as promoter. 200 bp regions upstream of two TSSs (p200a and p200b) were cloned in a plasmid upstream of the LacZ gene and transformed in wild type. A) LacZ assay on nitrocellulose membrane. Formation of blue colonies indicates expression of the LacZ gene. B) LacZ assay on nitrocellulose membrane of p200b, p200b cloned in inverted direction (p200b-inv), and a region of the SynRLM corresponding to p200b (Synp200b). C) LacZ assay in liquid culture of centromeric sequence and comparison with induced S. pombe nmt81 promoter.
Figure 7-8. Dissection of sequence required for the promoter function of p200b. Fragments of the p200b were cloned upstream of the LacZ gene and their activity as promoter assayed in liquid culture. A) Schematic representation of the fragments used in the dissection of the promoter region in p200b. B) LacZ liquid assay of the fragment shown in (A). C) Alignment of p50b and the corresponding sequence on the synthetic SynRLM. D) p50b was mutagenised and the activity of the mutant promoters (mut1-5) assayed in liquid culture. n.p.; LacZ without promoter region.
These analyses demonstrate that just 25 bp from the central domain of cen2 is sufficient to drive transcription and that alteration of this sequence ablates promoter activity. Several other mutants carrying single nucleotide changes were also created (mut1-4), but only mut5 completely disrupted promoter function.

In conclusion, transcripts from the LM region of the central domain from cen2 can be detected, but only when this region is carried in cells on an episomal plasmid. These transcripts are produced by RNAPII as they are 5’-capped, polyadenylated and, in some cases, spliced. Active promoters with these centromeric sequences have been characterised, however, it remains to be determined if these promoters are active at endogenous centromeric cc2 and if this transcription contribute to centromere function.

7.2.5 Investigating the role of transcription in CENP-A\textsuperscript{Cnp1} function

Transcription of centromeric DNA and centromeric transcripts themselves might play an important role in centromere structure and function. In many organisms, RNAPII and its associated transcription factors localise at kinetochore and RNAPII transcripts are produced from centromeres (Chan et al., 2012; Choi et al., 2011; Ohkuni and Kitagawa, 2011). However, the role of transcription at centromere remains unclear. Inhibition of RNAPII using α-amanitin, has been reported to affect the production of centromeric α-satellite transcripts and cause chromosome segregation defects suggesting a role for transcription in centromere integrity (Chan et al., 2012). However, drugs such as α-amanitin will inhibit all RNAPII driven transcription and not just the transcription associated with centromeres. Thus, it is possible that the reported defects in chromosomes segregation that caused by α-amanitin treatment are a consequence of indirect effects on gene expression rather than defects in the transcription of centromeric DNA itself. To allow a more specific test of the role of centromere transcription in centromere formation and/or maintenance, novel tools are required that specifically inhibit transcription at centromeres.

In eukaryotic cells, the multi-protein complex RNAPII catalyses the synthesis of messenger RNA. The largest subunit of the complex, Rpb1, possesses a C-terminal domain (CTD) that contains tandem repeats of the heptamer sequence Tyr\textsuperscript{1}-Ser\textsuperscript{2}-Pro\textsuperscript{3}-Thr\textsuperscript{4}-Ser\textsuperscript{5}-Pro\textsuperscript{6}-Ser\textsuperscript{7}. The CTD is not required for the RNAPII enzymatic activity but instead functions as a binding platform for other proteins, allowing the coupling of transcription with other nuclear processes such us RNA processing and export (MacKellar and Greenleaf, 2011; Proudfoot et al., 2002). The number of repeats
varies between organisms from 29 in fission yeast to 52 in mammalian cells and at least 10 repeats are required for viability in *S. pombe* and 16 repeats for a fully functional Rpb1 (Nonet et al., 1987; Schneider et al., 2010; West and Corden, 1995). Although deletion of some heptad repeats is tolerated, the complete removal of all of the CTD repeats is lethal, possibly due to the decreased transcriptional efficiency (Nonet et al., 1987).

Theoretically, the localised removal of the CTD from RNAP II should prevent synthesis of specific transcripts in that location. TEV (tobacco etch virus) protease has a specific recognition sequence (Glu-Asn-Leu-Tyr-Phe-Gln-Gly-Ala-Ser) and it has been successfully used in elegant *in vivo* experiments to specifically cleave the centromere-associated Rec8 cohesin subunit containing a TEV cleavage site (Yokobayashi and Watanabe, 2005). In order to cleave only the centromeric pool of Rec8, the TEV protease was fused to the C-terminal region of CENP-C<sup>Cnp3</sup> which allows the Cnp3C-TEV fusion protein to be targeted to centromeres (Tada et al., 2011; Yokobayashi and Watanabe, 2005).

We reasoned that the insertion of a TEV cleavage site (tcs) between the globular domain and the CTD of fission yeast Rpb1 (Rpb1-tcs-TEV) would render RNAPII inactive upon release of CTD by cleavage with the TEV protease. Clearly, loss of the CTD from all Rpb1 is expected to be lethal to cells. Therefore, systems for targeting the cleavage of Rpb1-tcs-CTD at only at centromeric regions were devised to prevent synthesis of transcripts at these locations.

### 7.2.5.1 Targeted Inhibition of RNAPII at specific locations

Fission yeast cells were created with Rbp1 containing a TEV cleavage site (tcs-Glu-Asn-Leu-Tyr-Phe-Gln-Gly-Ala-Ser) inserted upstream of the CTD repeats expressed from the endogenous gene (Rpb1-tcs-CTD) (Figure 7-9A). Cells expressing this Rbp1-tcs-CTD are viable, therefore, the addition of the TEV cleavage site does not strongly affect the functionality of Rpb1 or RNAPII. To determine if TEV cleavage of Rpb1-tcs-CTD interferes with transcription CFP-myc tagged TEV protease was fused to the tetracycline repressor DNA binding protein (TetR<sup>off</sup>-TEV) and placed under the control of the *adh21* promoter at the *leu1* locus (Figure 7-9A; the adh21- CFP-myc tagged TEV construct was previously described by Tada et al. 2011). This TetR<sup>off</sup>-TEV fusion protein can be tethered to any region in the genome containing the tetracycline operator sequences (TetO) and can be released from DNA by the addition of anhydrotetracyclin (reviewed in (Corbel and
Rossi, 2002). Analysis of total protein extract revealed that TetR\textsuperscript{off}-TEV is expressed and is able to release the CTD from the pool of total Rpb1 (Figure 7-9B). To test the ability of TetR\textsuperscript{off}-TEV to repress transcription by Rpb1-tcs-CTD, four TetO sequences (4xTetO) were placed upstream of a LacZ reporter expressed from p200b as described above (Figure 7-9C and Figure 7-7). The p200b plasmid was transformed into cells expressing TetR\textsuperscript{off}-TEV alone or in combination with Rpb1-tcs-CTD and expression of LacZ assessed. Although not completely abolished, β-galactosidase activity was reduced by three fold by the targeting of TetR\textsuperscript{off}-TEV in the presence of cleavable Rpb1 compared to the samples expressing wild-type Rpb1. This suggesting that TetR\textsuperscript{off}-TEV may interfere with transcription from p200b by cleaving Rpb1-tcs-CTD (Figure 7-9C).

The ability to repress transcription was also assessed by expressing the TetR\textsuperscript{off}-TEV in cells containing 4xTetO elements upstream of an ade\textsuperscript{+} reporter gene (Bayne et al., 2010). Tethering of theClr4 H3K9 methyltransferase represses 4xTetO-ade\textsuperscript{6} expression resulting in the formation of red colonies (low adenine; Figure 7-9D). In contrast, wild type cells or cells expressing the TetR\textsuperscript{off}-TEV alone formed colonies on –adenine plates and form white colonies on low adenine indicating that the ade\textsuperscript{6} gene is expressed (Figure 7-9D). Similarly, cells expressing both TetR\textsuperscript{off}-TEV cleavable Rpb1-tcs-CTD also form white colonies, however, the colonies are smaller in –adenine plates indicating that growth is slower and some repression might be occurring. To determine if this reduction growth is due to the targeting of TetR\textsuperscript{off}-TEV, these cells were also plated on medium containing anhydrotetracycline (AHT), which releases TetR from TetO sites. In presence of AHT, TetR\textsuperscript{off}-Clr4 does not repress the ade\textsuperscript{6} gene (white colonies only; Figure 7-9D). The growth of cells expressing both TetR\textsuperscript{off}-TEV and Rpb1-tcs-CTD in AHT is comparable to that of the control strains, suggesting that the slow growth of these cells on low and –ade plates might be due to the repression of ade\textsuperscript{6} by Rpb1-tcs-CTD cleavage (Figure 7-9D).

7.2.5.2 Inhibition of transcription at endogenous centromeres

The CENP-\textsuperscript{C}Cnp3–TEV fusion (described by Yokobayashi et al. 2005) was utilised in a preliminary attempt to direct the cleavage of Rpb-tcs-CTD at endogenous centromeres. In this construct, the TEV protease is fused to the C-terminus of CENP-\textsuperscript{C}Cnp3 (Cnp3C) allowing it to be recruited and localisation at centromeres
Figure 7-9. Targeted cleavage of Rpb1 CTD using TetR-TEV. A) Schematic representation of the constructs used. Rpb1-tcs-CTD: TEV cleavage site (tcs) was inserted upstream the CTD of the RNAPII gene; TetR-TEV:TEV protease fused to the tetracycline repressor protein (TetR\textsuperscript{off}). B) Western analysis of protein extracts expressing Rpb1-tcs-CTD alone or in combination with TetR\textsuperscript{off}–TEV. Anti-myc blot shows the expression of TetR\textsuperscript{off}–TEV while the anti-CTD blot indicates cleavage of the CTD. C) LacZ liquid assay on yeast containing TetR\textsuperscript{off}–TEV transformed with a plasmid containing 4xTetO repeats cloned upstream the p200b region and the LacZ gene. D) TetR\textsuperscript{off}–TEV expressed in cells containing 4xTetO repeats upstream of the ade6\textsuperscript{+} gene integrated at the ura4 locus. When spotted on medium containing limiting adenine (low ade), red colonies indicate ade6\textsuperscript{+} repression while white colonies indicate its expression. –ade: minimal medium lacking adenine; TetR-Clr4: fusion protein of the HMT with the TetR protein; AHT: anhydrotetracycline.
Chapter 7: Role of transcription in centromere function

(Tada et al., 2011; Yokobayashi and Watanabe, 2005) (Figure 7-10A, Cnp3-TEV). As a control TEV was also expressed without this centromere targeting motif and so that it is uniformly localised throughout the nucleus (Tada et al., 2011) (Figure 7-10A, TEV). Both the Cnp3C-TEV and TEV were expressed from the weak Padh21 promoter inserted close to the zfs1 locus (Tada et al., 2011; Yokobayashi and Watanabe, 2005).

Analysis of protein extracts from cells expressing Rbp1-tcs-CTD together with TEV or Cnp3C-TEV indicates that the TEV fusion constructs are expressed and that can cleave the CTD away from the Rpb1-tcs-CTD (Figure 7-10B). The TEV enzyme and Cnp3C-TEV do not affect the viability of the cells expressing non-cleavable wild type Rpb1 (Figure 7-10C). However, in the presence of Rpb1-tcs-CTD, the untargeted nuclear TEV enzyme impairs cell growth especially at higher temperatures (36°C), where colonies fail to form (Figure 7-10C). This suggests that indiscriminate cleavage of Rpb1-tcs-CTD occurs and affects global transcription and consequently cell viability.

The Cnp3C-TEV fusion protein allows most of the TEV enzyme to concentrate at centromeres (Yokobayashi and Watanabe, 2005). However, no effect on cell growth or viability was observed when both Cnp3C-TEV and Rpb1-tcs-CTD were expressed (Figure 7-10C, Rpb1-tcs-CTD + Cnp3C-TEV). The difference in phenotypes between cells expressing Rpb1-tcs-CTD combined with non-targeted TEV or Cnp3C-TEV may be explained by the fact that in cells expressing Cnp3C-TEV most of the enzyme is sequestered at centromeres and therefore most RNAPII and thus global transcription is presumably not affected.

As a preliminary assessment of Rpb1-tcs-CTD cleavage by TEV on chromosome stability, the cells plated on medium containing the microtubule-destabilising drug TBZ (Figure 7-7C). Cells with impaired centromere function frequently show sensitivity to TBZ, especially at lower temperature (25°C). Cells expressing Rpb1-tcs-CTD and TEV grow poorly on both rich plates (YES) with or without TBZ. This growth assay suggests that these cells are not more sensitive to TBZ. However, although the strain expressing Cnp3C-TEV and Rpb1-tcs-CTD grow equivalently to
Figure 7-10. Targeted cleavage of Rpb1 CTD to endogenous centromeres. A) Schematic representation of the constructs used. Rpb1-tcs-CTD: TEV cleavage site (tcs) was inserted upstream the CTD of the RNAPII gene; Cnp3C-TEV: TEV protease fused to C-terminus of CENP-C<sup>Cnp3</sup>; TEV: TEV protease without targeting motif. B) Western analysis on protein extracts of cells growing at 32°C and expressing Rpb1-tcs-CTD alone or in combinations with Cnp3C-TEV or TEV. The anti-myc blot shows the expression of Cnp3C-TEV and TEV constructs while anti-CTD blot indicates cleavage of the CTD. C) Spotting assay of cells expressing Rpb1-tcs-CTD, Cnp3C-TEV and TEV and combinations of them. Cells were grown at different temperatures in rich-medium (YES) and at 25°C on medium containing the microtubule-destabilising drug thiabendazole (TBZ, 12.5 µg/ml).
control cells (wt, Rpb1-tcs-CTD only, Cnp3C-TEV only) at 25°C, they display slightly reduced growth in the presence of TBZ. Although indirect, this suggests that Rpb1-tcs-CTD cleavage mediated by Cnp3C-TEV might have a negative effect on chromosome segregation.

7.3 DISCUSSION

Several studies in different organisms have suggested that transcription may be involved in centromere function (Chan et al., 2012; Gassmann et al., 2012; Nakano et al., 2008; Ohkuni and Kitagawa, 2011). To investigate the role of transcription at fission yeast centromeres, transcripts originating from endogenous centromere 2 or when cc2 carried on plasmids were analysed. Although the presence of transcripts within endogenous cc2 could not be detected, transcripts originating from plasmid borne LM sequences repeated in tandem were detected by northern analysis, indicating the presence of promoters within the cc2 (Figure 7-3B). These LM sequence were also detected by RT-PCR and their transcriptional start sites were mapped by 5’-RACE-PCR (Figure 7-4, 7-5). On these plasmids CENP-A<sup>Cnp1</sup> is not assembled on the centromeric LM sequences and H3 nucleosomes remain in place. These plasmids are considered to be in a similar chromatin state as endogenous centromeres in mutants where CENP-A<sup>Cnp1</sup> is lost and centromeric transcripts can be detected (e.g. <i>cnp1-1</i>, <i>mis6-302</i>; Choi et al. 2011). Consistent with this, forward transcripts could be detected from pcc1 in wild-type cells that are similar in size to transcripts that are produced from cc1 at endogenous centromere (Choi et al. 2011). However, the reverse transcripts emanating from endogenous cc1 in <i>cnp1-1</i> and <i>mis6-302</i> are not produced from the pcc1 plasmid (Figure 7-2). Although CENP-A<sup>Cnp1</sup> levels are reduced while H3 levels are elevated at centromeres in these mutants, some CENP-A<sup>Cnp1</sup> nucleosomes remain in place at centromeres while the plasmid borne pcc1 sequence completely lacks CENP-A<sup>Cnp1</sup> nucleosomes (Castillo et al., 2007; Pidoux, 2003). This difference in chromatin composition may explain the observed difference in the transcription pattern between plasmid borne and endogenous cc1 sequences. Alternatively, the presence of adjacent heterochromatin domains at endogenous cen1 may recruit factors that influence the pattern of transcription in the central domain.

The analysis of regions upstream the transcription start sites in the LM region, allowed the identification of some promoter elements required for transcription (Figure 7-8). Future analyses would involve determining if mutations in these
promoters, that prevent transcription, alter the establishment of CENP-A\(^{Cnp1}\) chromatin and the formation functional centromere on minichromosomes.

As an alternative approach to assess the role of transcription at centromeres, a new tool was developed to attempt to selectively repress transcription on centromeric sequences. Strains expressing a TEV cleavable Rpb1 subunit of RNAP II were created. To direct the cleavage of Rpb1, the TEV enzyme was either fused to the TetR DNA binding protein (TetR-TEV) or to the CENP-C\(^{Cnp3}\)-C kinetochore-targeting motif. Tethering of TetR\(^{off}\)-TEV by binding TetO DNA elements decreased LacZ expression, suggesting that removal of the CTD from RNAPII can affect transcription (Figure 7-9B). However, TetR\(^{off}\)-TEV did not alter the transcription when tethered upstream of the ade6 gene. It is possible that the tethering sites are too far away from the ade6 TSS to mediate robust inhibition of ade6 expression. Regardless, the analyses to date suggest that the TetR\(^{off}\)-TEV system can only repress weak promoters such as those identified in centromere sequences. The mild TBZ-sensitivity of cells expressing TEV cleavable Rpb1 and with TEV targeted to centromeres via the CENP-C\(^{Cnp3}\)-C region provides some indication that transcription play a role in centromere integrity but this requires further analyses. Cytological analysis of mitotic cells is required to determine if chromosome segregation is specifically affected in cells expressing TEV cleavable Rpb1 and Cnp3C-TEV. In addition, anti-CENP-A\(^{Cnp1}\) and other ChIP would reveal if cleavage of Rpb1 at centromeres causes a reduction in CENP-A\(^{Cnp1}\) and other kinetochore proteins.
8 DISCUSSION

8.1 Establishment of CENP-A<sup>Cnp1</sup> chromatin in fission yeast

CENP-A containing nucleosomes play a key role in centromere function as they create the interface between centromere DNA and kinetochore proteins (Earnshaw and Rothfield, 1985; Earnshaw et al., 1989; Palmer et al., 1987; 1991; Vafa and Sullivan, 1997). While the function of centromeres is conserved between different organisms, centromeric DNA sequences are rapidly evolving and do not share obvious homology between species. This, together with the existence of dicentric chromosomes and the formation of neocentromeres suggest that the location of centromeres are determined epigenetically by the presence of CENP-A (Allshire and Karpen, 2008; Sullivan et al., 2001). Several proteins are required for the incorporation of CENP-A, however little is known about how those proteins are recruited to centromeric regions and what are the contributors to the establishment of functional kinetochores.

Several studies in fission yeast have demonstrated that two elements are required for the establishment of a functional kinetochore: the heterochromatic regions and the central core DNA (Baum et al., 1994; Folco et al., 2008b; Kagansky et al., 2009). Indeed, heterochromatin flanking the central core is required to promote CENP-A<sup>Cnp1</sup> recruitment over the central domain when placed on a minichromosome (Folco et al., 2008b). Plasmids containing just the cc sequence but not the heterochromatic repeats are not able to acquire functional centromeres and do not recruit CENP-A<sup>Cnp1</sup>. Once CENP-A<sup>Cnp1</sup> chromatin is established, heterochromatin becomes dispensable for its maintenance, suggesting that heterochromatin is exclusively involved in de novo centromere establishment (Folco et al., 2008). However, the mechanism by which heterochromatin is involved in centromere formation is not understood. One hypothesis is that heterochromatin is involved in recruiting factors that destabilise histone H3 and favour CENP-A. Recent observations indicate that the need for flanking heterochromatin can be bypassed in cells expressing a mutant histone H3 with lysine 9 replaced by arginine (H3K9R) (Folco et al. unpublished–Allshire lab). In these mutants, it is possible that H3K9R is less stably associated with cc2 DNA than the wild type H3 and this may favour the deposition of CENP-A<sup>Cnp1</sup>. Alternatively, H3K9R could, by chance, form a binding site for a factor that...
promotes CENP-A<sup>Cnp1</sup> deposition specifically to the central core sequence. The substitution of the lysine 9 of H3 with an arginine could also mimic a deacetylated state of histone H3. Therefore, it is possible that heterochromatin may be required to recruit factors responsible for the deacetylation of H3. Indeed, the specific H3K9 deacetylase, Sir2 that acts at sites of heterochromatin, has been found to bypass the need of heterochromatic repeats for the establishment of CENP-A<sup>Cnp1</sup> chromatin on plasmids, indicating that deacetylation of H3 may be required for centromere establishment (Freeman-Cook et al., 2005; Shankaranarayana et al., 2003 and Allshire lab- unpublished). Typically acetylation neutralises the positive charge of lysine and by affecting the interaction between nucleosomes and DNA it makes the histones easier to displace and therefore it is thought to favour transcription (reviewed in Li et al., 2007). Thus, the deacetylation activity that may be associated with the heterochromatin or mimicked by the H3K9R mutant may contribute to centromere formation by affecting the transcription levels of the region as discussed in more details below.

The second element required for centromere formation is represented by elements within the central core DNA sequence. In fact, the findings presented in chapter 3 and 5, reveal that centromeric sequences must possess some intrinsic characteristics that are strictly dependent on the primary DNA sequence. Thus, alterations of the centromeric DNA by randomisation of the sequence (SynRLM), affected the ability to recruit CENP-A<sup>Cnp1</sup>, indicating that centromere identity is not entirely epigenetically regulated and other elements may be involved (Chapter 5). It is therefore possible that within the same species or between different species centromeres and neocentromeres share some sequence feature that explains the choice of region where a centromere is assembled. One characteristic of these cis-acting elements is that they are distributed along the entire centromeric sequence. In fact, results described in chapter 3 together with previous studies, indicate that there is no specific region of central core 2 that is absolutely necessary for establishing CENP-A<sup>Cnp1</sup> chromatin and combinations of centromeric regions are able to recruit CENP-A (Chapter 3 and Baum et al., 1994). Moreover, since the length of the central core sequence seems to be a critical parameter for CENP-A<sup>Cnp1</sup> assembly, in addition to the sole presence within the centromeric sequence, a certain threshold density of those elements may be required for proper CENP-A<sup>Cnp1</sup> deposition (Figure 3-2).

Thus, since two main contributors to centromere formation can be identified, a simple model can be proposed in which the effect of the heterochromatin must
cooperate with the features encoded by centromeric DNA sequences to direct CENP-A chromatin formation (Figure 8-1A). Once CENP-A chromatin and the kinetochore is established at that site, epigenetic mechanisms could ensure its maintenance and propagation at that site irrespective of specific binding factors or the surrounding chromatin context (Figure 8-1B). Therefore, it is clear that the cc2 sequence possesses some modular elements that strongly promote CENP-A\textsuperscript{Cnp1} deposition. A main focus of this study has been to investigate the nature of these features. Here, I will present three types of models that could explain the role of the DNA sequence and its interaction with heterochromatin.

8.2 MODEL 1: Sequence-specific binding factors promote CENP-A establishment

A simple explanation for the sequence dependence of CENP-A establishment might be that specific DNA binding sites in the central domain are required to recruit specific centromere proteins that mediate CENP-A\textsuperscript{Cnp1} deposition. As already discussed, the simplest point centromeres are the only centromeres known whose function is strictly dependent on the underlying DNA sequence and mutations in this sequence affect centromere function (Lechner and Carbon, 1991). Therefore, it is possible that although other organisms have evolved more complex centromeres they maintained some characteristics of the simplest centromeres such as the presence of specific binding sites for centromeric proteins. Indeed, in humans, the requirement of CENP-B binding motif for \textit{de novo} centromere formation on an artificial chromosome may suggest that there may be some sequence requirements for \textit{de novo} centromere formation even in higher eukaryotes (Ando et al., 2002; Ohzeki et al., 2002).

A large number of proteins are involved in assembling kinetochores and several possess the ability to bind DNA or contain predicted DNA binding motifs. Therefore, DNA binding domains could be involved in the initial recognition of specific features in centromeric DNA and contribute to the establishment of CENP-A chromatin at centromeres (Figure 8-2). In this model, the role of heterochromatin
Figure 8-1. Heterochromatin and central domain DNA are required for establishment of CENP-A chromatin. A) The flanking heterochromatin creates the correct environment condition (represented by a cloud in the figure) that allows the centromeric features (shaded circles) present on the central domain to recruit CENP-A and form functional kinetochore. The central domain consists of redundant elements that are capable of promoting CENP-A chromatin formation. B) Once CENP-A chromatin and the kinetochore are established, epigenetic mechanisms (red arrows) ensure its maintenance and propagation at that site irrespective of the surrounding chromatin context.
Figure 8-2. MODEL 1- Sequence-specific binding factors promote CENP-A establishment. In this model, DNA binding sites for specific centromeric proteins are present along the central domain and are responsible for the recruitment of CENP-A (dashed line). The surrounding heterochromatin may increase the stability of the factors within the centromere or create the correct environment for centromere establishment.
in centromere establishment might be to stabilise these factors within the centromeric sequence. In chapter 6, known centromere proteins containing DNA binding motifs (CENP-C\textsuperscript{Cnp3}, Scm3, and CENP-T\textsuperscript{Cnp20}) were tested for their ability to bind centromeric DNA \textit{in vivo} in absence of CENP-A\textsuperscript{Cnp1} assembled. However, the proteins tested were not found bound to the centromeric DNA present in the plasmid in conditions that do not allow CENP-A deposition (Figure 6-1). There are many other kinetochore-associated proteins, which could potentially bind centromeric DNA and could act to prime CENP-A assembly. A method to identify specific centromeric binding factors that promote centromere formation would be to employ an unbiased method that does not rely on testing candidate proteins. For example, several laboratories have successfully developed methodology that allows the affinity selection of plasmids from cells and the identification of proteins associated with specific sequences by mass spectrometric analyses (Akiyoshi et al., 2009; Ohkuni and Kitagawa, 2011). Such an approach could be applied to the plasmids containing centromeric domain DNA in the presence and absence of CENP-A chromatin and allow the identification of factors that bind centromeric DNA before CENP-A deposition and kinetochore assembly (Figure 8-3).

8.3 MODEL 2: Chromatin remodelling promotes CENP-A establishment

Neocentromeres and centromeres do not share any sequence similarity and this suggests that a more general feature influence centromere establishment. For example, non-specific activities such as transcription or replication, may initiate CENP-A assembly on centromeric sequences through chromatin remodelling events. In Drosophila, the overexpression of CENP-A\textsuperscript{CID} leads to formation of foci enriched in CENP-A\textsuperscript{CID} that are localised preferentially into euchromatin, suggesting that histone H3 must be removed from the regions where ectopic centromeres are formed (Heun et al., 2006). In Chapter 4, plasmids containing the full-length central domain of cc2 exhibited reduced levels of histone H3 associated with centromeric DNA when CENP-A\textsuperscript{Cnp1} chromatin is assembled (Figure 4-7). This suggest that CENP-A\textsuperscript{Cnp1} chromatin is assembled H3 must be removed and therefore, chromatin remodelling events occur.
Figure 8-3. Schematic representation of plasmids used for plasmid-pulldown experiment. Plasmids in which a LacO array was cloned will be used to identify factors that specifically bind the central domain independently on the presence of CENP-A. Plasmid with the central domain of cc2 will be transformed into cells expressing LacI fused to the epitope Flag and affinity purified using antibody anti-Flag. The protein purified will be analysed by Mass spectrometry. As controls, plasmids without the central core sequence or containing the SynRLM will be used. SynRLM was previously shown to not recruit CENP-A (Chapter 5) therefore, it is predicted to not bind centromeric protein.
8.3.1 Transcription as a chromatin remodeler

During transcription, RNAPII must remove the nucleosomes to gain access to DNA and transcribed genes efficiently. The disassembly and re-deposition mechanisms require several chromatin-remodelling factors and some of these factors are also required for proper centromere assembly (reviewed in Petesch and Lis, 2012). Moreover, in different organisms evidence suggests that transcription may have a role in maintaining centromere function (Chan et al., 2012; Gassmann et al., 2012; Nakano et al., 2008; Ohkuni and Kitagawa, 2011). Therefore, transcription may be the trigger for the remodelling of chromatin to facilitate the eviction of canonical histone H3 and deposition of other histone variants such as CENP-A. As previously described, the FACT complex cooperates with the ATP-dependent chromatin remodelling factor CHD1 for the maintenance of the level of CENP-A containing nucleosomes at centromeres (Foltz et al., 2006; Okada et al., 2009; Takahashi et al., 1991). In addition, RSF (remodeling and spacing factor1) a remodeling complex that cooperate with FACT during transcription, interacts with CENP-A and facilitates its incorporation into centromeric chromatin (Perpelescu et al., 2009). In fission yeast, the CHD1 homologue Hrp1 was shown to be required in maintaining the normal levels of CENP-A Cnp1 at centromeres, suggesting that Hrp1 may have a role in transcription-coupled nucleosome disassembly and incorporation of CENP-A Cnp1 (Choi et al., 2011; Walfridsson et al., 2007).

To determine whether transcription plays a role in centromere maintenance, I investigated whether transcription could be detected originating from the cc region of S. pombe centromere (Chapter 7). The presence of multiple TSSs appears to correlate with the modular nature of centromeric sequences in fission yeast (Figure 7-5). As previously described, the central core sequence seems to consist of repetitions of elements that are capable of promoting CENP-A Cnp1 (Chapter 3; Steiner and Clarke, 1994).

One possibility is that transcription at centromeres may induce displacement of canonical histone H3 and a mechanism that inhibits the re-deposition of H3 may act specifically at centromeres. Factors recruited through interactions with kinetochore proteins or with heterochromatin regions might control the eviction of H3 nucleosomes and the inhibition of its deposition at centromeres (Figure 8-4A).
Figure 8-4. MODEL 2: Transcriptional remodelling as a trigger for CENP-A establishment. A) During transcription nucleosomes are disassembled from their original position to allow the RNAP II to elongate through the sequences. The nucleosomes are then reassembled behind the elongating RNAP II and this process requires several chaperones and remodelling factors. However, at centromeres heterochromatin or other centromeric factors may inhibit the reassembly of H3 and favouring CENP-A nucleosomes at centromeres. B) Since transcription is a general mechanism occurring along the chromosomes, CENP-A nucleosomes are incorporated also at non-centromeric regions. However, a surveillance mechanism is responsible for the removal of the ectopically deposited CENP-A to ensure genome stability.
Since transcription is a general mechanism occurring along the chromosomes, it is possible that CENP-A nucleosomes are incorporated also at non-centromeric regions. However, the presence of CENP-A may induce formation of ectopic functional centromeres, therefore a mechanism able to remove the ectopically deposited CENP-A is required to ensure genome stability (Figure 8-4B). This model is supported by evidence in S. cerevisiae where the SWI/SNF complex acts to maintain the organisation of centromeres by removing the mislocalised CENP-A\textsuperscript{Cse4} from non-centromeric sites (Gkikopoulos et al., 2011). Indeed, deletion of one SWI/SNF component, Snf2, leads to an increase of CENP-A\textsuperscript{Cse4} delocalisation at sporadic sites along chromosome arms. In addition, SWI/SNF has been shown to be able to destabilise CENP-A\textsuperscript{Cse4} nucleosomes in vitro, supporting the role of SWI/SNF in the removal of CENP-A\textsuperscript{Cse4} from non-centromeric regions (Gkikopoulos et al., 2011). A similar role has been attributed to the histone chaperone FACT in fission yeast. FACT has no effect on CENP-A\textsuperscript{Cnp1} assembly at endogenous centromeres however, mutations impairing its function lead to widespread incorporation of CENP-A\textsuperscript{Cnp1} at ectopic loci, suggesting that FACT may be involved in preventing promiscuous residence of CENP-A\textsuperscript{Cnp1} (Choi et al., 2012). How then is CENP-A stabilised at centromeres? One possibility is that the presence of the kinetochore at centromeres prevents these chaperones to recognise CENP-A nucleosomes. Alternatively, kinetochore components or factors recruited by the adjacent heterochromatin may somehow inhibit the function of these chaperones specifically at centromeric regions.

8.3.2 Other chromatin remodelling processes involved in centromere function

Several biological processes such as replication and DNA damage involve remodelling events. For example, during DNA replication, several chaperones are involved in the disruption of the chromatin ahead of the replication fork and in the re-establishment following the passage of the replication machinery. In this process, the old histones are recycled together with newly synthesised histones to form new nucleosomes that are incorporated into newly replicated DNA (reviewed in Corpet and Almouzni, 2009). Therefore, replication may function as chromatin remodeller to remove H3 and placing specific histone variants. Indeed, in Drosophila during replication, the H3 histone variant H3.3 is deposited together with H3.1. However, H3.3 is removed in G1 and it is replaced with CENP-A nucleosomes suggesting that H3.3 acts as a "placeholder" for CENP-A in S-phase (Dunleavy et al., 2011). In addition to its remodelling ability, the replication machinery may play different roles
at centromeres. In *S. pombe*, a DNA polymerase α (Pol α) accessory protein, Mc11, and its binding partner Swi7, the catalytic subunit of Pol α, were shown to be involved in maintaining proper chromatin structure at fission yeast centromeres. The *mcl1* and *swi7* mutants show impaired CENP-A association and increased acetylation of histone H4 at the centromeric region (Natsume et al., 2008). Furthermore, in fission yeast centromeres are replicated in early S-phase and the *cc* sequences are primarily replicated by replication forks coming from the *otr* repeats (Kim et al., 2003). Nevertheless, the pre-RC is enriched at the central domain, suggesting that in these regions ORC proteins may possess other functions other than replication (Chapter 5).

DNA damage repair is another process that requires chromatin remodelling involving several histone chaperones (reviewed in Avvakumov et al., 2011). To access the site of damage, existing nucleosomes must be removed to allow the repair of the DNA and this is then followed by the reassembly of the histones previously present together with their specific modifications. Notably, HJURP (Holliday Junction Recognition Protein) the CENP-A histone chaperone specific for CENP-A loading was initially discovered as component of the DNA repair machinery (Kato et al., 2007). Furthermore, DNA repair mechanism may be involved in the recruitment of CENP-A in human and mouse cells (Zeitlin et al., 2009). Indeed, CENP-A nucleosomes are recruited together with the repair machinery to the region where DNA double-strand breaks are induced, suggesting a mechanism that may be required for neocentromere formation.

### 8.4 MODEL 3: Imposition of a precise level of transcription promotes CENP-A establishment

In an alternative model, transcription may function as a sensor to regulate the chromatin state of a specific region through the regulation of the levels of transcription. For example, regulation of transcription levels is required to maintain heterochromatin in mouse cells (Bulut-Karslioglu et al., 2012). In this study, binding sites for several transcription factors were identified within major satellite repeats and the knockdown of some of these transcription factors leads to an increase of the transcriptional levels and consequent loss of heterochromatic marks (Bulut-Karslioglu et al., 2012). Therefore, it was proposed that euchromatic regions might be evolved in a way that transcription arises from the cooperation of transcription factor binding sites found mainly at regulatory elements, such promoters and enhancers, which lead to the generation of defined mRNA species. In contrast,
heterochromatic regions lack this synergistic activity and instead several binding sites for transcription factors are spread along the sequence and low abundant cryptic transcripts are generated (Bulut-Karslioglu et al., 2012). This suggests that the regulation of the transcription allows the maintenance of a certain chromatin state. It is possible that an analogous mechanism could operate at centromeric DNA regions; the centromeric regions may contain binding sites for factors that function as transcriptional regulatory elements and the uncoordinated expression of cryptic transcripts may represent a signal for centromere determination (Figure 8-5A). Alternatively, these regulatory factors may be recruited by neighbours regions close to where centromeres are placed (e.g. surrounding heterochromatic domain) (Figure 8-5B).

Having the proper level of transcription seems to be important for centromere stability. For example, in budding yeast, specific transcription factors, Cbf1 and Ste12, regulate centromeric transcription together with Dig1, a repressor of Ste12-dependent transcription (Ohkuni and Kitagawa, 2011). Chromosome instability of cbf1Δ cells can be suppressed by transcription driven from an artificial promoter suggesting that although not essential, transcription contributes to centromere stability. Nevertheless, deletion of Dig1 increases centromeric transcription resulting in chromosome segregation defects, suggesting that altering the level of centromeric transcription affects centromere integrity (Blower and Sullivan, 2002; Chueh et al., 2009; Ohkuni and Kitagawa, 2011). Evidence suggests that even in mammalian cells the regulation of transcriptional activity within the centromere is required to maintain proper centromeric function. Indeed, targeting a transcriptional activator or a transcription silencer to the centromere of a HAC cause missegregation and loss of the HAC (Nakano et al., 2008). Furthermore, inhibition of RNAPII transcription by treatment with the transcription inhibitor α-amanitin leads to increased number of lagging chromosomes and reduced CENP-C association at centromeres (Chan et al., 2012). Recent analyses in fission yeast have shown that deletion of one of the subunits of the mediator complex, Med20, leads to an increase of RNAPII occupancy at centromeres and correspondent decrease of CENP-A\(^{Cnp1}\) levels (Carlsten et al., 2012). In med20Δ cells, the levels of centromeric transcripts are increased and this may cause eviction of CENP-A\(^{Cnp1}\). However, it is not clear if the reduction of CENP-A\(^{Cnp1}\) nucleosomes observed at endogenous centromeres is due to the lower amount of CENP-A\(^{Cnp1}\) protein observed in med20Δ cells.
Figure 8-5- MODEL 3: Regulation of transcription is required for CENP-A deposition. Highly transcribed regions and inhibition of transcription are both not compatible with centromere function. Therefore, the regulation and the “right” amount of transcription are required for proper centromere formation. A) Binding site for specific transcriptional regulator factors are present along the central domain region to maintain a proper transcriptional level compatible with centromere function. The right balance of inhibition and induction of transcription is important to maintain centromere function. B) Heterochromatin may control cryptic and non-specific transcription within the central core and cooperate with some other centromeric proteins that bind the central domain to recruit CENP-A.
Another possibility is that the role of the surrounding heterochromatin is regulating the transcription from the central domain and priming in this way the region for CENP-A assembly. Then, some other features (e.g. binding sites for centromeric proteins) present on the central domain may be responsible for the recruitment of CENP-A (Figure 8-5B). Thus, centromeres may possess a system that regulates transcription to a level that is compatible with centromere function. This mechanism may required the presence within the centromeric sequence of binding sites for factors that are responsible for the regulation of transcription levels within centromeres.

8.5 Further investigation of the role of transcription in CENP-A establishment

To test whether transcriptional activity is required for centromere function, terminator sequences may be placed few bases after the TSSs identified within the LM sequence. Or alternatively, to selectively inhibit transcription the tool previously described in chapter 7 could be used. This system utilises a TEV cleavage site inserted upstream of the CTD repeats of RNAPII and allows to direct the CTD cleavage through the tethering of the TEV protease enzyme. Thus, the TEV enzyme is either fused to a kinetochore-targeting motif possessed by CENP-C<sup>Cnp3</sup> (Cnp3C-TEV) or to a TetR protein (TetR-TEV). Tethering the Cnp3C-TEV to endogenous centromeres when combined with Rpb1-tcs-CTD induced TBZ sensitivity suggesting that the endogenous centromere structure was affected (Figure 7-10). These studies should be extended to give insight into the role of transcription in maintaining already pre-assembled centromeres.

To assess the function of transcription in centromere establishment, the fusion protein TetR-TEV can be tethered to minichromosome possessing TetO repeats (Figure 8-6A). Since several TSSs were identified across the LM region, the TetO repeats will be positioned in between different LM repeats (every 2 kb) to increase likelihood that the tethered could influence all centromeric sequence on the plasmid. The establishment of a functional centromere will be tested on these plasmids in the presence of heterochromatin flanking the LM repeats. If functional centromeres are not detected or the efficiency of centromere establishment is lower compared to the controls (e.g. plasmid not containing TetO repeats), this may suggest that inhibition of local centromeric transcription in the plasmid is affecting CENP-A<sup>Cnp1</sup> deposition.
Figure 8-6. Inhibition of transcription on plasmid-based minichromosomes. A) Cells expressing the Rpb1-tcs-TEV and TetR are transformed with a plasmid containing the 4xTetO repeats positioned in between the LM repeats that are flanked by heterochromatin. The efficiency of the establishment of functional centromeres on the plasmids will give information about the role of transcription at centromeres. B) The same experiment performed using plasmids without heterochromatin. If CENP-A is detected, the inhibition of transcription at the LM region is sufficient to bypass the need of heterochromatin. Therefore the role of heterochromatin may be to control and/or repress transcription.
In the model schematised in Figure 8-5B, heterochromatin represses transcription originating from the central domain and cooperates with some other features present on the centromeric sequence to recruit CENP-A. If the role of heterochromatin is to repress centromeric transcription, tethering the TetR-TEV to plasmids that do not possess heterochromatin and therefore inhibiting transcription at the central domain, may bypass the need for heterochromatin in CENP-A<sup>Cnp1</sup> establishment on a plasmid (Figure 8-6B).

### 8.6 Concluding remarks

A common feature of centromeres is the presence of a specific chromatin structure in which CENP-A nucleosomes mark the site of assembly of the kinetochore. Although centromeres are thought to be epigenetically regulated, the analysis presented here suggests that genetic mechanisms may contribute to centromere function of regional centromeres. In this study, fission yeast was used as model organism to study centromeric DNA elements. Since its centromeric DNA sequence is well defined and unique in its genome, <i>S. pombe</i> has been proven to be an efficient model system for study of centromeres. Although the centromeric sequence of fission yeast differs from those of other species, it is possible that they all share similar features and encode for similar processes. Thus, centromeric DNA may have evolved and be selected for specific characteristics that are compatible with centromere function and a good substrate for kinetochore assembly. However, it is not only the DNA sequence that is required for centromeres formation. As in fission yeast, it is possible that in other species elements encoded by centromeric sequences may cooperate with other biological process, making the dissection of the requirements for centromere function more difficult. It remains to be determined if the characteristics associated with these sequences, and their mode of action, are conserved at other centromeres.
Acknowledgements

I would like to thank Robin for giving me the opportunity to work and have fun in his lab. Thanks to Alison for being so patient and for teaching me everything I know about pombe. Thanks to members of the Allshire lab, past and present, for all support and for great discussion about science, the universe and everything. Thanks to my tea-break team for being always upstairs at 3:30pm.
So long, and thanks for all the fish.
**APPENDIX 1**

**Appendix 1** - Quantification of the CENP-A levels at the endogenous centromeres of H3 wild type (wt*) or mutated H3K9R cells transformed with plasmids containing different regions of cc2 (A-E). Quantification for CENP-A enrichment in wild type (grey)(n=3) and H3K9R (green)(n=3) was calculated relative to the actin gene (act1*). Quantification of the CENP-A levels on the plasmids can be found on figure 4-1.
Appendix 2- Quantification of the H3 levels at the endogenous centromeres of H3 wild type (wt*) or mutated H3K9R cells transformed with plasmids containing different regions of cc2 (A-E). Quantification for CENP-A enrichment in wild type (grey)(n=3) and H3K9R (green)(n=3) was calculated relative to the actin gene (act1*). Quantification of the H3 levels on the plasmids can be found on figure 4-2.
APPENDIX 3

APPENDIX 3. Agarose gels of 5’-RACE PCR and sequence of TSS identified.
References


Bayne, E.H., White, S.A., Kagansky, A., Bijos, D.A., Sanchez-Pulido, L., Hoe, K.-L.,


and RNAi-directed chromatin silencing. Genes & Development 19, 2301–2306.


Schueler, M.G., Dunn, J.M., Bird, C.P., Ross, M.T., Viggiano, L., NISC Comparative


Sugimoto, K., Yata, H., Muro, Y., and Himeno, M. (1994). Human centromere protein C (CENP-C) is a DNA-binding protein which possesses a novel DNA-binding


Wirbelauer, C., Bell, O., and Schübeler, D. (2005). Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone modifications show a promoter-proximal bias. Genes & Development 19,
1761–1766.


