The Pennsylvania State University

The Graduate School

Eberly College of Science

REGULATION OF DONOR PREFERENCE DURING YEAST MATING-TYPE SWITCHING

A Thesis in

Biochemistry, Microbiology and Molecular Biology

by

Sevinc Ercan

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2005

The thesis of Sevinc Ercan was reviewed and approved* by the following:

Joseph C. Reese
Associate Professor of Biochemistry and Molecular Biology
Thesis Advisor
Chair of Committee

Avery August
Assistant Professor of Immunology

David S. Gilmour
Associate Professor of Molecular and Cell Biology

Song Tan
Associate Professor of Biochemistry and Molecular Biology

Hong Ma
Professor of Biology

Jerry L. Workman
Investigator
Stowers Institute For Medical Research
Kansas City, MO
Special Member

Robert A. Schlegel
Professor of Biochemistry and Molecular Biology
Head of the Department of Biochemistry and Molecular Biology

*Signatures are on file in the Graduate School

ABSTRACT

Mating-type switching in *Saccharomyces cerevisiae* is directional. *MAT***a** cells choose *HMLα* for recombination and *MAT*α cells choose *HMR***a**. This is called "donor preference" and controlled by the recombination enhancer (RE). Donor preference during mating-type switching has been attributed to differences in chromatin structure for the left arm of chromosome III. I mapped the structure of ~45 kbp of the left arm of chromosome III in **a** and α cells in logarithmically growing cultures and in **a** cells during switching. Other than the RE, chromatin structure was identical in the two cell types. Changes in chromatin structure during switching were confined to RE and *HML*. My analysis indicates that primary chromatin structure does not cause the documented differences in recombinational frequency of the left arm of chromosome III in **a** and α cells.

In the process of addressing the role of chromatin in mating type switching, I carried out the first study of chromatin structure spanning an extensive region in an organism whose genomic sequence is known. This region includes thirty ORFs, six potential replication origins, two known silencers and a recombination enhancer, thus may serve as a representative of overall genome organization. Significant features of organized chromatin exist for the entire region. DNase I hypersensitive sites reside at the promoter region of nearly every gene, suggesting that a basic chromatin structural feature exists at every promoter. ~25% of the ORFs possess extended regions of positioned nucleosomes. My study on the chromatin structure of an extended region contributes to understanding how yeast organizes its primary chromatin structure.

A number of **a-**specific non-coding RNAs are transcribed from the RE locus, downstream of the two α2/Mcm1 operators. This transcription is cell cycle regulated. Additionally, proper mating-type switching and donor preference require a cell-cycle regulated factor. Mcm1 and Fkh1 regulate RE activity in **a** cells. Mcm1 binding is required for both RE transcription and Fkh1 binding. This requirement can be bypassed by inserting another promoter into the RE, which increases donor preference and opens the chromatin structure of RE. The data presented in this thesis suggest that the role of Mcm1 in **a-**cell donor preference is to activate RE by providing a promoter-related function such as recruiting the chromatin remodeling proteins and/or other transcription factors. Transcription activation and open chromatin enhances Fkh1 binding and the level of this binding determines the level of donor preference. My study on RE activation helps understanding the mechanism of RE function in increasing the rate of recombination between *HML* and *MAT*.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank my advisor Robert T. Simpson for his support and guidance. He was an excellent scientist and I learned valuable lessons from his approach to experimental science. I would like to thank Jerry L. Workman for opening his lab and helping me with my research; Joseph C. Reese for his help with my research and graduate studies. I also would like to thank my committee members and the past and present members of the Simpson, Workman, Reese and Tan Labs. John, Bob, Alexandra and others provided help and support.

Most of all, I would like to thank my family; Zehra, Ahmet Ali and Sevil Ercan, for always giving me the freedom to choose and supporting my decisions. Finally, I would like to thank my friends in State College. Mike, Sebastian, Nihal, Jale, Baris, Ayse, Mete and the others whose names I did not list here.

Chapter 1

Introduction

The genetic information of a eukaryotic cell is stored in the nucleus. The major constituent of the nucleus is a nucleic acid-protein complex called, “chromatin”. The events in the nucleus such as transcription, replication, recombination and repair happen in the context of chromatin. Therefore, our investigation of the components and the organization of chromatin help us understand how DNA functions during these processes.

One of the major changes that occur in a yeast (*Saccharomyces cerevisiae)* cell is mating-type switching. Switching is a gene conversion event that occurs between the mating-type locus and one of the two mating-type cassettes. Gene conversion is a type of homologous recombination in which the chromatin structure of the recombining sequences are different from each other, and these differences are important for recombination (Haber 1998). Furthermore, the chromatin structure of a *cis*-acting element that regulates donor preference during mating-type switching is also critical for proper gene conversion (Weiss and Simpson 1997). Thus, mating-type switching in yeast cells serves as a good model system to study how chromatin structure affects DNA recombination.

In recent years, both structural and regulatory roles were attributed to chromatin. Many different methods were established to determine these roles and understand the dynamic structure of chromatin. A number of methods that directly examine chromatin structure are based on mapping nuclease accessible sites on DNA (Simpson 1998).

Differences in the pattern of nuclease digestion reveal the sites where proteins interact with DNA and/or DNA assumes a different conformation. Hence, determining nuclease accessible sites in the chromatin is a useful method to monitor changes in the chromatin during mating-type switching.

In the first part of this chapter, I will describe the basic properties of chromatin structure and how DNase I nuclease is employed to analyze chromatin. In the second part, I will introduce the regulation of mating type and mating-type switching in *Saccharomyces cerevisiae*. Then, describe how a *cis*-acting element, the Recombination Enhancer (RE) regulates donor preference during mating-type switching.

1.1 Chromatin Structure and DNase I Mapping

Early observations on chromatin structure identified repetitive beads of nucleoprotein along the DNA (reviewed in (van Holde 1988)). These repetitive beads were later named "nucleosomes". Nucleosomes are comprised of a core particle and a linker DNA that attaches the core particles. The nucleosome core particle consists of a histone octamer and the DNA wrapped around it. Two molecules of each histone, H2A, H2B, H3 and H4, form the positively charged histone octamer. 147 bp of negatively charged DNA wraps around this octamer 1.65 times as a left handed superhelix (Luger et al. 1997). Each nucleosome particle is linked to another with the linker DNA. Another type of histone, H1, binds to the linker DNA between the nucleosomes (Allan et al. 1980).The structural role of the nucleosomes comes from its core particle. The regulatory role of the nucleosomes is mostly attributed to the unstructured amino-terminal tails of

the histones coming out of the core particle (Luger et al. 1997). These histone tails are subject to posttranslational modifications which regulate several functions of chromatin such as regulating transcription.

The primary structure of chromatin is defined as an array of nucleosomes that constitutes the repetitive beads of nucleoprotein. The higher-order conformations of the nucleosomes are less clear. However, it has been observed that a chain of nucleosomes could fold into 30 nm fibers (van Holde 1988). Recent studies showed that this fiber may be composed of two stacks of adjacent nucleosomes forming a helical structure (Dorigo et al. 2004). How the higher order structure of chromatin affects different DNA processes is not known. Certainly, some means of regulation of the higher order structure would be required for assuring the access of proteins to the DNA when necessary.

The primary structure of chromatin affects all the cellular processes that require access to DNA. Chromatin structure regulates transcription at different levels. First, nucleosomes physically block transcription factors' access to DNA. This kind of regulation was first identified at the promoters of the genes (Knezetic and Luse 1986; Lorch et al. 1987; Matsui 1987; Workman and Roeder 1987). Later, many studies showed that transcription activation of a gene involves alterations and/or removal of the histones (Wu, Bingham et al. 1979; Wu, Wong et al. 1979; Levy and Noll 1981). Second, posttranslational modifications at the histone tails regulate transcription in a complex manner. One model to explain the regulatory role of histone modifications in transcription is called "histone code" hypothesis (Jenuwein and Allis 2001). This hypothesis states that posttranslational modifications of the histone tails will result in distinct "read out" of the genetic information. Other levels of transcription regulation are

the presence of histone variants, higher order chromatin structure and nuclear localization.

Early studies on chromatin structure and transcription often employed nuclease mapping techniques (Gross and Garrard 1988). These studies found that nuclease sensitive sites appeared at the 5' end of genes upon transcription activation (Weintraub and Groudine 1976). Nuclease hypersensitive regions reflect the locations of important *cis*-acting regulatory sequences (Sippel et al. 1996; Simpson 1998). A localized hypersensitive site usually appears at a region where nucleosomes are either absent or have an altered confirmation.

One of the most informative nucleases that was used in previous studies to detect hypersensitive sites within a region of chromatin is DNase I (Weintraub and Groudine 1976; Wu and Gilbert 1981; Simpson and Stafford 1983; Staynov and Crane-Robinson 1988; Lutter 1989; Simpson 1998). DNase I is the first enzyme that was used to define nuclease sensitivity of transcribed domains. DNase I hypersensitive sites mark active regions of the genome such as promoters, enhancers and replication origins (Shimizu et al. 1991). DNase I cuts DNA at both the linker and the nucleosome core particle. While it cuts freely in the linker regions, it makes single stranded nicks at 10 nucleotide intervals when DNA is wrapped around the nucleosome core particle. The periodicity of DNase I cutting is useful in recognizing unusual geometries of the DNA around a nucleosome. DNase I is widely used for analyzing nucleosome position, conformation, and protein footprints in higher resolution chromatin mappings. In lower resolution methods however, DNase I marks the regions that are less protected by the primary chromatin.

DNase I digestion of chromatin isolated from nuclei is performed *in vitro*. Although isolation of nuclei is not ideal, previous studies showed that DNase I mapping of the chromatin structure was consistent with *in vivo* methyl transferase mapping (Simpson 1999). Therefore, isolation of the nuclei for DNase I mapping of the hypersensitive sites proved to be a good method to determine locations of important *cis* acting elements. Recently, Wang, X. and Simpson, RT. developed a method that employs DNase I to make cuts in *vivo* (Wang and Simpson 2001). However, the extent of cutting in this system is only suitable for high resolution mapping, where one can resolve base pairs on sequencing gel. By contrast, indirect end labeling can map 5-10 kb regions at medium level resolution. Therefore, it is a valuable method to map longer segments of chromatin. In this thesis, I will present data generated by mapping the chromatin structure of a ~ 45 kbp region at medium level of resolution with DNase I.

1.2 Yeast Mating Type Regulation

Saccharomyces cerevisiae has two haploid mating types, **a** and α. Cell type is determined by regulatory proteins that are encoded from the mating-type (*MAT*) locus. *MAT***a** and *MAT*α differ by approximately 700 bp of sequences, named Y**a** and Yα respectively (Figure 1-1). The Y region in the *MAT* locus contains the promoter and a portion of the coding sequence from the *MAT* genes. Encoded within *MAT* are regulatory proteins that control the transcription of several genes that are important for maintaining different properties of the two types (reviewed in (Herskowitz 1989; Haber 1998)).

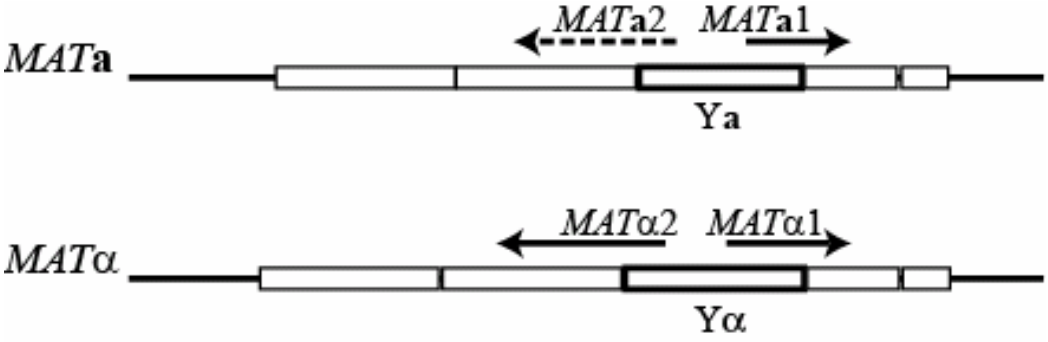

Figure 1-1: Structure of *MAT*a and *MAT*α alleles

*MAT***a** and *MAT*α are distinguished by their Y**a** (650 bp) and Yα (750 bp) regions. *MAT***a** encodes for two proteins, Mat**a**1 and Mat**a**2. *MAT*α encodes Matα1 and Matα2. (Adapted from (Haber 1998))

a and α cells express different regulatory proteins from the *MAT* locus. These proteins act in coordination with a constitutively expressed protein, Mcm1, to regulate the expression of **a**, α and haploid specific genes (Figure 1-1). In α cells, α1 and α2 are expressed. α1 associates with Mcm1 and activates a set of α-specific genes (Herskowitz 1989; Bruhn and Sprague 1994). On the other hand, α2, when associated with Mcm1, works with Tup1 and Ssn6 proteins to repress the **a**-specific genes. In **a** cells, **a**1 and **a**2 are expressed. **a**2 has no identified biological function. **a**1 functions in diploid cells. In **a** cells, Mcm1 activates transcription of the **a**-specific genes (Elble and Tye 1992; Keleher et al. 1992; Herschbach et al. 1994). In diploid cells, α2 and **a**1 are expressed and α1 is repressed. **a**1 associates with α2 to turn off haploid specific genes. Figure 1-2 summarizes the regulation of mating type specific genes.

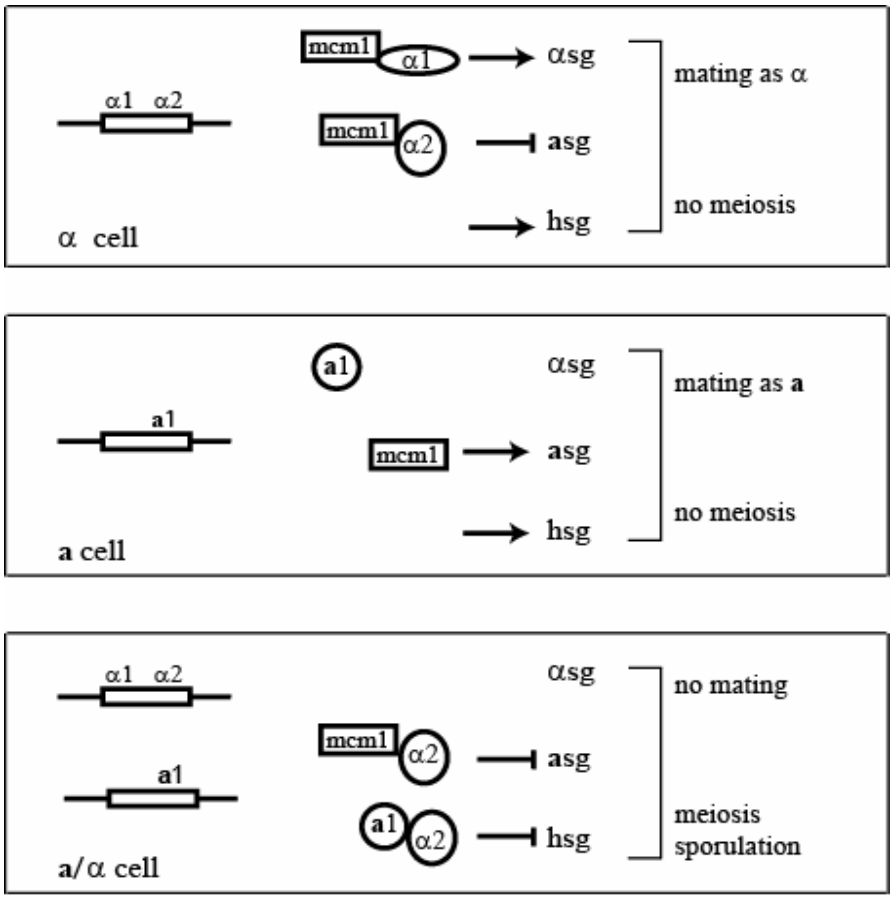

Figure 1-2: Regulation of yeast mating type by the regulatory proteins encoded from the *MAT* locus

Two different types of haploid cells (**a** and α) and the diploid type (**a**/α) express different sets of proteins from the *MAT* locus. These proteins are Matα1 (α1), Matα2 (α2), Mat**a**1 (**a**1) and the constitutively expressed protein Mcm1 (Mcm1). Depending on the cell type, a subset of these proteins are present to regulate the expression of α-specific genes (αsg), **a**-specific genes (**a**sg) and haploid specific genes (hsg). (adapted from (Herskowitz 1989))

Yeast cells of one haploid mating type can switch to the other. Switching ensures that both types of yeast are present in one population. This is required for establishing the diploid life cycle that involves mating of the two cell types. Mating-type switching is controlled by a number of different mechanisms. First, only mother cells can switch their mating type (Strathern and Herskowitz 1979). Second, switching starts at the G1 phase of

the cell cycle (Breeden and Nasmyth 1987). Third, the two mating-type cassettes are silenced and remain heterochromatic (reviewed in (Herskowitz 1989; Laurenson and Rine 1992; Haber 1998)). Fourth, more than 85% of the switching attempts result in change from one type to the other (Klar et al. 1982). In the following paragraphs I will discuss how these control mechanisms are established.

The first control mechanism on mating-type switching requires only the mother cell to start switching. After mitotic division, a haploid cell produces a bigger and a smaller cell, called "mother" and "daughter", respectively (Figure 1-3). HO endonuclease is the enzyme that initiates mating-type switching by introducing a double strand break at the *MAT* locus. After cell division, HO endonuclease expression is restricted to the mother cells, hence, only mother cells initiate switching (Strathern et al. 1982).

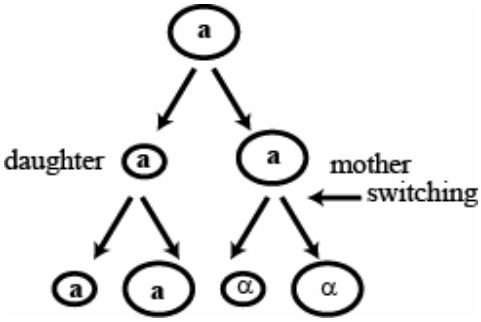

Figure 1-3: Mating type switching of a yeast cell
An **a** cell divides to produce a bigger (mother), and a smaller (daughter), cell. The mother cell switches mating type and gives rise to two α cells, while the daughter cell gives rise to two **a** cells. (Adapted from (Herskowitz 1989))

The second mechanism involves the regulation of *HO* expression during cell cycle. *HO* is expressed during G1 phase of the cell cycle and switching is completed before the end of the S phase (Breeden and Nasmyth 1987). This ensures that *MAT* locus is replicated after mating-type switching is completed. Therefore, when the mother cell

divides, it produces two cells of the other mating type. Since the daughter cell cannot switch mating type, after cell division, a daughter cell produces two cells of the same type. As a result, the population becomes a mixture of the two cell types (Figure 1-3).

The third control mechanism involves the silencing of the mating-type cassettes (reviewed in (Herskowitz 1989; Laurenson and Rine 1992; Haber 1998)). Mating type cassettes are the two alleles of the *MAT* locus located at the opposite ends of chromosome III (Figure 1-4). In most of the yeast strains, *HML* and *HMR* contain Yα and Y**a**, respectively. Therefore, these two cassettes are often referred to as *HML*α and *HMR***a**. Both cassettes have all the sequence information to express the regulatory proteins that are normally expressed from the *MAT* locus. However, transcription is strongly repressed by a heterochromatin structure (Ravindra et al. 1999). The transcriptional silencing is established by the help of two silencer sequences on either side of each cassette (*HML*-E, *HML*-I, *HMR*-E, *HMR*-I). These silencers interact with a number of proteins. Among these proteins are the histones, Silent Information Regulator (Sir) proteins, the DNA replication Origin Recognition Complex (ORC) proteins, several histone deactylases and chromatin assembly factors. These proteins together create ~3 kb of heterochromatin between the two silencer sequences. Transcriptional silencing of the mating type cassettes is crucial for a cell to maintain its mating type. Moreover, the heterochromatin domain over the mating type cassettes is important for proper gene conversion during mating-type switching.

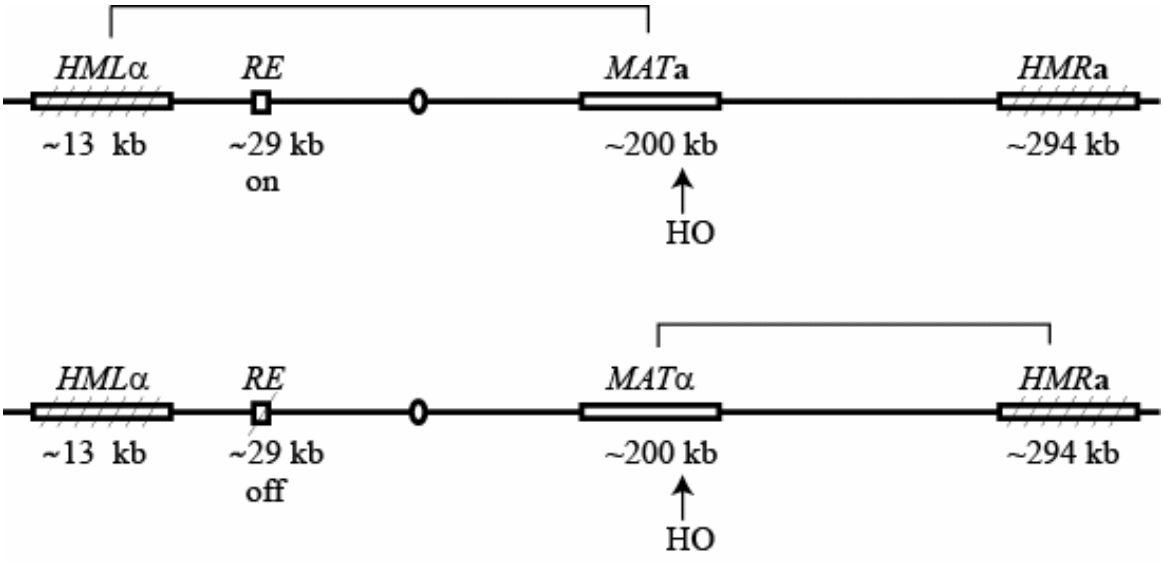

Figure 1-4: Mating type loci on chromosome III

The locations of all three mating type loci and the recombination enhancer (RE) are shown with the approximate distance from the left end of chromosome III. The two mating type cassettes on the left (*HML*) and the right *(HMR*) arm of chromosome III are silenced (shown with diagonal lines). RE is "on" in **a** cells and "off" in α cells. The centromere is depicted as a circle on the chromosome. The HO endonuclease cleavage site is shown with an arrow. The donor preference for each cell type (*MAT***a** or *MAT*α) is shown with a bracket from the *MAT* locus to the preferred cassette (adapted from (Haber 1998)).

The fourth mechanism that controls mating-type switching is called "donor preference". Donor preference ensures that **a** cells choose *HML*α and α cells choose *HMR***a** as donor during switching (Figure 1-4). The directionality of switching makes sure that most of the switching events will result in the change of the mating type. I will discuss mating-type gene conversion and donor preference in the last two subsections of this chapter.

1.2.1 Gene conversion

Mating-type switching is a gene conversion event that replaces the Y sequence at the *MAT* locus with Yα or Y**a** taken from *HML*α or *HMR***a**, respectively (Kostriken et al. 1983). HO endonuclease starts gene conversion by introducing a double strand break at its 24 bp recognition site at the *MAT* locus (Nickoloff et al. 1986). HO enzyme cleavage is stoichiometric; the enzyme cuts once and then becomes inactive (Jin et al. 1997). HO cutting generates 4 bp 3' overhanging ends which are further digested with 5' to 3' exonucleases to produce long 3' ended tails (Kostriken et al. 1983). One of the 3' tails of single stranded DNA invades the homologous sequences at the chosen donor (*HML*α or *HMR***a**). After invasion, 3' end acts as a primer for the synthesis of new DNA and copies the Y region of the donor. This type of recombination which involves gene conversion with DNA synthesis without crossover is called synthesis-dependent strand annealing (SDSA) (Paques et al. 1998).

Mating-type switching is a slow recombination event that employs a number of proteins (White and Haber 1990). Some of these proteins are Rad51, Rad52, Rad54, Rad55 and Rad57 (Hays et al. 1995; Rattray and Symington 1995; Clever et al. 1997). Deletion of each of these proteins exhibits different phenotypes (Rattray and Symington 1995; Sugawara et al. 1995; Haber 1998). Surprisingly, only Rad52 is essential for mating type switching. Recently, chromatin immunoprecipitation experiments were done to determine the association of these proteins with the *MAT* and the donor loci during switching (Sugawara et al. 2003; Wolner et al. 2003). Together with previous studies, these experiments suggest that Rad51, which has homology to the E. coli single strand

DNA binding protein RecA, binds to the single stranded regions at the 3' extended tails and catalyzes exchange reaction during recombination (Shinohara et al. 1992; Ogawa et al. 1993; Sung 1994). Rad52, Rad55 and Rad57 assist Rad51 binding to the single stranded DNA (Sung 1997). Rad54 is thought to play a role in opening up the chromatin structure at the *HM* loci by working as an ATP related chromatin remodeling enzyme (Jaskelioff et al. 2003).

1.2.2 Directionality of mating-type switching

Mating-type switching is based on a directional recombination system. Thus, **a** cells use *HML*α and α cells use *HMR***a** as the donor during recombination more than 85% of the time. This is called "donor preference". Several properties of donor preference in **a** and α cells are summarized in Table 1-1. This summary was generated using data reviewed in (Haber 1998). Directional recombination does not depend on the sequence of the recombining cassettes, rather it depends on the location of the cassettes on the chromosome (Weiler and Broach 1992). For instance, if cassettes are swapped in an **a** cell, this cell still uses *HML* as the donor, and switches from **a** to **a**. The regulatory proteins that are expressed from the *MAT* locus affect donor preference differently in the two cell types. The donor preference of α cells depends on the *MAT*α2 gene but not on *MAT*α1 (Szeto et al. 1997). In **a** cells, donor preference depends neither on *MAT***a**1 nor *MAT***a**2 gene (Wu and Haber 1995). The mechanism of donor preference in the two cell types does not mirror each other. In **a** cells, *HML*α is activated for recombination. Nevertheless, **a** cells can still use *HMR***a** when *HML*α is not available. In contrast, α cells

inactivate *HML*α, thus, use *HMR***a** as the donor. Moreover, when *HMR***a** is not present, α cells cannot use *HML*α instead and approximately one third of the cells fail to repair the double strand break and die. The remaining two third arrest at the G2/M checkpoint and find another way to repair the double stand break at the *MAT* locus (Wu and Haber 1996; Wu et al. 1996; Wu et al. 1997). Remarkably, activation or repression of *HML*α for recombination in **a** and α cells is not confined to the *HML*α sequences, but spread to ~40 kbp of the left arm of chromosome III. In **a** cells, this region is activated for homologous recombination and in α cells, the whole region is "cold" for homologous recombination (Wu and Haber 1995).

Table 1-1: The properties of donor preference during mating-type switching in **a** and α cells

Mating type	**a**	**α**
Donor preference	Use *HML*α > ~85% of the time	Use *HMR***a** >~90% of the time
RE activity	Active (on)	Inactive (off)
Transcription of the non-coding RNAs	Non coding RNAs are transcribed	No transcription
Recombination at the left arm of chromosome III	Left arm of chrIII is hot for homologous recombination	Left arm of chrIII is cold for homologous recombination
When *HML* is deleted	Use *HMR***a** 90% of the time	
When *HMR* is deleted		Fail to use HMR
When RE is deleted	Use *HMR***a** 90% of the time	No change
α2 binding prevented	No change	Use *HML*α 50% of the time & chromatin structure of the RE is altered
Mcm1 binding prevented	Use *HML*α 20% of the time & chromatin structure of the RE is altered	No change
RE in reverse orientation	No change	

Different recombination frequencies for the left arm of chromosome III according to the cell type led to the discovery of a small cis-acting sequence, the recombination enhancer (RE). RE lies in a 2.5 kb intergenic region, located ~29 kb from the left arm of chromosome III (Figure 1-4) (Wu and Haber 1996). The location of RE is important for correct donor preference. RE must reside close to *HML*. Deletion of the entire RE causes a dramatic reduction in **a** cell donor preference, from ~85-90% to ~10-15% *HML* use. The orientation of RE is not important for its function. RE also activates other types of homologous recombination at the left arm of chromosome III in **a** cells, and represses recombination in α cells (Wu and Haber 1995, 1996).

The mechanism of RE function is not clear. The whole RE includes two Matα2p/Mcm1p operators. A number of non-coding RNAs are transcribed from this region (Szeto et al. 1997). A minimal RE, ~750 bp, is located around the first operator and is composed of several conserved elements (Figure 1-5). These elements, also referred to as “domains”, were identified by sequence comparison between *Saccharomyces cerevisiae* and *Saccharomyces carlsbergensis* (Wu et al. 1998). Recent findings, discussed in the next section, assign important functions to these domains.

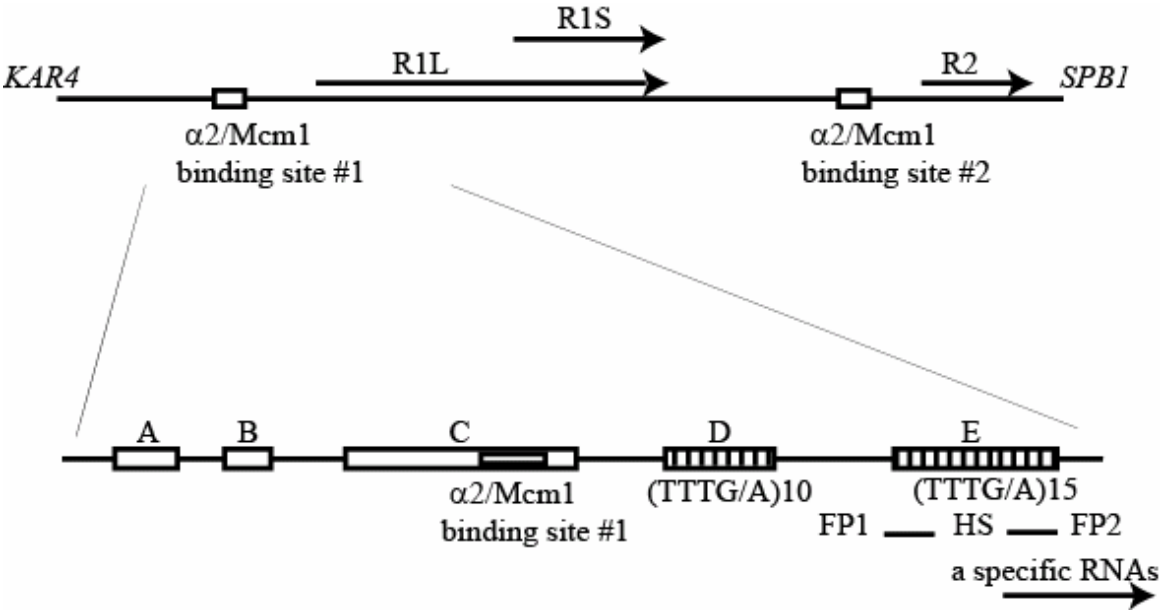

Figure 1-5: Functional domains of RE

RE resides between *KAR4* and *SPB1*. Locations of the two α2/Mcm1 binding sites are shown. RE is transcribed to produce three major non-coding RNAs named R1L, R1S and R2 in **a** cells. R stands for RNA and the number stands for the operator upstream of the transcription. The direction of transcription is depicted with arrows. The domains of RE are shown as boxes. TTTG/A repeats are shown with vertical tick marks. FP1 (Footprint #1), HS (hypersensitive site) and FP2 (Footprint #2) were identified as distinctive properties of the RE chromatin structure in *MAT***a** cells. (Upper panel adapted from (Szeto et al. 1997) and lower panel is adapted from Haber, JE., 1998)

1.3 Chromatin Structure and Transcription at the Recombination Enhancer Loci

Minimal RE consists of five conserved domains, A, B, C, D and E. The most important among these is the C domain that includes the first α2/Mcm1 operator. These operators also function in other parts of the genome by regulating transcription of **a**-specific genes. In α cells, α2 and Mcm1 bind to the operator and in cooperation with the global repressors, Tup1 and Ssn6, repress **a**-specific gene transcription (Smith and Johnson 1994). In **a** cells, α2 is not present, so Mcm1 binds to the operator by itself and

activates transcription. α2/Mcm1 operator functions similarly at the RE. In α cells, α2 and Mcm1p bind to the operator and turns RE "off", and in **a** cells, Mcm1p binds and turns it "on". Different from **a**-specific gene regulation, RE in α cells employs Tup1 independent from Ssn6 (Weiss and Simpson 1997). Furthermore, there are no ORFs downstream of the two α2/Mcm1 operators. Although there are no ORFs, a number of non coding RNAs are transcribed downstream of these operators in **a** cells (Szeto et al. 1997). In α cells, these non-coding RNAs are not transcribed. The chromatin structure of RE also differs in **a** and α cells (Weiss and Simpson 1997). RE has an open chromatin structure in **a** cells, whereas in α cells, it is packed into tightly positioned nucleosomes. The α2/Mcm1 operator governs the chromatin structure of RE similar to that of the **a**-specific genes. This operator also governs RE function. In α cells, mutations that prevent α2 binding reduce donor preference (Szeto et al. 1997). In **a** cells, a two bp mutation that prevents Mcm1 binding to RE also reduces the donor preference from 85-90% to ~20% (Wu et al. 1998).

Other RE domains were studied by mostly deletion analysis. These studies showed that the RE works redundantly. For instance, although the 750 bp minimal RE can give 75% *HML* preference to **a** cells, the presence of the rest of RE (~2kb) confers more than ~85% *HML* preference. Furthermore, individual domain deletions show different levels of reduction in donor preference (Wu and Haber 1996; Wu et al. 1998). For example, if one deletes the B domain, RE still works (Wu and Haber 1996). This is not true for the C domain where the operator is located. Interestingly, ~6 copies of the D domain or 5 copies of the E domain can provide donor preferences as high as 69% *HML* use. These two domains contain a stretch of TTTG/A repeats. One of the distinct

chromatin features in RE encompasses these repeats that house two protein footprints surrounding a nuclease hypersensitive region (Figure 1-5) (Weiss and Simpson 1997). Later, it was shown that three transcription activators, Fkh1, Fkh2 and Ndd1, bind to the TTTG/A repeats in RE (Sun et al. 2002). The mechanism by which these proteins affect RE function is not known. Nevertheless, the deletion of Fkh1 reduces donor preference of **a** cells, from ~85% to ~35%. Deletion of another protein, Ku80, also reduces donor preference in **a** cells, from ~85% to ~40%. Ku80 also binds to RE in **a** cells but not in α cells (Ruan and Simpson, unpublished).

Although it is known that Mcm1, Fkh1, Fkh2, Ndd1 and Ku80 bind to RE in **a** cells, the mechanism of RE function is not known. A number of recent publications investigated whether RE would govern the nuclear positions of the mating type loci, thus physically bringing the correct donor to the *MAT* locus. One of these studies showed that no such predetermined nuclear position occurs (Simon et al. 2002). However, after double strand break, the correct donor rapidly associates with *MAT*. The same study suggested that artificially tethering the wrong donor to *MAT* by using a bridge of lac repressor and operator, fails to affect donor preference. In contrast, another group suggested that artificial tethering of wrong donor affects donor preference, when this donor is *HML* (Kostriken and Wedeen 2001). This result once again stresses the fact that the mechanism of donor preference in **a** and α cells is different. Additionally, there are constraints in the mobility of chromosome III in α cells (Bressan et al. 2004). These studies overall show that nuclear localizations of the *MAT, HML*α and *HMR***a** loci may or may not be important for donor preference, depending on the cell type. Moreover, it is not clear how the correct donor is positioned next to *MAT* after HO break is induced and

whether pre-organized locations of the recombining sequences affect this positioning. Therefore, the mechanism of donor preference is most likely to be controlled by commitment to switching rather than pre-arranged nuclear localization in **a** cells.

After its discovery in 1996, RE proved to have a complex mechanism of action. First, deletion of the conserved domains of RE, by themselves or in combination, affect RE activity at different levels. Second, deletion of the proteins that are required for RE function, do not totally diminish the donor preference of **a** cells. For instance, Fkh1 deletion reduces donor preference from ~85% to ~35%, higher than the use of *HML* in α cells (<10%). Ku80 deletion reduces donor preference to only ~40%. Therefore, I think that the RE may have more than one mechanism of action to ensure correct donor preference. Exploring all these mechanisms may help us understand how mitotic recombination is affected by a small *cis*-acting sequence. Hence, *Saccharomyces cerevisiae* RE serves as a model for understanding how proteins interact with their corresponding *cis*-acting sequences to affect processes that require changes over long chromosomal distances.

In this thesis, I set out to find the mechanisms of RE action. I wanted to know how RE activates homologous recombination of the left arm of chromosome III in **a** cells and how transcription affects RE function and how this transcription relates to the binding of regulatory proteins to RE. The relationship between recombination and transcription has been studied in a number of other systems. These studies showed that the transcription through the donor sequences increases the rate of recombination and this increase might be due to the chromatin remodeling activity within the elongation complex (Aguilera et al. 2000; Chavez et al. 2000; Saxe et al. 2000; Gallardo and

Aguilera 2001). Another well studied system that connects transcription to recombination involves V(D)J recombination in immune cells. This type of recombination is also directional and the donor sequences that undergo recombination are marked by the presence of an active promoter upstream (reviewed in (Oltz 2001)). Although active RE has open chromatin and is transcribed, it does not serve as the donor during mating type gene conversion. Understanding how RE transcription activates recombination may help us understand the relationship between transcription and recombination.

In this first chapter of my thesis, I have introduced DNaseI mapping of primary chromatin structure and yeast mating-type switching. In the second chapter, I will discuss the primary chromatin structure of a ~45 kbp region at the left arm of chromosome III and present evidence that the primary chromatin structure does not underlie the difference in the frequency of homologous recombination between the two mating-types. In the third chapter, I will discuss the role of Mcm1 protein and transcription activation at RE. In the final chapter, I will summarize my studies on donor preference during mating-type switching and make concluding remarks.

Chapter 2

Global chromatin structure of 45,000 base pairs of chromosome III in a and α-cell yeast and during mating-type switching

2.1 Abstract

Directionality of yeast mating type switching has been attributed to differences in chromatin structure for the left arm of chromosome III. I mapped the structure of ~45 kbp of the left arm of chromosome III in **a** and α cells in logarithmically growing cultures and in **a** cells during switching. Distinctive features of chromatin structure were occurrence of DNase I hypersensitive sites in the promoter region of nearly every gene and some replication origins and presence of extended regions of positioned nucleosomes in ~25% of the ORFs. Other than the recombination enhancer, chromatin structure was identical in the two cell types. Changes in chromatin structure during switching were confined to the recombination enhancer and the donor. This unbiased analysis of an extended region of chromatin reveals that significant features of organized chromatin exist for the entire region and these features are largely static with respect to mating type and mating-type switching. My analysis shows that primary chromatin structure does not cause the documented differences in recombinational frequency of the left arm of chromosome III in **a** and α cells.

2.2 Introduction

Mating type interconversion in *S. cerevisiae* is directional. HO endonuclease initiates switching by making a double strand cut at the *MAT* locus on chromosome III. Repair of this break in an α cell preferentially uses silent mating type information encoded at *HMR***a**, while in an **a** cell, recombination usually takes place with sequences located at *HML*α (Haber 1998). The recombination enhancer (RE), located ~30,000 base pairs (bp) from the left end of chromosome III and highly conserved between yeast species, controls this directionality (Wu and Haber 1996; Wu et al. 1996; Szeto et al. 1997; Wu et al. 1998). In **a** cells, the RE has nuclease hypersensitive sites indicating unusual DNA geometry and footprints suggesting protein binding sites. Some of the protein candidates which may create these footprints are the two transcription activators, Fkh1p and Fkh2p, and their associated protein Ndd1p which were shown to bind to the RE directly only in **a** cells (Sun et al. 2002). In α cells, organized chromatin abutting the binding site for Mcm1p and Matα2p replaces the active **a**-cell configuration (Weiss and Simpson 1997). How these structural differences at the RE facilitate interactions between *MAT* and *HML*α in **a** cells is not known. It is known, however, that extensive segments of the left arm of chromosome III are altered in their recombination properties in the two cell types. In **a**-cells, at least 40 kbp at the left end of this chromosome is in a hyperactive state for recombination, while over 100 kbp in this region is recombinationally "cold" in α cells (Wu and Haber 1995).

Alterations in chromatin structure have been invoked as a possible explanation of these differences in recombination potential (Wu and Haber 1995). Chromatin structure

of only two regions of chromosome III has been reported. RE structure is strikingly different in α and **a** cells. *HML*, has distinctive chromatin structure but it is identical in the two cell types (Weiss and Simpson 1998). Differences in the chromatin structure of the left arm of chromosome III in the two cell types and changes in chromatin structure accompanying switching are unstudied and unknown. In this study, I mapped the chromatin structure of ~45 kbp of the left arm of chromosome III. The analysis was carried out for both α and **a** cells and for the latter cell type during mating type switching.

The single most informative probe of chromatin structure is probably DNase I (Weintraub and Groudine 1976; Wu and Gilbert 1981; Simpson and Stafford 1983; Lutter 1989; Staynov and Proykova 1998). In general, DNase I hypersensitivity marks the sites where chromatin is more open, making hypersensitive site synonymous with enhancer, promoter, replication origin, or other similar features of chromatin. All of these features enhance the utility of DNase I as the primary instrument in dissection of chromatin structure of large regions of a genome.

In the process of addressing the role of chromatin in mating type switching, I carried out the first nonbiased study of chromatin structure of an extensive region in an organism whose genomic sequence is known. In contrast to higher organisms, where chromatin structure undergoes major changes when genes are repressed or transcribed, hypersensitive sites are present at the 5' ends of most genes and some replication origins in the left arm of yeast chromosome III. A surprising finding is the presence of organized arrays of positioned nucleosomes on ~25% of the genes examined. A unique change in chromatin structure occurs at the RE during mating type interconversion in **a**-cells, even though the structure of the remainder of the left arm of chromosome III is nearly identical

in the two cell types. The chromatin structure of the entire region is same for the two cell types except in the RE. This is in contrast to expectations of chromatin structural differences of the left arm of chromosome III as the explanation for observed cell type specific differences in recombinational frequency (Wu and Haber 1995).

2.3 Materials and Methods

2.3.1 Cell growth

Standard yeast media, YP, with an appropriate carbon source (2% dextrose, 2.5% lactose) were used. For comparison of the chromatin structures of **a** and α cells, FY24 α and FY23**a** cells (Winston et al. 1995) were grown at 30ºC to OD_{600} =1.0 in YPD.

For chromatin structure analysis of switching cells, JKM161**a** (*ho HMLα Mata HMRa ade1-112 lys5 leu2-3 ura3-52 trp::hisG, ade 3:: GalHO),* provided by J. Haber, was used. Induction of HO expression was done as described (White and Haber 1990). Cells were grown to OD_{600}=0.2 in 900 ml YPL at 30°C. They were synchronized by adding alpha factor to a concentration of 10 μg/ml. When >90% of the cells were unbudded, 100 ml of 20% (w/v) galactose was added to induce HO expression and cutting for 30 min. After induction, cells were harvested by filtration and suspended in prewarmed YPD. At appropriate time points, NaN_3 was added to 1 L culture to a final concentration of 0.2% and cells were harvested.

For mapping of the *FLO8* chromatin structure, wild type (*SPT6*) and mutant (*spt6-1004*) cells were grown at permissive (30^oC) temperatures until OD_{600} ~ 0.8.

Temperature shift was done by mixing warm media to the culture to bring up the temperature to 39^{o}C and cells were grown at 39^{o}C for additional 80 minutes. The SPT6 temperature sensitive strain, along with the wild type, was provided by F. Winston (9).

HML chromatin structure was mapped in **a** cells lacking sir3 (Δ*sir3*), along with the wild type cells. These strains were obtained from the yeast deletion library (Open Biosystems).

2.3.2 Nuclear DNA preparation and analysis

Nuclei from 1 L cells of OD_{600}=0.4-1.5 were prepared as described (Roth and Simpson 1992). After harvesting, cells were washed twice with 30 ml Sorbitol Buffer (1.4 M Sorbitol, 40 mM HEPES; pH 7.5, 0.5 mM Mg Cl_2), containing 1 mM PMSF, 10 mM B-mercaptoethanol. Cell pellet was weighed, resuspended in 4XW (weight in grams) ml of Sorbitol buffer containing 0.5 mg/ml Zymolyase (Seigaku), for ~30 min, at 30^{o}C. After Zymolyase treatment, the following steps were performed at 4^{o}C. Cells were washed twice with Sorbitol Buffer + 1 mM PMSF and resuspended in 20 ml Ficoll Buffer (18% Ficoll 400, 20 mM PIPES; pH 6.5, 0.5 mM Mg Cl_2) + 1 mM PMSF. The sample was homogenized by Thomas Teflon pestle tissue homogenizer with drill. After homogenization, it was layered onto 20 ml Glycerol Buffer (7 % Ficoll 400, 20% Glycerol, 20 mM PIPES; pH 6.5, 0.5 mM Mg Cl_2) + 1 mM PMSF. The nuclei were precipitated by spinning in Sorvall rotor HB-6 for 30 minutes at 11,500 rpm. Pellet was resuspended in 20 ml Ficoll Buffer+ 1 mM PMSF by vortexing for 5 minutes and samples were centrifuged at 4,500 rpm for 15 minutes. The supernatant, containing the

nuclei was centrifuged for 30 minutes at 11,500 rpm. The pellet, containing the nuclei, was washed in 10 ml Digestion Buffer (10 mM HEPES; pH 7.5, 0.5 mM Mg Cl_2, 0.05 mM $CaCl_2$) and re-centrifuged for 15 minutes. After wash, nuclei were resuspended in 2 ml of digestion buffer and digestion was performed for each 400 µl aliquot. Nuclear DNA was cut with DNase I (Worthington) at concentrations of 0.1-0.8 U/ml or 2.5- 10 u/ml MNase I 37^{o}C, for 5 minutes. DNA was recovered after ProteinaseK and RNaseA treatment and Phenol/Chroloform extraction.

2.3.3 Southern blots

For mapping blots, DNA from equal cell numbers (based on weight of cell pellets) was purified, cut with *Bam*HI, subjected to electrophoretic separation on a 1.2% agarose gel, transferred to Hybond-NX membrane (Amersham), cross linked with UV light and hybridized with specific probes. The probes were prepared by PCR amplification of ~250 bp regions abutting each *Bam*HI cut site. The specificity of each probe was determined by running Blast searches for probe sequences. Undigested control in the maps also serves as a measure of the specificity of each probe, because it is expected that cross hybridization would result in the appearance of multiple bands in these samples. Probes were gel purified and random primer labeled with [α-^{32}P]dATP. Blots were exposed to a PhosphorImager screen and analyzed using ImageQuant v.5.

For mating-type switching assays, DNA was recovered from cells by glass bead disruption and Phenol/Chloroform extraction. After digesting DNA with StyI, samples were separated in denaturing gels and Southern Blot was performed. 1.2% Agarose gel

was prepared in Buffer 1 (50 mM NaCl, 4 mM EDTA). The gel was soaked in Buffer 2(30 mM NaOH, 2 mM EDTA) for 2 hours. Samples were prepared in 2X Loading buffer (50% Glycerol, 1 M NaOH, 0.05 % Bromocresol Green). After boiling for 5 minutes, they were loaded. After dye reached 2/3, the gel was neutralized in 0.1 M Tris.Cl; pH 7.5 for 1 hour. Blotting was performed and *MAT* specific probe was used to detect 2.2 kb distal fragment, 930 bp *MAT***a**, 1880 bp *MAT*α, and 723 bp HO cut fragment (White and Haber 1990).

2.3.4 RNA analysis

Total RNA was prepared by glass bead cell disruption as described previously (Zhang and Reese 2004). For Northern blotting, 20 μg of each RNA sample was subjected to electrophoresis in a formaldehyde/agarose gel, followed by transfer to a nylon membrane, and hybridization to specific probes. Probes for Northern blot hybridization were prepared by amplification from genomic DNA by PCR. Gel purification of the probes was followed by random primer labeling (Stratagene). EtBr staining of rRNA was used as loading control. Blots were stripped for re-probing by boiling in 0.1% SDS. FLO8 probe: Chr V, 375938-375232, YCL065: Chr III, 14001-14119, Alpha2: Chr III, 12957-13257, scR1: Chr V, 441736-442428.

2.4 Results

Restriction endonuclease *Bam*HI cuts the leftmost ~55 kbp of chromosome III into fragments that range from 6 to 11 kbp in size (Figure 2-1). Since an individual indirect end label experiment can map 5-7 kbp of chromatin at medium level resolution,

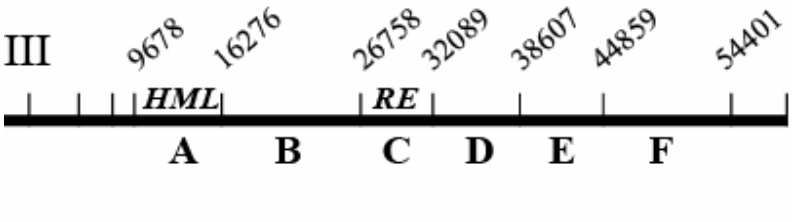

Figure 2-1: BamHI sites on the leftmost 58 kbp of yeast chromosome III
Restriction sites are indicated by the vertical tic marks and, for those used for mapping additionally by the coordinates of the sites. Locations of *HML* and the recombination enhancer are indicated. Letters below the line indicate the subfigure in Figure 2-2 presenting the chromatin map for that region.

mapping of the entire region of interest was possible using a single set of DNase I digests, repetitively separated by gel electrophoresis and/or repetitively probed on Southern blots. Since all the probes were generated by polymerase chain reaction amplification of genomic DNA, were about the same length, were labeled in parallel using the same conditions for random primer labeling, and were hybridized and washed identically, it is possible to make semi quantitative conclusions about relative nuclease susceptibilities from inspection of the blots. Mapping of the first 8 kbp of the chromosome was not possible due to common subtelomeric sequences between chromosome III and other chromosomes. Figure 2-2 presents indirect end label maps of chromatin from 9 to 54 thousand map units (kmu) for both **a** and α cells.

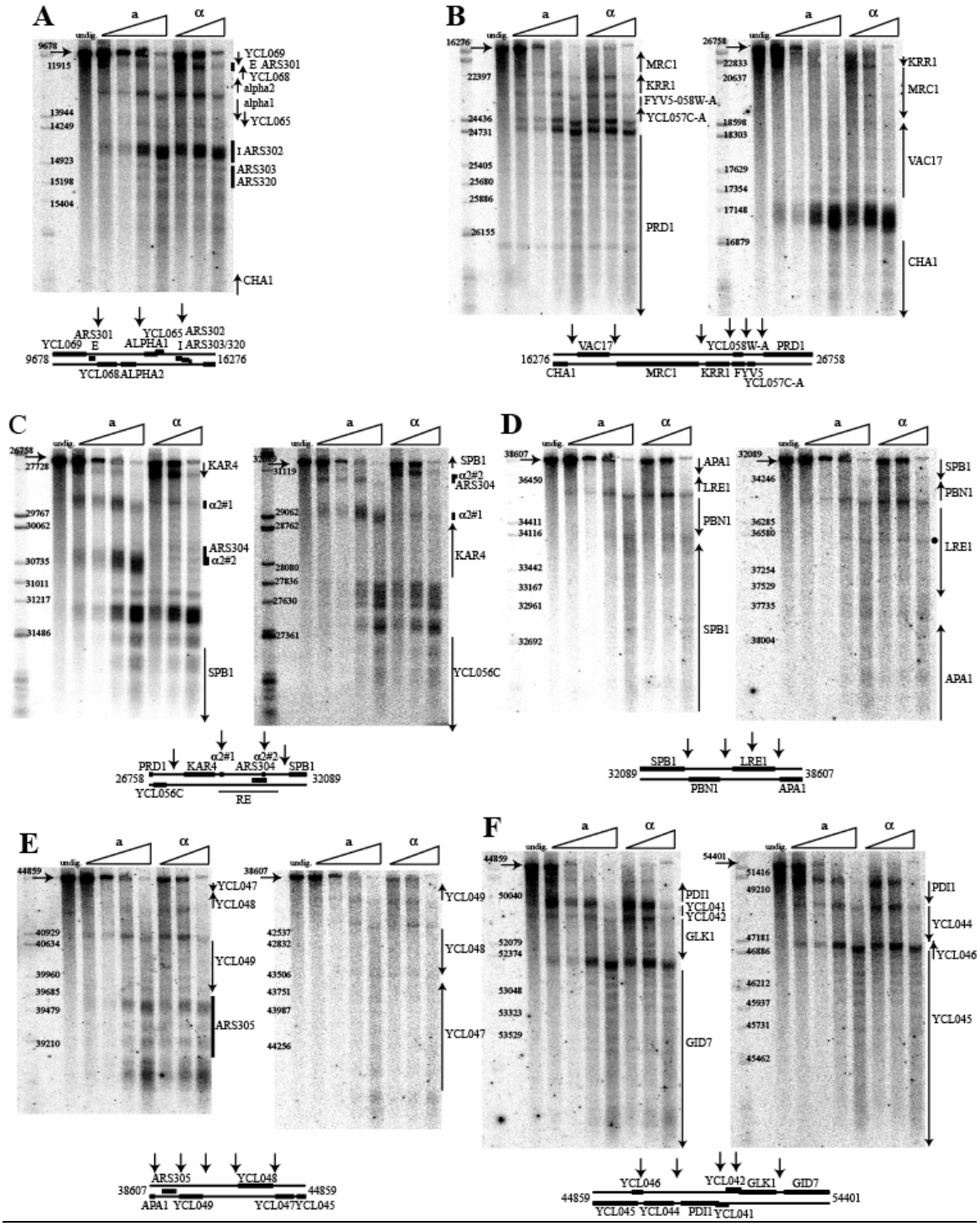
A
B
C
D
E
F
a
α
undig.
YCL069
E
ARS301
YCL068
alpha2
alpha1
YCL065
I
ARS302
ARS303
ARS320
CHA1
MRC1
KRR1
FYV5-058W-A
YCL057C-A
PRD1
VAC17
YCL058W-A
FYV5
KAR4
α2#1
ARS304
α2#2
SPB1
YCL056C
RE
APA1
LRE1
PBN1
YCL049
YCL048
YCL047
ARS305
PDI1
YCL041
YCL042
GLK1
GID7
YCL046
YCL044
YCL045

Figure 2-2: Chromatin structure of an extensive region of the left arm of chromosome III

For each panel, an undigested control is shown together with four levels of digestion of **a**-cell nuclei (0.1, 0.2, 0.4, 0.8 U/ml) and three levels of digestion of α-cell nuclei (0.2, 0.4, 0.8 U/ml). The genetic features of each region are shown to the right of each gel pattern and below, where the arrows indicate major DNase I digestion features. The direction of transcription of each gene was shown below each gel by placing the genes on Watson or Crick strand of DNA fragment that was mapped. Numbers adjacent to size standards are coordinates for chromosome III. Each parent fragment is shown with an arrow and the coordinate for the restriction site to the left of each gel. All maps are bi-directional except for that presented in panel A, where similar subtelomeric sequences between chromosomes precluded mapping to the left of 9678. A: The region from 9678 to 16276. Locations of two ARS sequences are shown, as are the positions of the E and I silencers. B: The region from 16276 to 26758. C: The region from 26758 to 32089. Locations of the two Mcm1p/Matα2p binding sites in the recombination enhancer are shown. D: The region from 32089 to 38607. E: The region from 38607 to 44859. Location of ARS 305 is indicated. F: The region from 44859 to 54401.

Naked genomic DNA was also cut with DNase I and mapped. The digestion of naked DNA with DNase I generated uniform smears in almost all cases and I did not observe bands corresponding to the hypersensitive sites detected in the chromatin maps Figure 2-3. Features of the genome sequence are presented at the side of each experimental map and summarized at the bottom of each data set, together with vertical arrows that indicate major DNase I cutting sites. A list of the genetic features within the mapped region, obtained from *Saccharomyces* Genome Database (http://www.yeastgenome.org) and the transcriptional frequency of most of the genes (Holstege et al. 1998) are summarized in Table 2-1.

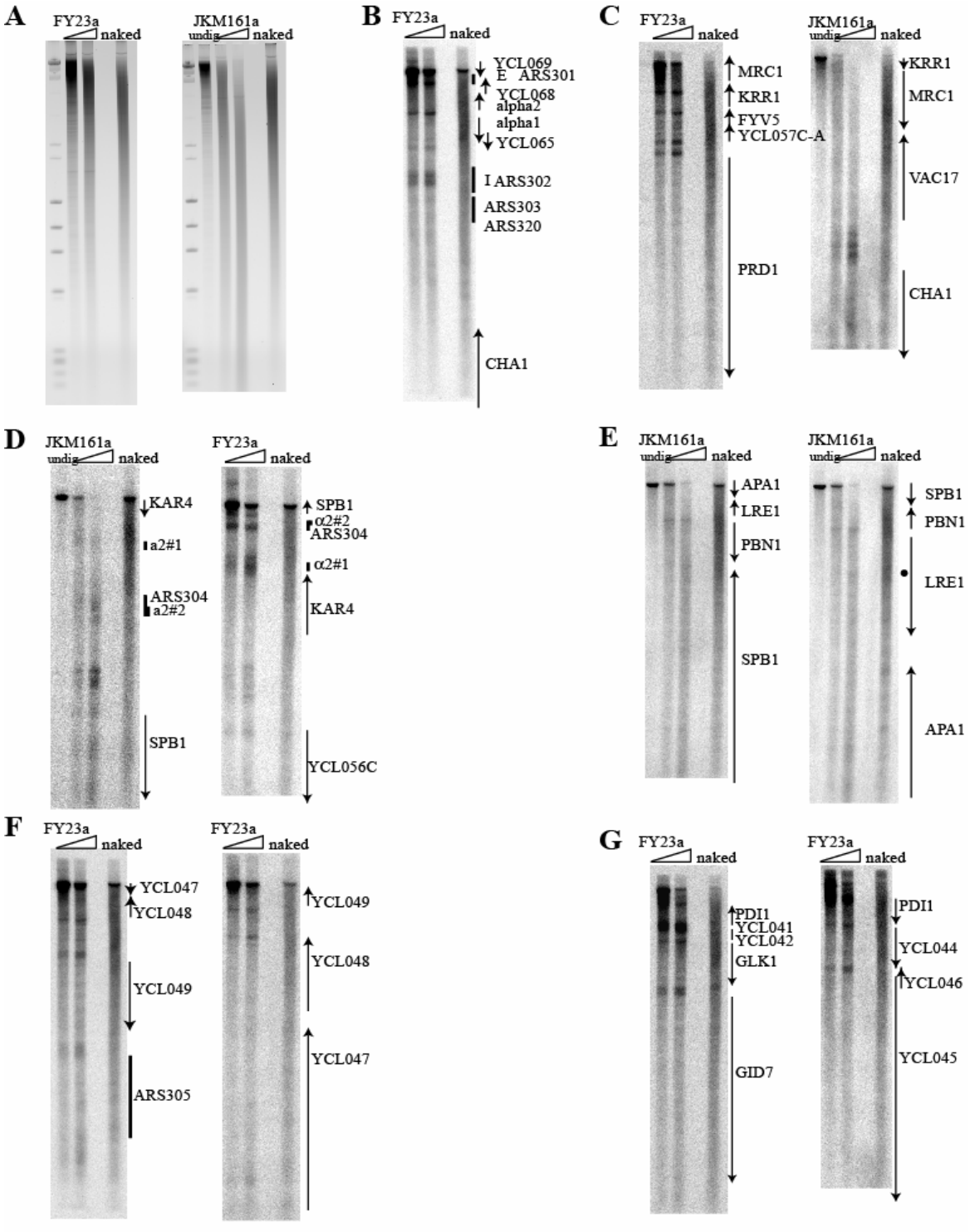
A
FY23a
naked
JKM161a
undig
naked
B
FY23a
naked
YCL069
E ARS301
YCL068
alpha2
alpha1
YCL065
I ARS302
ARS303
ARS320
CHA1
C
FY23a
naked
JKM161a
undig
naked
MRC1
KRR1
FYV5
YCL057C-A
PRD1
KRR1
MRC1
VAC17
CHA1
D
JKM161a
undig
naked
FY23a
naked
KAR4
a2#1
ARS304
a2#2
SPB1
SPB1
α2#2
ARS304
α2#1
KAR4
YCL056C
E
JKM161a
undig
naked
JKM161a
undig
naked
APA1
LRE1
PBN1
SPB1
SPB1
PBN1
LRE1
APA1
F
FY23a
naked
FY23a
naked
YCL047
YCL048
YCL049
ARS305
YCL049
YCL048
YCL047
G
FY23a
naked
FY23a
naked
PDI1
YCL041
YCL042
GLK1
GID7
PDI1
YCL044
YCL046
YCL045

Figure 2-3: DNase I maps of naked DNA

Two blots were employed to map naked DNA over the entire region. Along with naked DNA, both blots contain DNA prepared from nuclei that were digested with 0.2 and 0.4 u/ml of DNase I. The first blot includes DNA from FY23**a** cell nuclei. Second blot contains DNA from JKM161**a** cell nuclei and an undigested control. For both blots, the same naked DNA sample was used. Naked DNA sample was prepared by digesting FY23**a** DNA with 0.1 u/ml of DNase I at 37 ^{o}C for 5 min. For each gel, the naked DNA lane is indicated. The genetic features of each region are shown to the right of each gel pattern. A: The level of digestion of naked DNA compared to that of the chromatin sample is shown in ethidium bromide staining of the two gels. The left and right panels show naked DNA along with DNA from FY23**a** and JKM161**a** nuclei, respectively. B: the region from 9678 to 16276. C: the region from 16276 to 26758. D: the region from 26758 to 32089. E: the region from 32089 to 38607. F: the region from 38607 to 44859. G: the region from 44859 to 54401.

Table 2-1: Features and expression profiles between coordinates 9 and 55 kbp of chr III

Feature	Locus description	Expression level (copies/cell)	Transcription frequency (mRNAs/hr)
YCL069W	Hypothetical ORF	n/a	n/a
ARS301	Autonomously Replicating Sequence		
YCL068C	Hypothetical ORF	0.2	n/a
YCL067C	HMLALPHA2	n/a	n/a
YCL066W	HMLALPHA1	n/a	n/a
YCL065W	Hypothetical ORF	n/a	n/a
ARS302	Autonomously Replicating Sequence		
ARS303	Autonomously Replicating Sequence		
ARS320	Autonomously Replicating Sequence		
YCL064C	CHA1 catabolic serine (threonine) dehydratase	4.5	16
YCL063W	VAC17	0.2	0.4
YCL061C	MRC1 protein involved in replication checkpoint	0.2	0.6
YCL059C	KRR1 involved in cell division and spore germination	1.9	5.8
YCL058W-A	Identified by homology to Ashbya gossypii	n/a	n/a
YCL058C	FYV5	n/a	n/a
YCL057C-A	Hypothetical ORF, has similarity to proteins in S. pombe, C. elegans, D. melanogaster.	n/a	n/a
YCL057W	PRD1	0.9	n/a
YCL056C	Protein of unknown function; (GFP)-fusion protein localizes to the cytoplasm	2.0	3.5
YCL055W	KAR4 transcription factor involved in karyogamy	1.3	3.3
ARS304	Autonomously Replicating Sequence		
YCL054W	SPB1 Putative methyltransferase	1.8	3.3
YCL052C	PBN1 Protease B, nonderepressible form	0.9	0.5
YCL051W	LRE1 involved in laminarase resistance	0.2	n/a
YCL050C	APA1	4.3	n/a
ARS305	Autonomously Replicating Sequence		
YCL049C	Hypothetical ORF	0.7	n/a
YCL048W	Hypothetical ORF	n/a	n/a
YCL047C	Hypothetical ORF	0.8	1.9
YCL046W	Hypothetical ORF	n/a	n/a
YCL045C	Hypothetical ORF	1.6	1.9
YCL044C	Hypothetical ORF	0.3	0.2
YCL043C	PDI1 protein disulfide isomerase	6.7	11.6
YCL042W	Hypothetical ORF	0.4	0.4
YCL041C	Protein required for cell viability	n/a	n/a
YCL040W	GLK1 Glucokinase	3.7	3.8
YCL039W	GID7	0.2	0.2

n/a: For indicated ORFs, data are not present in the analysis of Holstege, et al.

2.4.1 Chromatin structure features common to both cell types

Most of the structure map was very similar for the two cell types. Specifically, there was no indication of a generalized increase in nuclease susceptibility for the highly recombination competent **a**-cell nuclei compared to the nuclei from less recombinogenic α cells. Visual comparison of the rate of disappearance of the parent restriction endonuclease DNA fragment in the two cell types during the time course of DNase I digestion Figure 2-2 revealed minor variations but no consistently enhanced rate of nuclease digestion for **a** versus α cells. There is one significant exception to this generalization. For a region where chromatin structure was known to differ for the two cell types, the RE (Weiss and Simpson 1997), the parent fragment was lost at lower nuclease concentrations in **a** than in α-cell nuclei (Figure 2-2C), consistent with a protected chromatin domain in α cells. Because the DNase I sensitivity of the RE mirror known chromatin structure differences, the absence of similar differential nuclease sensitivity for other regions strongly suggests that there are no global differences in chromatin structure of **a** and α cells for the leftmost 55 kb of chromosome III.

The first region mapped, from 9 to 16 kmu, includes the *HMLα* locus (Figure 2-2A). The E and I silencers which flank the locus were hypersensitive to DNase I; these sites are also locations of autonomous replicating sequences (ARSs) 301 and 302, respectively. Whether the altered chromatin structure results from silencer or replication origin function is uncertain. Two additional ARSs in this region, 303 and 320, were not nuclease sensitive. As expected from the high resolution analysis of this locus (Weiss and Simpson 1997), the intergenic promoter region between the α1 and α2 genes was also

nuclease sensitive, in spite of the transcriptional inactivity of the two genes. This is an exception to the generality that silenced gene promoters are nuclease insensitive while transcriptionally competent gene promoters are nucleosome-free and nuclease sensitive. The possibility of a role for Rap1p binding in both repression and maintaining a nucleosome-free sensitive site is attractive (Yu and Morse 1999). Both in the YCL065 ORF and between the I silencer and the 3' end of the *CHA1* locus, there were patterns of nuclease sensitivity, a "ladder", indicative of arrays of positioned nucleosomes.

Mapping the next region, from 16 to 27 kmu, revealed a number of features of chromatin structure which were characteristic of much of the entire ~45 kbp segment of chromosome III (Figure 2-2B). Hypersensitive sites occurred near the 5' and/or 3' ends of ORFs, most frequently the former. Strong hypersensitive sites were present in two intergenic promoters of divergently transcribed genes, YCL057 - *PRD1* and YCL063 (*VAC17*) - *CHA1*. The hypersensitivity at the *CHA1* promoter is consistent with previously published data (Moreira and Holmberg 1998). Note that there are two distinct hypersensitive sites marking the 5' end of YCL057 and *PRD1*. Separate hypersensitive sites for the 5' end of some genes were possible to observe when the region between two genes was long enough to resolve in the mapping gels. Some genes, predicted to have low transcriptional activity (Table 2-1) (Holstege et al. 1998), had positioned nucleosomes over extensive regions of the ORF. Examples in the Figure 2-2B map are *PRD1* and *VAC17* which displayed long stretches of regular DNase I cutting sites with a periodicity of ~160 bp. These regions, suggestive of highly organized, sequence specific chromatin structure, were unexpected. Both genes are low activity with 0.9 or 0.2 mRNA molecules per cell, respectively (Holstege et al. 1998). The coexistence of a domain of

positioned nucleosomes within an ORF with a nuclease hypersensitive promoter region 5' of the gene suggests that establishment of transcriptional competency can proceed in the absence of extensive levels of transcription. Note that all these data derive from studies of unique, genomic yeast genes, obviating population considerations that may accompany studies of multicopy minichromosomes. The caveat that structures and expression may vary in cell-cycle dependent fashion is apparent; there is no indication of cell-cycle specific regulation of the cited genes (Spellman et al. 1998).

The region from 27 to 33 kmu of chromosome III contains the recombination enhancer. The only cell type specific features of chromatin structure were located in this segment of the chromosome (Figure 2-2C) and are discussed in depth below.

From 32 to 39 kmu, centromere proximal of the RE, little was distinctive in the DNase I map of chromatin structure (Figure 2-2D). Nuclease sensitive sites were present at the intergenic 3' ends of a gene pair, *SPB1-PBN1*. An unusual hypersensitive site was present in the middle of the *LRE1* gene which was highlighted with a dot located to the right of the mapping data. In mapping nearly 45 kbp of chromatin, this is the only DNase I hypersensitive site that was located internally in an open reading frame sequence. Diffuse ladders of cutting sites were present in the 3' halves of the *SPB1* and *APA1* genes. These were not nearly as sharp, defined bands as those present to the right of the I silencer or in the *PRD1* and YCL063 ORFs. The pattern in the 32 to 39 kmu region is suggestive of an imprecisely organized chromatin structure.

Many features of DNase I cutting in the 39 to 45 kmu region of chromosome III was reminiscent of inferred structures for more distal regions of the chromosome (Figure 2-2E). Hypersensitive regions were present in the 5' flanking regions of YCL048

and YCL049. Diffuse ladders of cutting sites suggested a poorly organized set of positioned nucleosomes in the ORFs for YCL047 and YCL049. Note that the 3' regions of the two convergent ORFs, YCL048 and YCL047 did not show strong hypersensitivity, suggesting that the 3' of the genes do not posses the same chromatin structure with the 5' end of the genes. Differing from features of previously described regions was the presence of a long segment with distinctive chromatin structure in a genomic sequence that lacked any identified ORF, downstream of YCL049. A DNase I hypersensitive site 3' of the ORF abutted three diffuse cutting sites, suggesting the presence of three nucleosomes. The region was bracketed on the opposite end by a strong DNase I hypersensitive site at the location of ARS305. Whether the chromatin structure of this region is critical for function of this ARS, or is a consequence of the activity of the replication origin, remains to be determined.

The density of genes in the final chromatin region whose structure was mapped, 45 to 54 kmu of chromosome III, is very high, making it difficult to distinguish the several DNase I hypersensitive sites association with the 5' or 3' flanking sequences of a particular gene or ORF (Figure 2-2F). This region did exemplify two different types of gene structure. YCL039 (*GID7*) and YCL045 are predicted to have low activity (0.2 and 1.6 copies of mRNA, respectively). Both of these ORFs were packaged in arrays of positioned nucleosomes. In contrast, *PDI1* and *GLK1* are genes of moderate activity (6.7 and 3.7 copies of mRNA, respectively). These two genes have strong DNase I hypersensitive sites at their 5' and 3' flanking regions. The presence of two different types of genes in close proximity to each other indicates that long-range organization of chromatin regions with similar transcriptional regulation and chromatin structure, thought

to occur in some cases in higher organisms (Hebbes et al. 1994; Anguita et al. 2001; Litt et al. 2001), is not a rule in yeast.

To gain more insight into the relationship between transcription and 5' hypersensitivity, I tested whether transcription from cryptic promoters would induce formation of a hypersensitive site similar to the ones found at the 5' end of the ORFs. The transcription from the cryptic promoter of *FLO8* was induced by using a temperature sensitive strain of the elongation factor, encoded by *SPT6* (Kaplan et al. 2003). At higher temperature, *spt6-1004* cells initiated transcription from the *FLO8* cryptic promoter, monitored by the short FLO8 transcript, with northern blot analysis (Figure 2-4A). The DNase I map of the chromatin structure around *FLO8* did not identify a DNase I hypersensitive band around the cryptic promoter (highlighted with a rectangular box in Figure 2-4B). Instead, an overall increase in sensitivity to nuclease digestion was evident by the smear throughout the ORF in *spt6-1004* cells compared to wild type cells (*SPT6*). This smear may mask the hypersensitive site formed at the 5' end of cryptic transcription. The formation of a smear along the gene is consistent with the hypothesis that *spt6-1004* may cause a defect in the chromatin structure over the whole gene (Kaplan et al. 2003). Since the results from the cryptic promoter were inconclusive, I decided to test whether transcription from a normally silenced and heterochromatic region would generate a DNase I hypersensitive band. *HML* locus was de-repressed by deleting *SIR3*. When *SIR3* was deleted, α2 and *YCL065* were expressed (Figure 2-5A). *HML* chromatin was mapped in cells lacking Sir3 (Δ*sir3*) or in wild type cells (wt **a**) (Figure 2-5B). In Δ *sir3* cells, the transcription of *YCL065* and α2 caused the appearance of two strong hypersensitive sites

at the 5' and 3' end of *YCL065*. This result suggests that transcription may induce formation of DNase I hypersensitive sites abutting a gene.

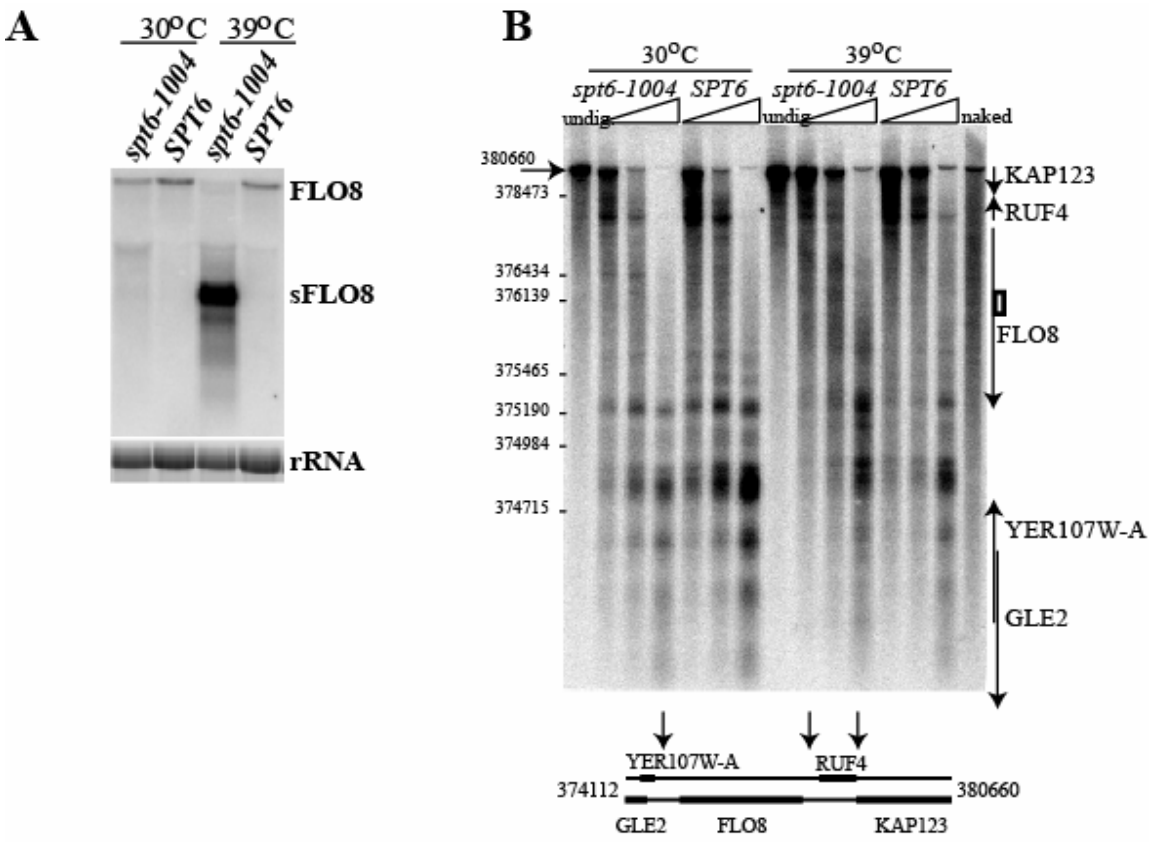

Figure 2-4: Chromatin structure of a region containing *FLO8* in wild type and *spt6-1004* at permissive and non permissive temperatures

For each panel, an undigested control is shown together with three levels of digestion of wild type and mutant nuclei. A naked DNA control is shown on the last lane of the gel. The genetic features of the region are shown to the right of the gel and below, where the arrows indicate major DNase I digestion features. The rectangle on the *FLO8* represents the approximate location of the cryptic promoter. Numbers to the left of the figure are chromosomal coordinates. A: Northern blot analysis for *FLO8* transcription. Transcript from the cryptic promoter is designated as s*FLO8*. rRNA is used as loading control. B: Chromatin structure of a region containing *FLO8* was determined by digesting highly purified nuclei from wild type (*SPT6*) and mutant (*spt6-1004*) cells grown in permissive (30°C) and non permissive (39°C) temperatures. Temperature shift was done by mixing warm media to bring up the temperature to 39°C and cells were grown at high temperature for an additional 80 minutes.

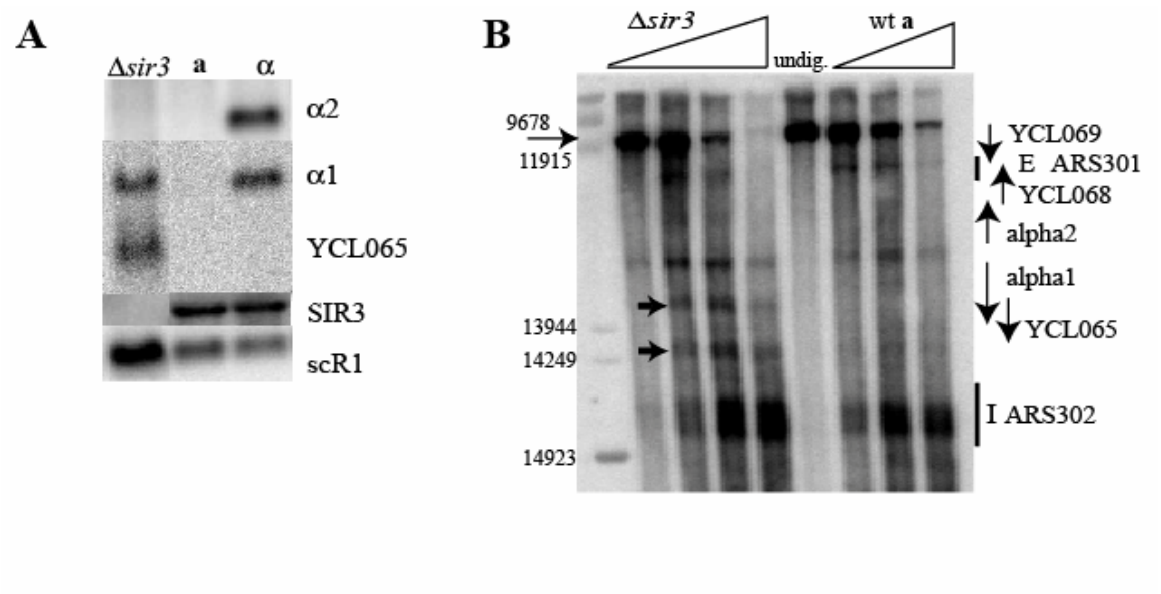

Figure 2-5: Silencing and chromatin structure at *HML*

A: Total RNA from mutant (Δ*sir3*) and wild type **a** and α cells were isolated and analyzed with Northern blot for expression of α1, α2, *YCL065* and *SIR3*. scR1 is used as loading control. B: Highly purified nuclei from Δ*sir3* and wild type (wt **a**) cells were digested with increasing levels of DNase I shown together with an undigested control. The genetic features of the region are shown to the right of picture. The numbers on the left indicate the coordinates in the genome. The two hypersensitive sites that occur in the mutant cells are depicted with arrows.

2.4.2 Distinctive chromatin structures in a- or α-cell nuclei

High resolution analysis of micrococcal nuclease digestion sites in the RE previously showed major differences in chromatin structure between **a** and α cells (Weiss and Simpson 1997). This conclusion was reinforced by the results of the current DNase I study of the region of chromosome III from 27 to 33 kmu (Figure 2-2C). The RE DNA contains two binding sites for Mcm1p and Matα2p, α2#1 and α2#2, separated by ~1350

bp, and a region with extensive TTTG/A repeat sequences just to the right of α2#1 (Oliver et al. 1992).

To the left of α2#1, covering the *KAR4* ORF, positioned nucleosomes were present in the nuclei of both cell types. In contrast, the DNase I pattern of this region indicated two domains. The ORF for *KAR4* was generally nuclease resistant, with diffuse, minor cutting sites marking linker regions between nucleosomes that were not precisely positioned. The intergenic promoter region between the 5' ends of *KAR4* and *YCL056* is more sensitive to DNase I in both cell types. One strong site adjacent to *YCL056* probably relates to its transcription; the origin of the other sites in this region is unknown. At the other end of the RE, a strong DNase I hypersensitive site was present in both cell types flanking the *YCL054* ORF. Three diffuse cutting sites occupy the ~200 bp space between this site and the 5' end of the ORF.

Both the operator sites, α2#1 and α2#2, are DNase I hypersensitive in **a** cells. The cut sites encompass a broad region, larger than the 11 bp site that interacts with Mcm1p (Sauer et al. 1988) a *trans*-acting factor that binds these sites in **a** cells. Other proteins, Fkh1p, Fkh2p, Ndd1p were shown to bind to the RE domains around the α2#1 operator (Sun et al. 2002). In α cells, site α2#1 was slightly sensitive to DNase I while site α2#2 was not cut preferentially at all. This is surprising, since the half-life of Matα2p is very short (Hochstrasser et al. 1991) and the protein is expected to be absent from its binding site in isolated nuclei from α cells (Murphy et al. 1993). Thus, only Mcm1p should bind these DNA sequences in isolated nuclei.

Between the α2#1 operator and the hypersensitive site to the 5' side of *SPB1* in α-cells, a continuous ladder of DNase I cutting sites signaled the presence of an array of positioned nucleosomes. Organized chromatin spanned the second operator without apparent interruption, at least at the level of resolution afforded by indirect end label mapping of DNase I cutting sites. The highly organized chromatin in α-cells (Weiss and Simpson 1997) may fold into a higher order structure that precludes access of the nuclease to the second operator. The region between the two operators was relatively resistant to this nuclease and was featureless in its digestion pattern in **a**-cells.

2.4.3 Chromatin structure changes during mating type interconversion in a-cells

Next, I evaluated changes in chromatin structure during mating type interconversion in **a** cells. Normally, expression of HO endonuclease is confined to mother cells in the G1 phase after cell division (Strathern and Herskowitz 1979). This group comprises a small proportion of a logarithmically growing yeast culture, so others have developed a system that allows for highly synchronous switching of an entire yeast culture (Connolly et al. 1988); I have employed this system for the study of chromatin structure during switching.

Asynchronously growing cultures of **a** cell *S. cerevisiae* were treated with the peptide α-factor to arrest cells in G1. The cells used are an **a** strain that expresses the HO endonuclease gene under galactose (GAL) control. Galactose was added to the medium to induce synthesis of HO endonuclease. After 30 minutes, cells were transferred to

glucose medium which represses expression of GAL controlled genes. Cells then progressed through a typical mating type interconversion.

Others have devised a Southern blotting protocol that allows one to follow the time course of all aspects of switching in **a** cells (Connolly et al. 1988). Distinctive DNA fragments mark the *MAT***a** locus, prior to or after cutting by HO endonuclease, the *MAT*α locus, and a loading control derived from common sequences to the right of *MAT* (Figure 2-6B). Figure 2-6A documents the switching kinetics in my experiments. At -30 min (end of α-factor synchronization, beginning of HO induction), cells are *MAT***a**, as shown by the *MAT***a** fragment. At -15 (mostly) and 0 (completely), HO has cut the *MAT* locus. The fragments are now shorter than the *MAT***a** species. Over the next 60-90 min, interconversion occurs with the bulk of the cells becoming *MAT*α, signaled by the large *MAT*α fragment. A small fraction of cells correct the double strand DNA break by recombination at *HMR***a**, leading to regeneration of the *MAT***a** fragment. An important control for experiments which require isolation of nuclei and DNase I digestion at closely spaced time points during synchronized switching is demonstration that arrest of cellular metabolism with sodium azide freezes the DNA changes during switching. Nuclei isolated from azide treated cells arrested at the beginning or end of galactose induction of HO expression (-30 and 0) or 30 minutes into switching (30) contained the identical DNA fragments as DNA isolated by a rapid, hot phenol/glass bead lysis procedure (Figure 2-6C).

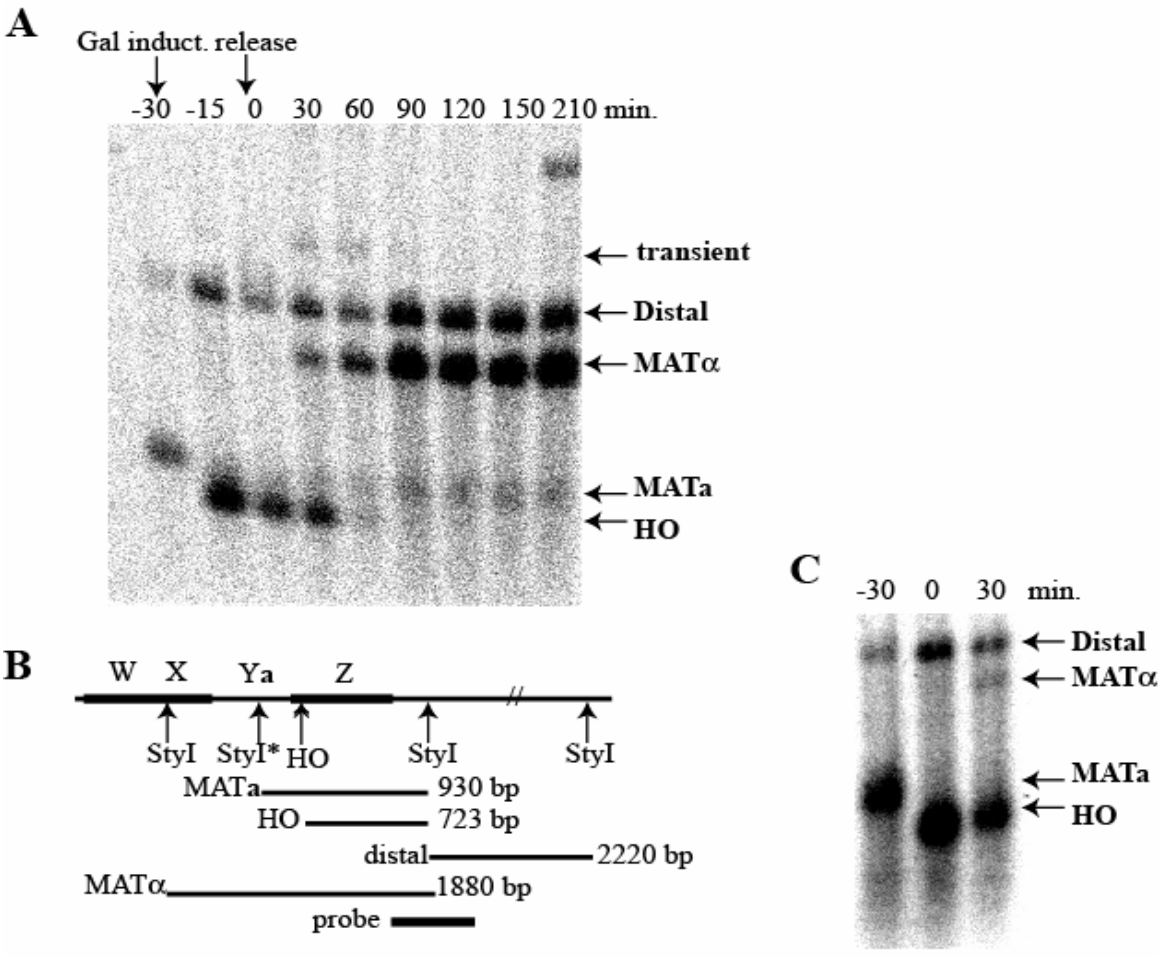

Figure 2-6: Synchronized switching of mating type

A: Time course of induction of HO endonuclease and mating type interconversion. DNA cut with *StyI* was blotted with a probe whose location is indicated in panel B. As shown in B, this probe detects a cell-type independent distal 2220 bp fragment, an 1880 bp fragment from *MATα*, a 930 bp fragment from *MAT***a**, and a 723 bp fragment when *MAT* has been cut by the HO endonuclease. Yeast **a**-cells, arrested at G1 by incubation with α factor, are induced to express HO in galactose (Gal induct; -30) and simultaneously released from the cell cycle block and transferred to dextrose medium to initiate switching (release; 0 min). Additional samples are taken over the next 3.5 hours, as indicated. C: Mating type assay in cells after nuclei purification. NaAz was added to cells to stop switching during nuclei preparation.

I evaluated chromatin structure throughout the ~45 kbp region shown in Figure 2-2 during synchronized switching of **a**-cells. Aliquots of cultures were taken after α-factor

arrest, at the end of galactose induction of HO endonuclease expression, and at intervals during the first 60 minutes of switching. Sodium azide was added to arrest metabolism, nuclei were prepared and digested with various amounts of DNase I. Purified DNA was then cut to completion with *Bam*HI and analyzed by indirect end labeling.

The region from 9 to 16 kmu contains *HML*α, the donor for recombinational repair of the double strand break at *MAT***a**, and is therefore a strong candidate for showing changes in chromatin structure during switching. In contrast to this expectation, there were no gross changes in DNase I digestion patterns in this region of chromosome III during the switching event (Figure 2-7A). The minor DNase I sensitive site at the E silencer, the major site at the I silencer, and the sites at the intergenic promoter and 3' end of the α1 gene were unaltered from arrested cells to near completion of the mating type switch. However, an additional band (highlighted with a dot), indicative of a localized DNase I sensitive site was present at the border of α1 and YCL065 at the 10 and 20 minutes samples. This site coincides with the region where the strand invasion is thought to occur (Haber 1998). The nucleosome ladder pattern extending rightward from the I silencer appeared to be less well defined during switching, particularly at the 20-40 minute points (Figure 2-7A). The region downstream of *CHA1* appears to become more sensitive to DNase I from 30-40 minutes. This sensitivity may be caused by the replication origin function of the ARSs around that region, but this remains to be confirmed. During synchronized switching, DNase I susceptibility of the rest of the region shown in Figure 2-2 was essentially identical to that in randomly growing **a** cells.

Chromatin structure of the RE differed between **a** and α cells (Figure 2-2C). Additional differences from the resting state in **a** cells were found when RE chromatin

structure was examined during a synchronized switching event (Figure 2-7B). Chromatin structure of the RE in the **a** cell population arrested at the G1/S boundary by α factor was similar to that in randomly growing cells. Structure was unchanged after induction of HO endonuclease expression. In contrast to these similarities, a new DNase I sensitive site (highlighted with a dot on the right side of the mapping picture) appeared early in the switching process and continued throughout switching. The site was located just to the right of the α2#2 operator, downstream from the transcription start sites for one of the three non-coding RNAs transcribed from the recombination enhancer, identified by Szeto and Broach (Szeto et al. 1997). At the end of switching, the RE chromatin must be remodeled from the **a**-cell structure, characterized by hypersensitive sites, to the α-cell array of positioned nucleosomes. Even though the majority of cells have α information at *MAT* after 60 minutes of switching (Figure 2-6A), chromatin structure at the RE has features characteristic of the **a**-cell pattern, e.g. similar nuclease sensitivity of the two operators (Figure 2-7B).

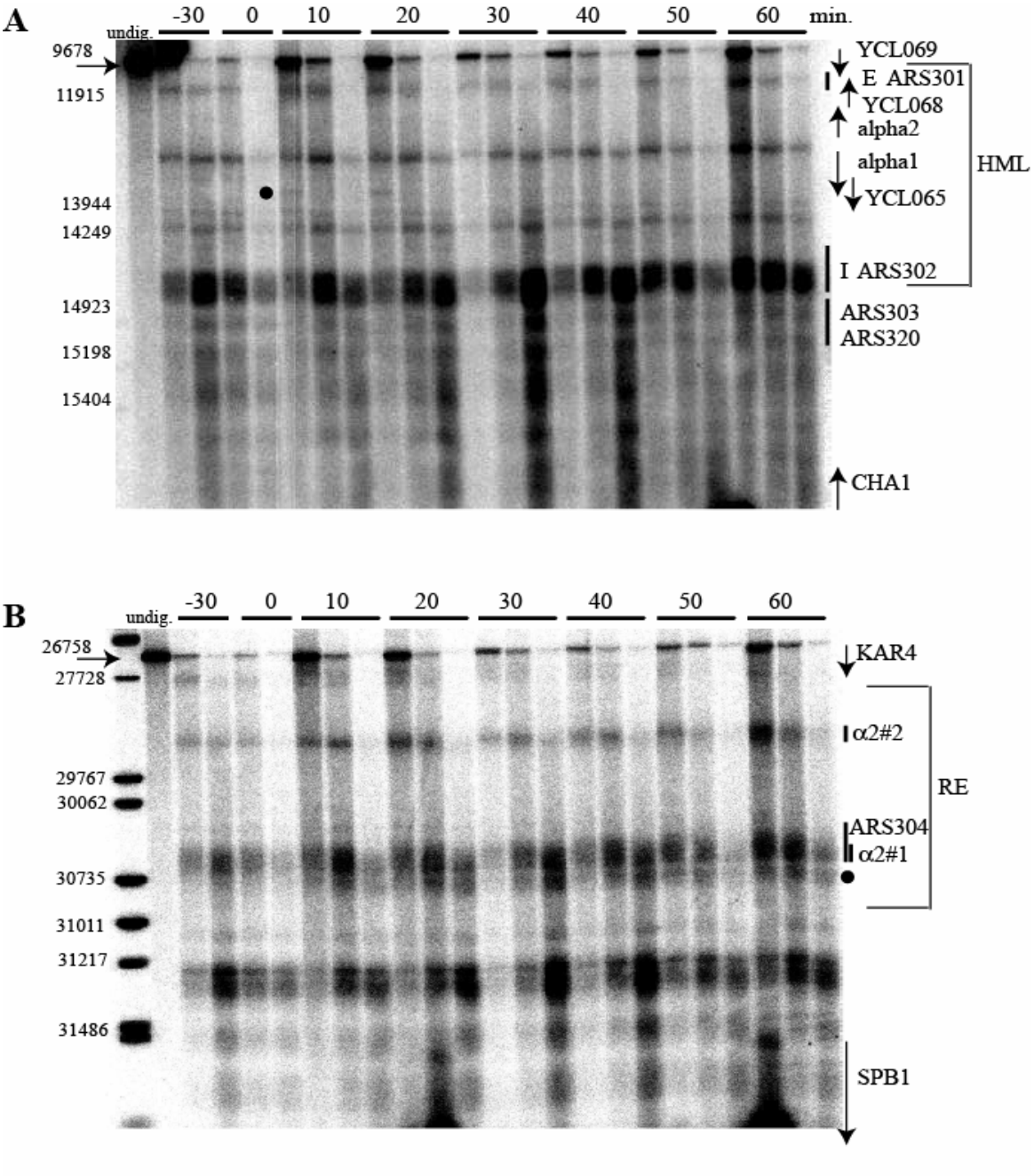

Figure 2-7: Chromatin structure during mating type interconversion

Nuclei from azide arrested samples taken during synchronized mating type switching were digested with DNase I and mapped by indirect end labeling. A: the *HML* region, comparable to Fig. 2A. B: the RE region.

2.5 Discussion

Increasingly, a role for chromatin structure is being implicated in studies of DNA function in transcription, replication, recombination and repair. Many of the studies leading to these conclusions are inferential, using genetic or other indirect approaches. Actual studies of chromatin structure have generally been confined to areas of particular interest to the investigator; promoters, enhancers, locus control regions, and such. In contrast, in this chapter, I present an unbiased examination of the DNase I susceptibility of an extensive region of yeast chromosome III. The region is of interest in its own right, since chromatin structure has been proposed as a contributory factor to directionality of mating type interconversion (Haber 1998). This inference derives from the low recombinational activity of the left most 100 kbp of this chromosome in α-cells and the oppositely high recombinational capacity of at least 40kbp of the same region in **a**-cells. A corollary to the hypothesis that chromatin structure is a factor in directionality of mating type interconversion is that there might be changes in chromatin structure during the switching event.

2.5.1 Chromatin structural features of the left arm of chromosome III

The mapped region includes thirty genes or ORFs, six potential replication origins, two known silencers at *HMLα*, and two Matα2p/Mcm1p operators at the recombination enhancer. Thus, the ~45 kbp chromatin segment includes most of the elements commonly found in the yeast genome and is a valid model for large scale characterization of chromatin.

As suspected from previous studies of specific genes, DNase I hypersensitive sites are present at the 5' flanking regions of ORFs. The surprising finding is that a hypersensitive site is present in the promoter region of nearly every gene in the region examined. The generality of occurrence of DNase I sensitive sites in 5' flanking regions favors a hypothesis that every ORF that is competent for transcription will have chromatin structure that leads to such a site. Occurrence of hypersensitive sites in the promoter region of nearly all genes, be the gene active at different levels or inactive, suggests that some basic chromatin structural feature exists at every promoter. Recent studies suggest that this basic structure might be the presence of less nucleosomes at the intergenic regions (Lieb et al. 2001; Lee et al. 2004). However, this is possibly not the only reason because I do observe differences in the DNase I cutting intensity between 5' and 3' end of the genes, although they are both intergenic regions. The 5' ends of the genes show stronger sensitivity to DNase I digestion than the 3' ends, suggesting that other factors contribute to the chromatin structure of the promoter regions. Some of the other factors that might contribute to creation of hypersensitive sites are binding of transcription factors, changes in DNA confirmation caused by protein binding, altered or remodeled nucleosomes at the promoter regions. The identification of the structural element of the 5' DNase I hypersensitivity, proteins involved in its creation, and how it relates to further aspects of transcriptional control will be important in understanding eukaryotic gene regulation.

While hypersensitive sites are present upstream of many of the genes in the ~45 kbp structure map, there is no striking correlation between intensities of cutting and transcription levels. Most yeast genes are transcriptionally competent; only a few are

truly silenced in the fashion of larger eukaryote heterochromatic genes. The examples of the yeast genes that are truly silenced are the silent mating type loci, **a**-cell specific genes in α cells, and the RE in α-cells. A clear example of DNase I sensitivity that correlates with active regions of chromatin and contrasts with chromatin that is silenced is found in the RE. The two Matα2p/Mcm1p binding sites are DNase I sensitive in **a** cells but less sensitive or not sensitive, for α2#1 or α2#2, respectively, as part of an organized chromatin domain in α-cells. The correlation of chromatin structure with transcription of the silent RNAs from the RE is clear. In **a** cells the non-coding RNAs are transcribed downstream of the both operators but not in α cells (Szeto et al. 1997).

The relationship between the 5' hypersensitive sites with transcription was further investigated by inducing transcription from regions that are normally repressed. The mapping results from the *FLO8* gene's cryptic promoter failed to identify a 5' hypersensitive site upon activation of transcription. However, the nuclease sensitivity of the entire gene increased upon the inactivation of Spt6. This may mask the detection of a DNase I hypersensitive site at the 5' end of transcription. On the other hand, when the *HML* locus was de-repressed, the transcription of *YCL065* and α2 caused appearance of two hypersensitive sites, suggesting that transcription factor binding and/or transcription induce formation of DNase I hypersensitive sites abutting a gene.

In addition to transcriptional competence, other features of chromatin structure reflect DNA function – a prime example is replication origin activity. Replication origin ARS305 is hypersensitive and has chromatin structural features flanking the consensus sequence; this is known to be a functional origin in actively growing yeast (Huang and

Kowalski 1996). In contrast, ARS303 and ARS320 are not known to be active and do not show strong sensitivity to DNase I digestion (Vujcic et al. 1999). The remaining two ARS consensus sequences are hypersensitive but this may relate to their being part of the two silencers, E and I, at *HML*α (Dubey et al. 1991). The parallel between locus control regions in higher organisms and silencers is obvious. Silencers are examples of chromatin regions where *trans*-acting factors interact with *cis*-acting elements, creating either transcriptionally competent or repressed chromatin domains. The protein-binding sites, such as silencers, are often nuclease sensitive, reflecting interactions (or their dissociation during chromatin isolation) of the *trans*-acting factors. The resultant chromatin domain can be nuclease sensitive, when it includes active genes, or nuclease resistant, when its constituent genes are repressed. In contrast, regulation of individual genes usually involves chromatin structure that is specific for the gene.

The most surprising feature of this unbiased investigation of a large region of chromosome III in yeast is occurrence of a number of regions where DNase I susceptibility indicates regions containing positioned nucleosomes. Cutting of linker regions occurs somewhat more frequently than core particle sequences, leading to a fuzzy DNA ladder when DNase I is used to assess regions of positioned nucleosomes (Staynov and Proykova 1998). The contrast with a MNase digestion pattern is apparent in comparison of the data in this work (e.g., Figure 2-2C) with previous MNase maps (Weiss and Simpson 1997). Given this limitation in mapping chromatin structure by DNase I digestion, it is quite striking that at least 7 sites occur in the mapped region where data strongly suggest the existence of an extended region of positioned nucleosomes. The frequency of such domains of positioned nucleosomes (~25% of

ORFs) is much higher than anticipated. Possible mechanisms for generating such organized chromatin domains include sequence/structure specific interactions with histones, usually thought to result from nonisotropic flexibility or bending of DNA (Shrader and Crothers 1989), statistical positioning of nucleosomes resulting from exclusion of histone binding by sequence specific DNA binding protein(s) (Kornberg and Stryer 1988), active organization of repressed domains by a repressor and corepressor that interact with histones (Ravindra et al. 1999), or nonhistone protein binding to specific DNA sequences separated by a number of base pairs that is at or near an integral multiple of the nucleosome repeat length (Thoma 1986), ~160 bp for budding yeast. The distinctive observation about chromatin structure made in the current study, suggesting a much wider occurrence of positioned nucleosomes in organized chromatin domains than previously envisioned, reinforces the possible role of chromatin structure in regulation of expression of a wider cast of genes and demands more extensive high resolution analysis of chromatin structure in studies of transcriptional regulation. Of interest, one ORF with positioned nucleosomes that overlapped an oppositely transcribed gene and did not have a promoter nuclease sensitive site, *YCL046* (Figure 2-2F), was expressed at a nearly 10-fold higher level in the Young laboratory genome wide analysis of transcriptional frequency in yeast strains depleted of nucleosomes by removing expression of histone H4 (Wyrick et al. 1999). This suggests that this gene is repressed by the positioned nucleosomes over its promoter. The significance of this regulation will be clear when a function is assigned to this gene. None of the other genes identified as having positioned nucleosomes was dramatically up regulated by nucleosome depletion.

2.5.2 Chromatin structure and mating-type switching

When homothallic yeast divides, the mother cell expresses HO endonuclease. This enzyme cleaves genomic DNA at the *MAT* locus, forcing the cell to heal the double strand break by a recombinational event using *HML*α or *HMR***a** as the donor locus. The donor is selected based on the sequences at *MAT*; most of the **a** cells recombine with *HML*α, while nearly all α cells use *HMR***a** as donor (Strathern and Herskowitz 1979). Directionality results from sequences at the recombination enhancer (Haber 1998). When RE is activated in **a** cells, ~100 kbp of the left arm of chromosome III, in general, and *HML*, in particular, are highly available for recombination. In contrast, in α cells, the RE is unavailable for binding of presumed *trans*-acting factors and the left arm of chromosome III is recombinationally cold (Wu and Haber 1995). A repressive chromatin structure encompasses the RE in α cells (Weiss and Simpson 1997); propagation of this structure has been suggested to lead to recombinational restriction for ~40 kbp of the left arm of chromosome III. Although differences in chromatin structure have been invoked to explain the differences in recombinational capacity for this region of chromosome III in **a** and α cells, no experimental analysis of chromatin for the extensive regions proposed to differ in the two cell types has been available until the current study.

The surprising result of this experimental analysis is that there are no differences in the DNase I chromatin maps between **a**- and α-cell nuclei other than those at the RE. Particularly revealing are the closely similar rates of decrease of the parental restriction endonuclease DNA fragment during DNase I digestion of the two cell type nuclei, indicating similar overall chromatin structures over the extensive region analyzed

(Figure 2-2). Validity of this analysis is testified to by differences in parental fragment digestion at the RE, where chromatin from **a** cells is known to have nuclease accessibility in excess of that from yeast α-cell nuclei (Weiss and Simpson 1997). Since general DNase I sensitivity did mirror known chromatin structure differences at the RE, the absence of similar differential nuclease sensitivity for other regions strongly supports the argument that there are no global, pervasive differences in chromatin structure of **a** and α cells for the leftmost 55 kb of chromosome III. Thus, if general chromatin structure plays a role in the differences in recombination capacity for the left arm of chromosome III in the two cell types, the structural elements responsible either were not preserved in isolated nuclei or were not detected by DNase I digestion. While a recently developed a DNase I *in vivo* expression method can be used for chromatin structure studies (Wang and Simpson 2001), the level of expression of the nuclease is suitable for analysis of single strand nicking by primer extension only. Higher activities will be necessary to apply *in vivo* DNase I expression to double strand cuts studied by indirect end label analysis. Alternatively, recombinational frequency differences may result from other, higher order chromatin structure or unknown features of nuclear organization in the two cell types.

Even more surprising is the near lack of detectable changes in the primary chromatin structure during the switching event. Clearly, the bulk of the population of **a** cells undergoes recombinational repair using donor sequences at *HMLα.* This major alteration in DNA sequence, including recombination and replication, occurs in the near absence of detectable changes in chromatin structure as mapped by DNase I digestion

and indirect end label analysis of *HML*α. However, a small DNase I sensitive site appears within 10 minutes of switching at a region where strand invasion is thought to start Figure 2-7A. If this site in fact reflects invasion, the changes in chromatin structure of the donor region during recombination is local and rapid. Moreover, strand invasion for most of the cells may start 10 to 20 minutes after start of the recombination, suggesting that the finding of the right donor happens early during switching. Since the chromatin maps presented in this study deal with structural analysis of a population of cells, it is conceivable that any individual cell goes through the switching recombination in a very short time, making detection of its individual chromatin structure changes impossible in the face of most molecules being pre- or post-switching. I also detect chromatin structure changes at the recombination enhancer that are correlated with the switching process. Since I can monitor chromatin changes at the RE and *HML* during switching, the chromatin structure alterations at the other loci must be localized to a small region and rapidly progressing from identical pre- to post-switching structures and therefore almost undetectable in my analyses of chromatin structure at selected time points. Alternatively, the recombination process may occur with the participants sequestered in a region of the nucleus inaccessible to DNase I and thereby be experimentally transparent in terms of chromatin structure changes.

The one region that does have significant alterations in chromatin structure during the switching process is the RE, the same region which exhibits distinctive chromatin structures in nuclei from the two yeast mating types. The two operators where Mcm1p can bind are hypersensitive to DNase I in **a** cells. During the time course of synchronized mating type interconversion, a region of chromatin just proximal to the proximal operator

becomes hypersensitive to DNase I. Nuclease sensitivity appears and persists over a time course consistent with the timing of mating type switching in these **a**-cells. This observation is consistent with transient binding of a factor(s) that activates the recombination enhancer for its role in the mating type switching process. Some of the candidates were shown to directly bind to the RE (Sun et al. 2002). Whether these factors contribute to transcription of the sterile RNAs and whether these non-coding transcripts (Szeto et al. 1997) play a role in activation of an extensive region of the left arm of chromosome III for recombination remains to be determined. I have evidence suggesting that transcription activation is critical for switching directionality (Chapter 3), in agreement with a prescient suggestion from James Broach (Szeto et al. 1997)

Chapter 3

Directionality of Mating-Type Switching and Transcription from the Recombination Enhancer Loci

3.1 Abstract

Saccharomyces cerevisiae mating-type switching is a gene conversion event which exhibits donor preference. *MAT***a** cells choose *HMLα* for recombination and *MAT*α cells choose *HMR***a**. Donor preference is controlled by the recombination enhancer (RE), located between *HMLα* and *MAT***a** on the left arm of chromosome III. A number of **a**-specific non-coding RNAs are transcribed from the RE locus. This transcription is cell cycle regulated. Additionally, proper mating-type switching and donor preference require a cell-cycle regulated factor. Mcm1 and Fkh1 activate RE in **a** cells. The results presented in this chapter suggest that Mcm1 binding is required for both RE transcription and Fkh1 binding. This requirement can be bypassed by inserting another promoter into the RE. The insertion of this promoter increases donor preference and opens the chromatin structure around the conserved domains of RE. Additionally, the level of Fkh1 binding positively correlates with the level of donor preference. I conclude that the role of Mcm1 in RE is to facilitate chromatin opening and transcription activation. This facilitates Fkh1 binding and the level of this binding determines the level of donor preference.

3.2 Introduction

Saccharomyces cerevisiae has two haploid mating types, **a** and α. Mating type of a haploid cell is determined by regulatory proteins that are encoded from the mating-type locus (*MAT*). *MAT* locus is located near the centromere of chromosome III. The same chromosome also harbors two transcriptionally silent and heterochromatic mating type cassettes, *HML* and *HMR*, which are located at the far left and right arm, respectively (schematically represented in Figure 1-4). In most *S. cerevisiae* strains, these two cassettes contain the sequences for the α and **a** alleles, thus often referred to as *HML*α and *HMR***a** (reviewed in (Rine and Herskowitz 1987; Haber 1998)). Yeast cells of one mating-type can switch to the other, as often as every generation. Mating-type switching is a gene conversion event between the *MAT* locus and one of the two mating type cassettes. Gene conversion starts with the introduction of a double strand break at the *MAT* locus by HO endonuclease. This double strand break is repaired by replacing the allele present at the *MAT* locus with a copy of the allele taken from one of the two donor loci, *HML*α or *HMR***a** (reviewed in (Haber 1998)). Mating-type switching is controlled by a number of different mechanisms. First, only mother cells can switch their mating types (Strathern and Herskowitz 1979). Second, switching starts at the G1 phase of the cell cycle (Breeden and Nasmyth 1987). Third, more than 85% of the switching attempts result in change from one mating-type to the other (Klar et al. 1982).

The first control mechanism on mating-type switching requires only the mother cell to start switching. A haploid cell divides to produce a bigger and a smaller cell, called “mother” and “daughter”, respectively. After division, HO endonuclease

expression is restricted to the mother cells, hence, only mother cells initiate switching (Strathern et al. 1982). The second mechanism involves the regulation of *HO* expression during cell cycle. *HO* is expressed during G1 phase of the cell cycle and switching is completed before the end of the S phase (Breeden and Nasmyth 1987).

The third mechanism that controls mating-type switching is based on a directional recombination system. Thus, **a** cells use *HML*α and α cells use *HMR***a** as the donor of gene conversion more than 85% of the time (Klar et al. 1982). This is called "donor preference". Donor preference does not depend on the sequence of the recombining cassettes, rather it depends on the location of the cassettes on the chromosome (Weiler and Broach 1992). A small cis-acting sequence, the recombination enhancer (RE) controls donor preference (Wu and Haber 1996). RE is located within a ~2.5 kb intergenic region ~29 kb from the left arm of chromosome III, between *HML* and *MAT*. The location of RE is important for correct donor preference. Deletion of the entire RE causes a dramatic change in the donor preference of **a** cells, where *HML* preference drops from ~85% to ~15% (Wu and Haber 1996). RE includes two α2/Mcm1 operators. A minimal RE of ~750 bp, is located around the first operator and is composed of a number of conserved sequence elements, named A,B,C, D and E and 300 bp downstream of the E region (Wu et al. 1998).

The mechanism by which RE governs directional recombination is still unknown. One of the most remarkable features of RE is that it activates the whole left arm of chromosome III for homologous recombination in **a** cells (Wu and Haber 1995). The same region is repressed for recombination in α cells. It has been shown that the differences in the primary chromatin structure of this region do not cause the differences

in recombinational frequency between the two cell types (Ercan and Simpson 2004). One mechanism by which RE may function is by governing the localization of the mating-type cassettes in the nucleus. The mobility of the left arm of chromosome III differs in **a** and α cells (Bressan et al. 2004). It was also demonstrated that the proximity of the recombining cassettes may not play a significant role in committing to recombination (Kostriken and Wedeen 2001; Simon et al. 2002).

While the mechanism of RE action is not clear, it is known that RE is activated in **a** cells and repressed in α cells. In **a** cells, Mcm1 binds and activates RE, and in α cells, Mcm1p binds along with α2p and represses its activity (Wu et al. 1998). The chromatin structure of RE also differs according to the cell type (Weiss and Simpson 1997). In **a** cells, RE chromatin has an open structure with two protein footprints bordering an unusual nuclease hypersensitive site. Recently, it was shown that in **a** cells, several transcription factors, Fkh1, Fkh2 and Ndd1 bind to RE (Sun et al. 2002). How these proteins affect RE function is not known. Nevertheless, among these, the deletion of Fkh1 protein reduces donor preference of **a** cells for *HML* from ~85% to ~35%. Interestingly, a number of non-coding RNAs, sizes ranging from 0.3-1.3 kb are transcribed from the RE locus in **a** cells (Szeto et al. 1997). The three major RNA species that are apparent in our northern blot analysis are named R1L&S and R2 and illustrated schematically in Figure 1-5. In α cells, there is no transcription. It is not known whether transcription from RE affects its activity. In addition, the relationship between transcription and the binding of other proteins to RE is unclear.

In this chapter, I present data suggesting that a cell cycle regulated factor is required for proper mating-type switching and donor preference. The transcription of the

non-coding RNAs is also cell cycle regulated. Additionally, the results presented in this chapter suggest that the RE is activated by a promoter, providing a transcription associated function that enhances Fkh1 binding in **a** cells. Mcm1 binding, which was shown to be required for RE function, is also required for the transcription of the **a**-specific non-coding RNAs. This requirement can be bypassed by inserting another promoter into the RE. The promoter insertion opens the chromatin structure around the conserved domains of RE and increases Fkh1 binding. These results suggest that the role of Mcm1 at RE is to activate transcription and facilitate Fkh1 binding. Moreover, the level of Fkh1p binding positively correlates with the level of donor preference.

3.3 Materials and Methods

3.3.1 Strains and cell growth

Standard yeast media, YP (including yeast extract and peptone), with an appropriate carbon source (2% dextrose, 2.5% lactose, 2% Raffinose) were used. For cell cycle regulation of RE transcription, and mating-type switching, JKM161**a** (*ho HML*α *Mat***a** *HMR***a** *ade1-112 lys5 leu2-3 ura3-52 trp::hisG, ade 3:: GalHO),* provided by J. Haber, was used. Cells were grown to OD_{600}=0.2 in 900 ml YP-Lactose at 30°C. They were synchronized by adding alpha factor at a concentration of 10 μg/ml. When >90% of the cells were unbudded, the culture was divided into two and 50 ml of 20% (w/v) galactose was added to half of the cells for HO induction. After 45 minutes of induction, cells in both switching and cycling cultures were filtered, washed and transferred in

prewarmed YP-Dextrose media. 20 ml of cells were collected at 10 minute intervals for RNA preparation. For the cell-cycle regulation of mating- type switching, in JKM161 cells, HO was induced for 45 min, then culture was divided into two. Half of the culture was transferred into media containing alpha factor, and the other half was released from arrest.

For mating-type switching and northern blot analysis, derivatives of CW157 (*HMLα, MAT***a**, *HMRα+BamHI, ade1, ura3-52, leu2-3,112, ade 3:: GalHO* (Sandell and Zakian 1993), provided by J. Haber, were used (Wu et al. 1998). These strains were generated as explained in the next section. *FKH1* was deleted from the genome by one step PCR replacement using the *TRP1* marker. For switching, cells were grown in YP-Lactose (or Raffinose), then Galactose was added to the media for 45 minutes. After induction, cells were collected by filtering and transferred into prewarmed YP-Dextrose media containing 0.2 mM $CuSO_4$. The switched cells were collected by harvesting 3-4 hours after the transfer. For preparation of RNA from synchronized cells, samples were collected similar to the collection of the cycling cells, described above except 0.2 mM $CuSO_4$ was added to the media into which cells were transferred at 0 min. RNA preparation for asynchronously growing cells were done with samples that were grown in YP-Dextrose media up to OD_{600} ~0.4-0.6, then 0.2 mM $CuSO_4$ was added for ~45 minutes before harvesting.

For ChIP analysis, cells were grown to OD_{600} ~0.6-0.8 in 200 ml YP-Dextrose, then 0.2 mM $CuSO_4$ was added for ~45 minutes before formaldehyde crosslinking.

3.3.2 Cloning and insertions at the RE

Insertion of this promoter and other sequences to the genome was performed with pop in/out with use of YIp5 (Scherer and Davis 1979). 660 bp RE was amplified from FY**a** cells for wild type RE, and CW157 cells for RE containing Mcm1 binding site mutation, from coordinates: Chr III 28922-29580 and cloned into the BamHI site in YIp5 (YIp5-RE). 165 bp *CUP1* promoter fragment, amplified from pYEX 4T-2 (Clontech) or a control sequence, amplified from the first 165 bp of *STE6* gene, was inserted at the BstBI site of YIp5-RE. The *CUP1* promoter was inserted in both directions. The primers used for the *CUP1* promoter amplification are; CUP1F: GAAGCGCTTCGAAGAGCGATGCGTCT TTTCCGC and CUP1R: GAAGCCGTTCGAACGATGACTTCTATATG ATATTG. Two TATA boxes were replaced by using primers; Forward: CCTTGTCTTGTATCAATTGCA**CCCGGG** TATCTTCTTGTTAGTGCAA TATCAT Reverse: GGTGCAATTGATACAAGACAAGGAGTTATT TGCTTCTTCTC**GGGCCC**TGATTCTGACAATCCATATTGC. The bold letters indicate the bases that replaced the TATA boxes. Forward primer was used with CUP1R, and reverse primer was used with CUP1F to amplify two fragments from the *CUP1* promoter. These two fragments were cut with MfeI, religated and cloned into the BstBI site of YIp5-RE. The resulting plasmids, YIp5-RE+inserted fragment, were cut with SpeI, and transformed into CW157 cells. DNA fragments containing the *URA3* marker were popped out from the genome by homologous recombination by growing the transformants in 5-FOA plates. The presence of the insertions and desired mutations in the genome was confirmed by PCR and sequencing.

3.3.3 Analysis of donor preference

For mating-type switching analysis, derivatives of CW157 were used. The donor preference analysis was performed as described previously (Wu and Haber 1995, 1996). CW strains carry *HML*α and *HMR*α containing a single base pair mutation that creates a BamHI site (*HMR*α+BamHI). HO induced switching from *MAT***a** to *MAT*α or *MAT*α+BamHI was monitored by Southern Blot analysis (Wu and Haber 1995). DNA isolated from switched cells was digested with HindIII and BamHI. A specific probe (ChrIII 13044-13278) was used to give fragments that correspond to the two donor loci *HML*α and *HMR*α+BamHI (~5 kb), and the switch products, *MAT*α and *MAT*α+BamHI, 4.4 kb and 3.1 kb respectively. Donor preference was shown as percent *HML* use. The bands corresponding to each donor use were quantified by using ImageQuant v.5.

3.3.4 RNA analysis

Total RNA was prepared by glass bead cell disruption as described previously (Zhang and Reese 2004). For Northern blotting, 20 μg of each RNA sample was subjected to electrophoresis in a formaldehyde/agarose gel, followed by transfer to a nylon membrane, and hybridization to specific probes. Probes for Northern blot hybridization were prepared by amplification from genomic DNA by PCR. For R1 L&S probe is complementary to Chr III, 29712-30517, R2 to Chr III, 30817-31317. Gel purification of the probes was followed by random primer labeling (Stratagene). For each blot, EtBr staining of rRNA or an RNA polymerase III-transcribed transcript, scR1, was

used as loading control. scR1 probe was generated as was explained in Chapter 2. Blots were stripped for re-probing by boiling in 0.1% SDS.

Qiagen Oligotex kit was used to prepare Poly-A RNA from total RNA of FY**a** cells.

The primer extension on total yeast RNA was performed as described (http://www.fhcrc.org/labs/hahn/methods/mol_bio_meth/primer_ext.html). The primer, complementary to Chr III 29520-29555, was purified from polyacryamide gel before kinase labeling. 2.0 µl 5X annealing buffer (25 mM Tris pH 8.3, 375 mM KCl, 5 mM EDTA), 10 ng kinase labeled primer and total RNA was brought to a volume of 10 µl. The mixture was boiled for 1 minute and incubated at 42^{o}C for 45 minutes. The following mixture was added to the samples after incubation; 4 µl 5X synthesis buffer (250 mM Tris pH 8.3, 375 mM KCl, 22.5 mM $MgCl_2$, 75 mM DTT), 0.3 µl each dNTP (10 mM dATP, dCTP, dGTP, TTP),14.3 µl H_2O, 0.5 µl M-MLV Reverse Transcriptase (200 u/ul GIBCO/BRL). After incubation at 37^{o}C for 30 minutes, samples were ethanol precipitated and RNaseA treated. After precipitation, the pellets were dissolved in sequencing loading buffer and run on sequencing gels.

3.3.5 Chromatin immunoprecipitation

The C-terminally myc-tagged Fkh1 was generated by using a PCR directed method with a 13MYC-KanMx cassette (Longtine et al. 1998). Chromatin immunoprecipitations were performed as described previously (Hecht and Grunstein 1999; Sharma et al. 2003). Cells were grown to OD_{600} ~1 and crosslinked with 1%

formaldehyde for 15 minutes. After quenching with 150 mM Glycine for 5 minutes, cells were harvested and broken by glass beads for 45 minutes in FA buffer (50 mM HEPES/KOH pH 7.5, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate,1 mM PMSF, 2 ug/ml Leupeptin, 1 ug/ml Pepstatin) with 0.2% SDS. The chromatin was fragmented by sonication using Branson 450, Setting 4, duty cycle 90%, 12 pulses for 5 times. Samples were cooled on dry ice/ethanol bath between pulses. After clarifying the lysates once by centrifuging @ 12,000rpm for 30 sec, samples were sonicated for 10 more times. Samples were spinned 20 minutes at 14,000 rpm at 4°C. From these extracts Fkh1-myc was immunoprecipitated in FA buffer by using 1 µl of antibody against myc tag (Roche) and then collected by ProteinA-sepharose (Amersham Biosciences). Protein A beads were washed with 1 mls of each: 10 minutes with FA buffer, 5 minutes with FA buffer containing 1 M NaCl, transferred into fresh tubes and washed with FA buffer containing 0.5 M NaCl for 10 minutes, 5 minutes with TEL buffer (0.25 M LiCl; 1% NP-40; 1% sodium deoxycholate; 2 mM EDTA; 10 mM Tris-HCl, pH 8.0), and twice with TE pH 8.0. After eluting the DNA with 450 µl elution buffer (25 mM Tris, pH 7.5, 2 mM EDTA, 0.2M NaCl, 0.5% SDS) at 65^{o}C for 30 minutes, samples were treated with Proteinase K, crosslink was reversed overnight at 65 ^{o}C and DNA was prepared by phenol/chloroform extraction and ethanol precipitation in the presence of 50 µg/ml glycogen. DNA was analyzed by PCR.

3.3.6 Nuclear DNA preparation and RE maps

Nuclei were prepared as described in Chapter 2. Nuclear DNA was cut with MNaseI (Worthington) at concentrations of 2.5-10 u/ml, at 37^{o}C, for 5 min. DNA was purified, cut with ScaI and subjected to electrophoretic separation on 1.5 % agarose gels, transferred to Hybond-NX membrane (Amersham), cross linked with UV light and hybridized with a specific probe. The probe was prepared by PCR amplification of a ~250 region specific to the distal end of the parental fragment including RE. Undigested control in the maps serves as a measure of the specificity of the probe. Naked DNA controls were generated by digesting 100 ng PCR products (amplifying Chr III, 28125-29774) and 30 μg Calf Thymus DNA with 50 u/ml MNaseI for 5 minutes at 30^{o}C. DNA was recovered by Phenol/Chloroform extraction and ethanol precipitation. Samples equal to 1 ng of naked DNA were treated with the DNA purified from nuclei preparations.

3.4 Results

3.4.1 Cell cycle regulation of RE transcription and mating-type switching

A number of non-coding RNAs are transcribed from the RE locus only in **a** cells. However, the effect of this transcription on RE activation is not known. I set out to find how transcription affects RE activity and donor preference during mating-type switching. In order to understand the role of transcription in RE activity, I first wanted to monitor the transcription from RE during switching and the cell cycle. Since mating-type switching is a cell cycle regulated event, it was anticipated that the transcription from the

RE locus would also be cell cycle regulated. In a previous study, I had determined the kinetics of switching in my experimental system (Figure 2-6) (Ercan and Simpson 2004). Northern blot analysis was used to monitor transcription from RE during switching and cell cycle progression (Figure 3-1). Samples were taken at 10 minutes intervals after release from alpha factor arrest (0 min). The timing of transcription was estimated by comparing the transcription of RE RNAs to HTA1 transcript. *HTA1* encodes for a histone protein and its RNA can be detected mostly during the S phase (Hereford et al. 1981; Hereford et al. 1982). The transcription of the non-coding RNAs was cell cycle regulated and expression peaked around late G1 and early S phase (Cycling cells, Figure 3-1). Note that the cycling cells showed another peak of transcription corresponding to the second G1/S. On the other hand, the second peak of RE RNAs in the switching cells did not appear because most of the cells changed from **a** to α within 1-1.5 hours, consistent with the kinetics of switching. Additionally, the second peak of the HTA1 transcript broadened but was still appeared, indicating that these cells continued cycling. The results at this resolution suggest that the kinetics of transcription is similar between the R1L&S and R2. The analysis of transcription in cycling and switching cells suggests that the RE transcription happens in parallel to switching during cell cycle. Transcription starts at late G1 and continues through the S phase, coinciding with the timing of mating-type switching. However, at this resolution I cannot distinguish whether transcription precedes or is coincidental with switching for each cell.

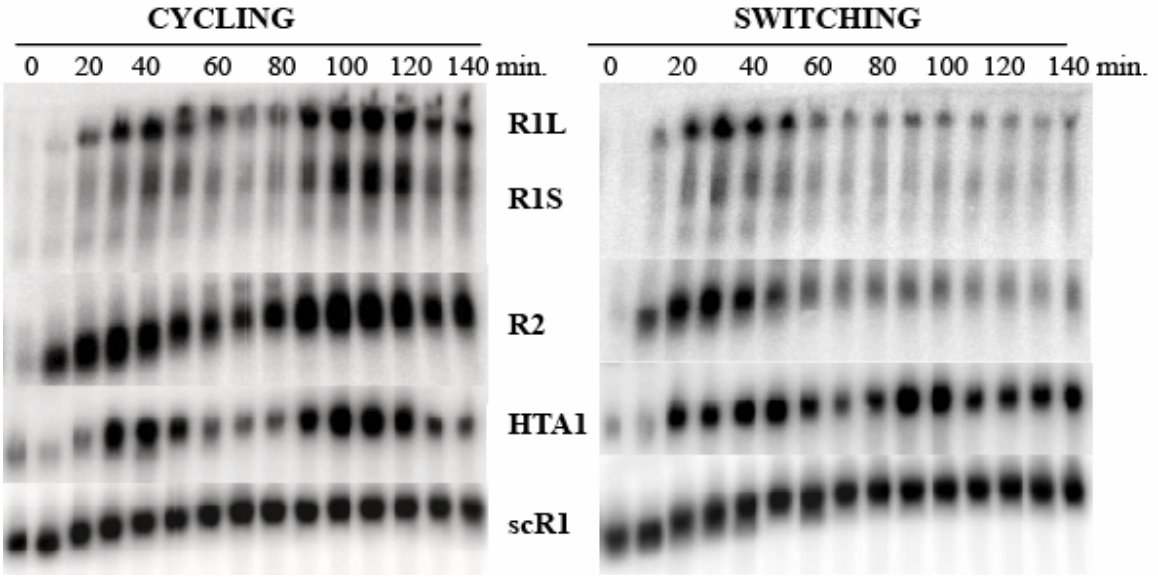

Figure 3-1: Transcription of the non-coding RNAs is cell cycle regulated

a cells were grown in lactose containing media and were arrested at G1 phase of the cell cycle with alpha factor. In the left panel (Cycling) cells were released from alpha factor arrest by filtration at 0 min. In the right panel (Switching) cells were induced for HO expression for 45 minutes with galactose and then released from arrest into fresh dextrose containing media at 0 min. RNA was prepared from aliquots of cells collected at 10 min intervals and R1L, R1S, R2 transcription was analyzed with northern blot. *HTA1* encodes for histone H2A, transcribed during S phase and *scR1* was used as loading control.

In order to understand the role of cell cycle in switching, I wanted to determine whether the cells would switch before the normal timing. For this, cells were arrested at the G1 phase of the cell cycle by alpha factor and HO cutting was induced for ~45 minutes. After the induction, half of the cells were released into fresh dextrose containing media (Lane 4, Figure 3-2A). The other half was transferred into fresh dextrose containing media with alpha factor, thereby forced to switch during G1 (Lane 3, Figure 3-2A). The switch products were analyzed by a previously developed Southern blot method (Connolly et al. 1988). Cells not only switched slowly during G1 but also exhibited less donor preference. After two hours, all the released cells completed switching with high

donor preference, ~80% *HML* use (Lane 4). In contrast, only slightly more than half of the arrested cells completed switching, and with a lower donor preference, ~ 53% *HML* use (Lane 3). This suggests that passage through late G1 and S phase is required for proper switching and correct donor preference. I tested whether the arrested cells died after two hours because they failed to repair the double strand break at the *MAT* locus. After 2 hours of switching at G1, the cells were grown in fresh dextrose containing media for 3 additional hours (Lane 5). The remaining cells also repaired their DNA, indicating that they were not dead.

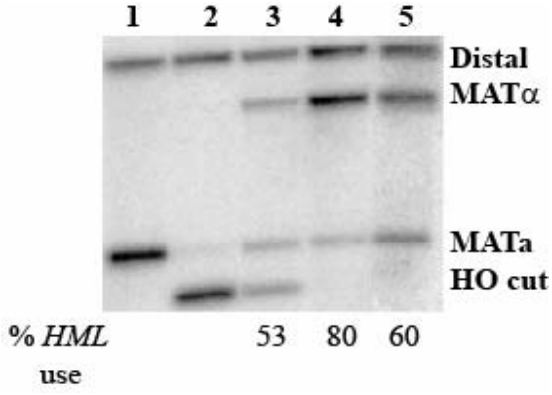

Figure 3-2: Cell cycle regulation of mating-type switching and donor preference

DNA cut with *StyI* was blotted with a probe specific to the *MAT* locus. This probe detects a cell-type independent distal 2220 bp fragment, an 1880 bp fragment from *MATα*, a 930 bp fragment from *MAT***a**, and a 723 bp fragment when *MAT* has been cut by the HO endonuclease. **a** cells, arrested at G1 by incubation with α factor (Lane 1), are induced to express HO in galactose (Lane 2). Half of the cells were transferred into dextrose medium with alpha factor for 2 hours (Lane 3), the other half was transferred into fresh dextrose containing media, thereby released from arrest (Lane 4). Some of the cells from Lane 3 were released from arrest and allowed to switch for additional 3 hours (Lane 5).Donor preference is shown as percent *HML* use at the bottom of each lane.

Since the non-coding RNAs are cell-cyle regulated, I wanted to test whether RE transcription was the cell-cycle regulated factor necessary for correct donor preference. For this, another promoter was inserted at the RE to activate transcription during G1.

However, this promoter failed to increase transcription, as well as the donor preference when cells were forced to switch during G1, shown in Appendix A Figure A-1.

In order to understand the nature of the non-coding RNAs, I determined that they are polyadenylated and transcribed by RNA Polymerase II (Figure 3-3A). This was done by isolating total RNA from **a** cells and separating polyadenylated RNA by binding them to an oligo-dT column. The fraction that contained the polyadenlylated RNA (Beads) and the rest of the RNAs (Sup) were analyzed by northern blot. *GYP1* messenger is polyadenylated, thus serves as a positive control, while scR1 serves as a negative control. RE RNAs bound to the oligo-dT column, indicating that they were polyadenylated. Previously, the Broach Lab reported that the transcription start site of the non-coding RNAs from the first Mcm1/α2 operator was ~ 250 bp downstream of the operator (Szeto et al. 1997). In order to narrow the region where transcription starts, I performed primer extension assay on total RNA, isolated from **a** cells. Different amounts of total RNA was titrated for optimal extension by the RNA polymerase. Transcription started ~250 bps downstream of Mcm1/α2 operator, within the E region (Figure 3-3B).

3.4.2 Mcm1 binding is required for RE transcription

Mcm1 is a constitutively expressed protein that regulates transcription of a number of genes including the **a**-specific genes. In α cells, α2 and Mcm1 bind to the α2/Mcm1 operator and repress transcription of the downstream gene by several mechanisms including organizing the chromatin structure into tightly positioned nucleosomes (Simpson 1990; Roth et al. 1992). In **a** cells, Mcm1 binds to the operator

and activates transcription. Previous studies showed that the α2/Mcm1 operator, located within the ~750 bp RE, is important for donor preference (Szeto et al. 1997; Wu et al. 1998). A mutation that prevents Mcm1 binding abolishes RE activity in **a** cells, where the donor preference reduces from ~85% to ~20 % *HML* use (Wu et al. 1998).

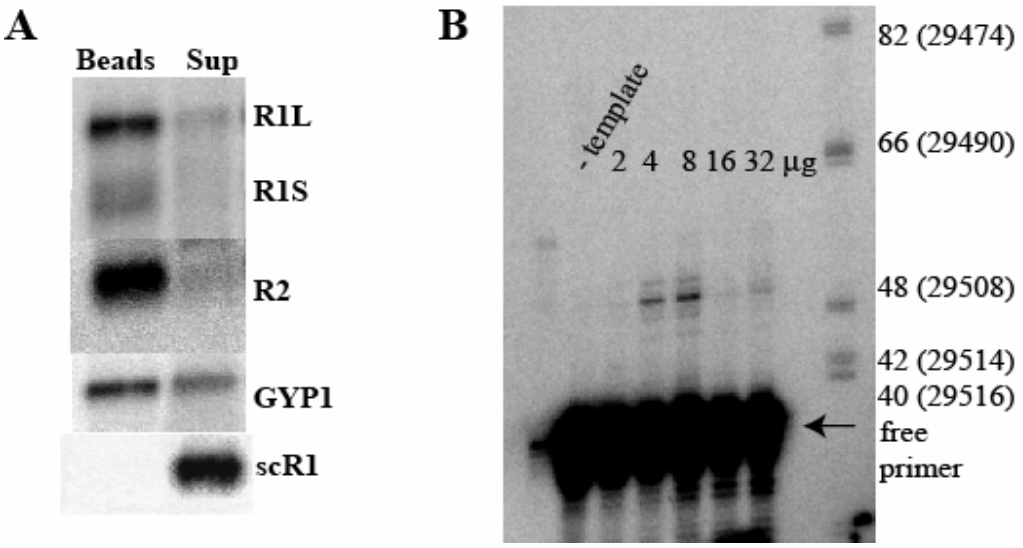

Figure 3-3: Non-coding RNAs from RE are polyadenylated and the transcription of one or more of the non-coding RNAs starts within the E region

A: RNA isolated from **a** cells was passed over an oligo dT column. The eluate and the supernatant were analyzed by northern blot. GYP1 was used as a positive and scR1 was used as a negative control for polyadenylation. B: Primer extension analysis was performed with total RNA of indicated amount. The chromosomal coordinates, corresponding to the marker is shown to the right of the gel. Free primer is shown with an arrow. E domain resides between coordinates 29436 and 29551.

Although RE contains two α2/Mcm1 operators, there are no ORFs in this locus. On the other hand, the **a**-specific non-coding RNAs are transcribed downstream of these operators (Szeto et al. 1997). Previous studies argued that these RNAs may not have a role in RE activity because the deletion of most of the transcribed sequences did not cause drastic changes in donor preference (Wu and Haber 1996). However, deletions towards the transcription start site (E region) gradually and significantly decrease donor

preference (Sun et al. 2002). Since the non-coding RNAs are **a-**specific, I reasoned that Mcm1 may activate transcription. I performed northern blot analysis to determine whether a mutation that prevents Mcm1 binding to RE would reduce transcription of the non-coding RNAs. RNA was prepared from **a** cells that were collected at 20 minutes intervals after release from alpha factor arrest. The non-coding RNAs were transcribed in cells with wild type RE (W) (Figure 3-4). In contrast, there was no transcription from the first operator in cells containing the Mcm1 binding site mutation at this operator (M), and the transcription from the second operator was also reduced. These results suggested that Mcm1 binding is required for RE transcription.

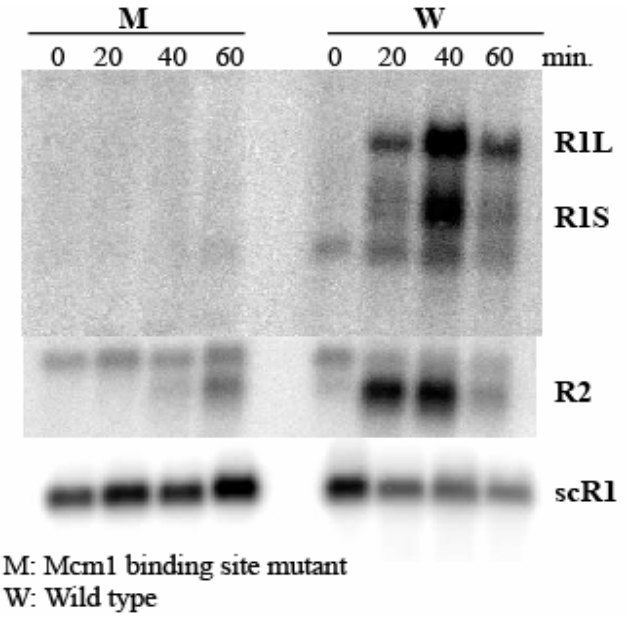

Figure 3-4: Mutation that prevents Mcm1 binding to RE also abolishes transcription of the non-coding RNAs

Cells containing a two base pair mutation at the first Mcm/α2 operator in RE (M), or wild type cells (W) were synchronized at G1 with alpha factor and released from arrest at 0 min. RNA was prepared from cells collected at 20 min intervals and R1L, R1S, R2 transcription was analyzed by northern blot. scR1 was used as loading control.

3.4.3 A promoter insertion into the RE can bypass the requirement for Mcm1 binding for correct donor preference of **a** *cells*

Since Mcm1 both activates transcription and RE function, I wanted to test whether Mcm1 protein itself or its transcription activator function is required. This was tested by inserting a well defined promoter (165 bp fragment that contains UAS2 and transcription start site of the *CUP1* gene) into the RE already containing the Mcm1 binding site mutation (MC). The insertion site and other features of RE are schematically represented in (Figure 3-5A). First, I wanted to determine how *CUP1* promoter affected RE transcription by northern blot analysis (Figure 3-5B). The transcription from the cells containing the wild type RE (W) and the RE with the Mcm1 binding site mutation (M) is also shown. The *CUP1* promoter activated transcription in MC cells (Figure 3-5B). Note that the MC cells exhibited longer RNAs than that of the wild type (indicated by arrows pointing to the RNA species appeared after the addition of *CUP1* promoter) suggesting that the transcription was governed by the *CUP1* promoter. In order to control for the effect of *CUP1* promoter insertion, this promoter was also inserted into the wild type RE (WC). To determine the donor preference in these strains, a previously developed method was employed (schematically described in Figure 3-5C, left panel) (Wu and Haber 1995, 1996). These strains contain *MAT***a,** *HML*α and a BamHI marked *HMR*α indicated as *HMR*α+BamHI. During switching, cells choose *HML*α or *HMR*α+BamHI and give rise to *MAT*α and *MAT*α+BamHI, respectively. These two alleles can be distinguished by Southern Blot analysis. Donor preference of **a** cells was determined by calculating %*HML* use, which can be used as a measure of RE activity. The insertion of *CUP1* promoter into the mutant RE increased the donor preference from ~25% to ~61%, M and

MC, respectively (Figure 3-5C). This result suggests that the requirement for Mcm1 binding can be bypassed by another promoter. The insertion of *CUP1* promoter into wild type RE did not affect the donor preference significantly (~80% *HML* use (WC), compared to ~86% in cells containing only wild type RE (W)).

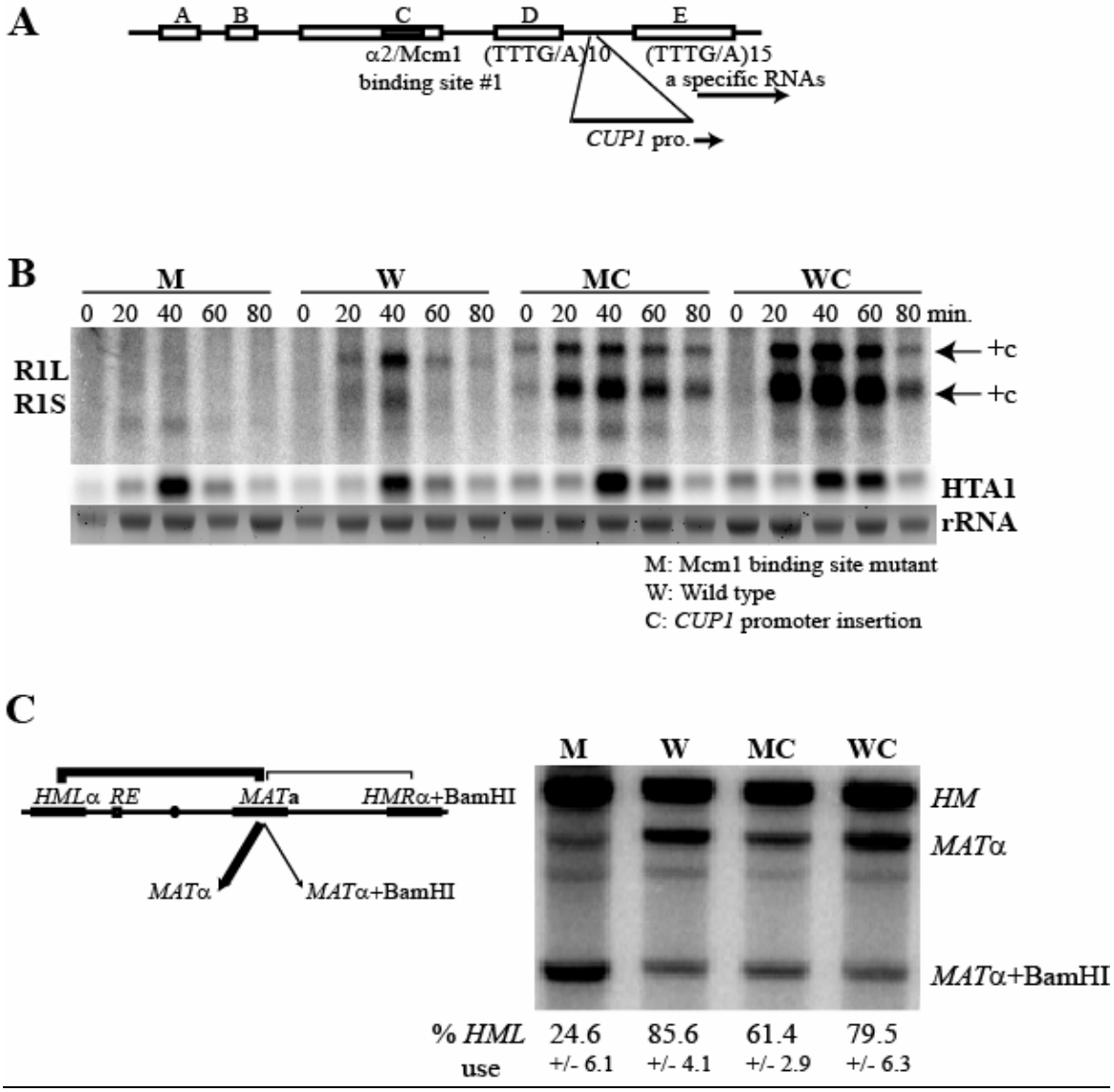

Figure 3-5: Transcription and donor preference of cells with/without the *CUP1* promoter insertion at RE

A: Schematic representation of *CUP1* promoter insertion B: Northern blot analysis of the transcription in cells containing wild type RE (W), RE with the Mcm1 binding site mutation (M), *CUP1* promoter insertion at the mutant RE (MC) and *CUP1* promoter insertion at the wild type RE (WC). Cells were grown in YP-Lactose and synchronized at G1 with alpha factor. Galactose was added to the media for 45 min to induce HO cutting. At 0 min, cells were released into prewarmed YP-Dextrose media containing 0.2 mM $CuSO_4$. Aliquots of cells were collected at 20 min. intervals and RNA was isolated. "+c" indicates the RNA observed after the *CUP1* promoter insertion. HTA1 transcript marks the S phase. rRNA was used as loading control. C: Donor preference analysis of the cells from (B). The analysis is schematically explained at the left panel. Strains that were used in this experiment carry a BamHI marked *HML* in place of the *HMR***a**. HO induced switching from *MAT***a** to *MAT*α or *MAT*α+BamHI was monitored by Southern Blot analysis. DNA isolated from cells that completed switching (3-4 hours after the start of switching) was digested with HindIII and BamHI. A specific probe was used to give fragments that correspond to the two donor loci *HML*α and *HMR*α+BamHI (~5 kb), and the switch products, *MAT*α and *MAT*α+BamHI, 4.4 kb and 3.1 kb respectively. Donor preference is shown as percent *HML* use at the bottom of each lane.

In order to test whether a control sequence that does not contain a promoter would also increase donor preference, a portion of the *STE6* open reading frame that is approximately the same size as the *CUP1* promoter was inserted into the RE containing the Mcm1 binding site mutation (MS). The control sequence failed to increase transcription; compare transcription of M and MS (Figure 3-6A) as well as donor preference, from 25% to 26% *HML* use in M and MS cells, respectively (Figure 3-6B), suggesting that a function of the promoter is required for the RE activity.

Reversing the orientation of the promoter reduced the donor preference only slightly. When *CUP1* promoter was inserted in reverse direction (MR), donor preference reduced to ~55% (Figure 3-6B). This result suggests that the direction of transcription and the sequences that are transcribed do not influence donor preference significantly. The transcription from these cells in both directions is shown in Figure 3-6C.

I also wanted to determine whether the level of transcription from the *CUP1* promoter was important for donor preference. Transcription was reduced by deleting the TATA boxes of the inserted promoter (MCΔT) (Figure 3-6A). Previous studies showed that the deletion of the TATA boxes at the *CUP1* promoter reduces transcription without preventing the activators function at this promoter (Shen et al. 2001). When the TATA boxes were deleted, donor preference was increased even more, from ~61% to ~76%, M and MCΔT cells, respectively (Figure 3-6B). Notably, these cells produced RNAs of normal size, suggesting that the promoter lacking the TATA boxes is activating transcription from the native start site. The increase in donor preference of MCΔT cells, 77% *HML* use, compared to 61% *HML* use (MC cells), may be a reflection of transcribing the normal non-coding RNAs. It may also reflect the possibility that when started from the *CUP1* promoter, transcription passes through the entire E domain, hence preventing other regulatory proteins from binding to this region. These results suggest that while the activity of the *CUP1* promoter is important for RE activation, the direction or the level of transcription may not affect donor preference. This also suggests that the transcription activators, not the transcription itself, are necessary for RE function.

In order to understand the mechanism of RE activation by the *CUP1* promoter insertion, I generated several strains and performed northern blot and donor preference analysis. The results of these studies are presented in Figure 3-7. Northern blot analysis was performed with RNA isolated from asynchronously growing cells in media lacking copper or containing 0.2 mM $CuSO_4$. Northern blot analysis of the transcription from the *CUP1* promoter in RE suggested that when inserted at the RE, the *CUP1* promoter is not

as sensitive to $CuSO_4$ (Figure 3-7A). When cells were grown in media lacking copper, or media containing 0.2 mM $CuSO_4$, RE transcription did not change drastically.

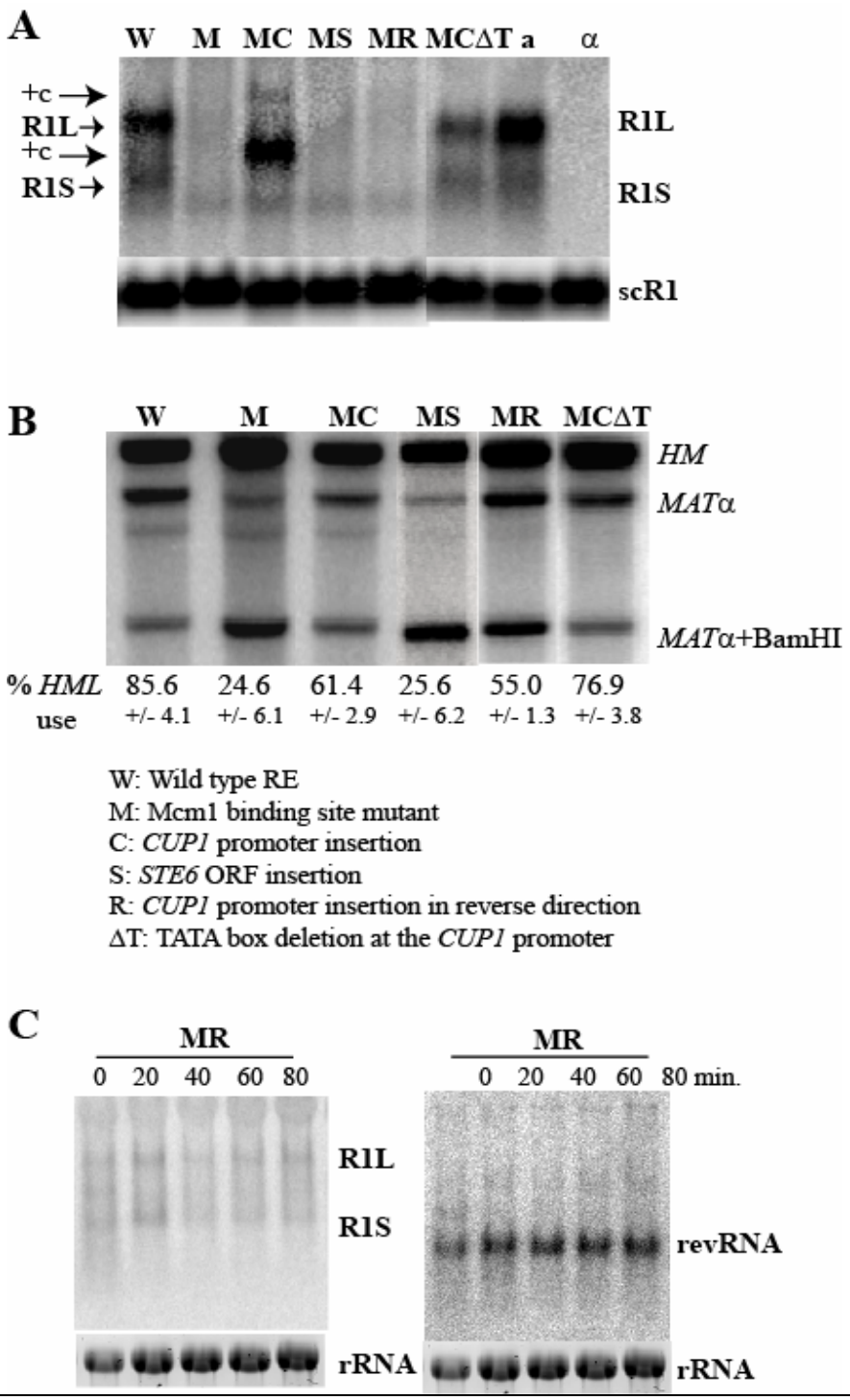

Figure 3-6: Northern blot and donor preference analysis of *CUP1* promoter insertion into RE

A: Northern blot analysis of transcription from **a** and α cells along with **a** cells containing wild type RE (W), RE with Mcm1p binding site mutation (M), CUP1 promoter (MC), control sequence (MS) and CUP1 promoter lacking the TATA boxes (MCΔT) inserted at the mutant RE. All cells were collected from asynchronously growing cultures, 0.2 mM CuSO4 was added to these cultures 45 minutes before harvesting. scR1 was used as loading control. B: Donor preference analysis of the cells from part A. Analysis was done as described in Figure 3-5C. Donor preference is shown as percent HML use at the bottom of each lane. C: Northern blot analysis of transcription downstream and upstream of reverse CUP1 insertion. Experiment was performed as described for Figure 3-4.

One exception is the cells containing the *CUP1* promoter lacking TATA boxes inserted into the mutant RE (MCΔT). In these cells, addition of copper to the media increased transcription, as well as donor preference, from ~59% to ~77% *HML* use. When the major transcription activator of the *CUP1* promoter, *ACE1* was deleted in the cells containing the *CUP1* promoter inserted into the mutant RE (MCΔA), RE transcription was reduced (Figure 3-7A). These cells also demonstrated lower donor preference of ~44% *HML* use (Figure 3-7B). The deletion of the TATA boxes in the *CUP1* promoter in these cells (MCΔTΔA) changed neither transcription, nor donor preference; compare 43.7% *HML* use in MCΔA cells to 43.6% in MCΔTΔA cells (Figure 3-7B). These results suggest that the *CUP1* promoter, when inserted in RE functions somewhat differently from *CUP1* transcription. This may be because the promoter sequence itself may disrupt the nucleosomes in that region of the RE. The control sequence that was used in my studies is a portion of an ORF, thus may not disrupt nucleosomes.

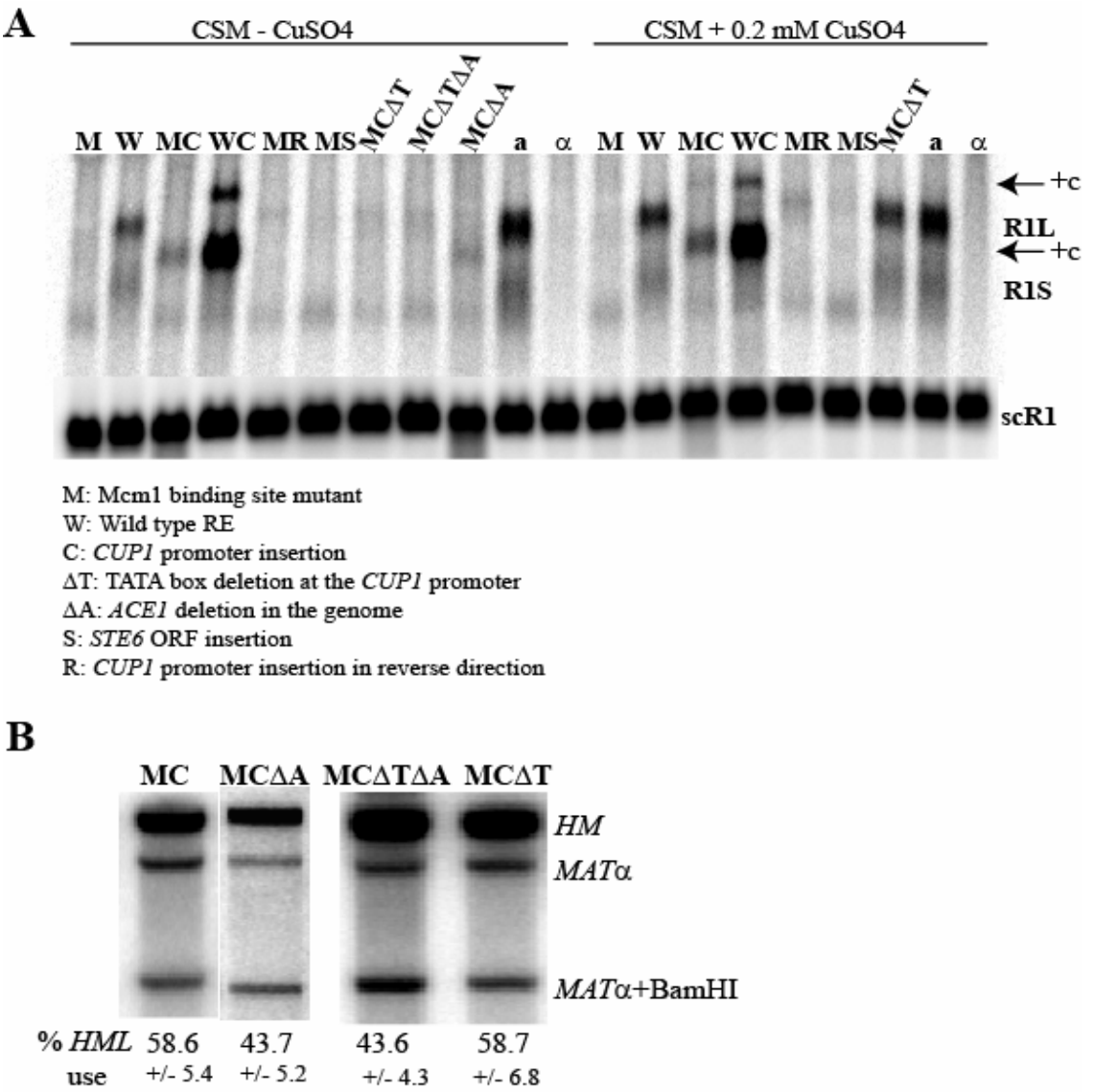

Figure 3-7: Analysis of *CUP1* promoter insertion into RE

A: Northern blot analysis of RE transcription in cells with genotypes indicated below. RNA was prepared from cells that were grown asynchronously in complete synthetic media lacking copper or containing 0.2 mM $CuSO_4$. B: Donor preference analysis of cells with the indicated genotype. All switching were performed in media lacking copper. Southern blot analysis was performed as described in Figure 3-5C. Donor preference is shown as percent *HML* use at the bottom of each lane.

3.4.4 CUP1 promoter insertion at the RE enhances Fkh1 binding

A number of transcription activators, Fkh1, Fkh2 and Ndd1, bind to RE in **a** cells. Among these, the deletion of Fkh1 protein significantly reduces donor preference of **a** cells, from ~85% to ~35% (Sun et al. 2002). Since it was known that Fkh1 is the transcriptional activator of several genes, I wanted to know whether Fkh1 activates RE transcription (Kumar et al. 2000; Zhu et al. 2000). For this, RE transcription was analyzed by northern blot in cells lacking Fkh1 (Δ*fkh1*) (Figure 3-8A). The northern blot indicated that the overall transcription from the RE region is not affected by the deletion of *FKH1*. Since the donor preference in cells lacking Fkh1 is low, while RE transcription occurs normally, transcription by itself is not sufficient for RE activity.

Figure 3-8A presents that the *FKH1* deletion does not affect RE transcription, suggesting that the function of Fkh1 in RE may be different from its role in transcription activation. I also wanted know whether the function of Fkh1 could be replaced by the *CUP1* promoter insertion. For this, the donor preference of **a** cells lacking Fkh1 with (CΔ*fkh1*) or without (Δ*fkh1*) a *CUP1* promoter insertion were determined. Transcription from these cells is shown in Figure 3-8A. The *CUP1* promoter insertion failed to increase the donor preference of cells lacking Fkh1, ~31% to ~33% *HML* use in Δ*fkh1* and CΔ*fkh1* cells, respectively (Figure 3-8B), suggesting that the *CUP1* promoter cannot replace the function of Fkh1 in donor preference.

Since the deletion of Fkh1 protein did not affect RE transcription, I hypothesized that the activation of RE in **a** cells consists of a sequence of events that starts with binding of Mcm1 in order to open chromatin and activate transcription, followed by Fkh1

binding. In order to test this hypothesis, I wanted to determine whether Mcm1 was required for Fkh1 binding and the *CUP1* promoter insertion would increase Fkh1 binding. For this, chromatin immunoprecipitation (ChIP) method was used.

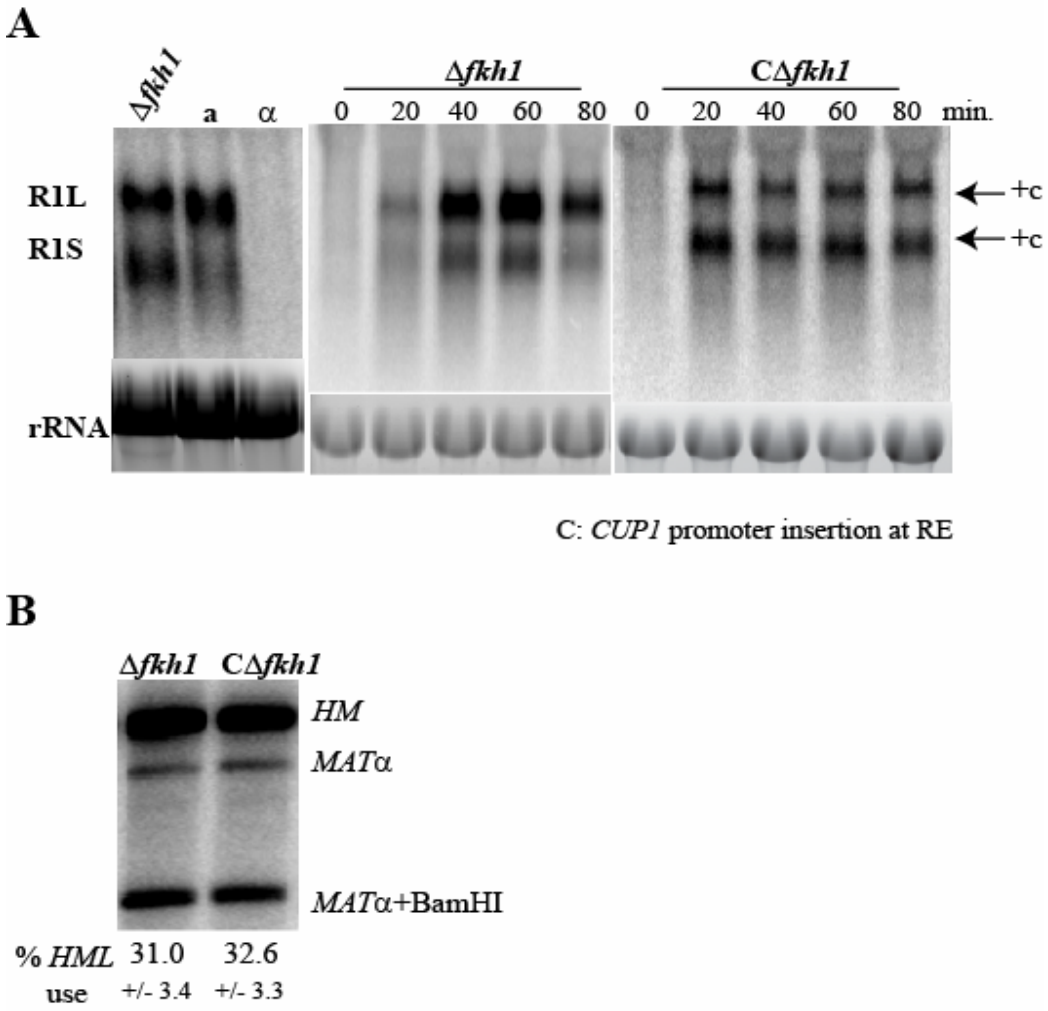

Figure 3-8: Northern blot analysis of transcription and donor preference of cells lacking Fkh1 protein

A: RNA from asynchronously growing Δ*fkh1* cells along with **a-** and α cell-RNA is shown on the left. Northern blot analysis was performed as described in Figure 3-6 for the asynchronously growing cells. rRNA was used as loading control. RNA from cells lacking the Fkh1 protein (Δ*fkh1*) and with the CUP1 promoter insertion (CΔ*fkh1*) was prepared from aliquots taken at 20 min. intervals after release from alpha factor arrest as explained in Figure 3-4. B: Donor preference of CΔ*fkh1* and Δ*fkh1* cells was determined as described in Figure 3-5C.

Fkh1 protein was precipitated by using an antibody against the myc tag that was attached to the C-terminus of the protein. Cells were grown asynchronously in dextrose containing media, and 0.2 mM $CuSO_4$ was added to the media for 45 minutes before crosslinking with formaldehyde. The level of RE DNA enrichment was determined by PCR analysis of the immunoprecipitations from the cells with wild type RE (W), Mcm1 binding site mutation (M), *CUP1* promoter (MC), or control sequence (MS) inserted into the mutant RE. The level of Fkh1 binding was expressed as the fold enrichment of RE DNA in tagged versus untagged Fkh1 containing cells. Fold enrichment of RE DNA is plotted as a graph at the right panel of Figure 3-9A. The mutation that prevents Mcm1 binding also reduced the level of Fkh1 binding from ~5 fold (W) to no enrichment (M), suggesting that Fkh1 requires Mcm1 binding at the RE. The insertion of the *CUP1* promoter at the mutant RE (MC) increased the level of Fkh1 binding from no enrichment (M) to 3 fold enrichment (MC). This result suggests that the *CUP1* promoter may activate RE in the absence of Mcm1, by enhancing Fkh1 binding. Moreover, the level of Fkh1 binding positively correlated with the level of donor preference.

Since the level of Fkh1 binding positively correlates with the level of donor preference, I wanted to know how this protein functions at the RE. Figure 3-8A presented that *FKH1* deletion does not affect RE transcription and a promoter insertion cannot bypass the requirement for Fkh1 for donor preference. Thus the role of Fkh1 in RE may be different from its role in transcription activation. In addition to its role in activation, Fkh1 was shown to be associated with transcription elongation (Morillon et al. 2003). Fkh1 could bind both to the promoter and towards the coding region of the *CLB2* gene. I tested whether Fkh1 binding was higher towards the middle of the transcribed region by

using ChIP assay. Figure 3-9B demonstrates that Fkh1 binding peaked around its binding sites, suggesting that the role of Fkh1 in RE may be different from its role in transcription elongation.

In order to understand how Mcm1 and the *CUP1* promoter enhances Fkh1 binding, I tested whether they were required for the open chromatin structure of RE in **a** cells. For this, the overall chromatin structure was mapped in α cells, **a** cells lacking Fkh1, **a** cells containing wild type RE (W) or RE with the Mcm1 binding site mutation (M) and with *CUP1* promoter insertion (MC) or control sequence insertion (MS). The map was performed at medium resolution with indirect end labeling by using MNase I (Figure 3-10). The chromatin structure of RE has been mapped with DNase I at medium resolution and with MNase I at higher resolution by primer extension in both cell types (Weiss and Simpson 1997; Ercan and Simpson 2004). In **a** cells, RE chromatin was open and sensitive to nuclease digestion. In α cells the nuclease digestion pattern suggested the presence of tightly positioned nucleosomes. Our MNase I mapping of the region comprising the conserved domains of RE at medium resolution was consistent with these studies. In **a** cells, the cutting pattern across the RE (Figure 3-10, Lanes 5&6) was similar to that of the naked DNA (Lane 7), indicating that this region was sensitive to nuclease digestion. In α cells, the sites that were apparent in the naked DNA digestion (Lane 7) were mostly protected and the spacing of the MNase I cutting sites suggested the presence of positioned nucleosomes (Lanes 9&10). *FKH1* deletion in **a** cells did not affect the overall chromatin structure. The cutting pattern in these cells (Lanes 2&3) was similar to that of the **a** cells (Lane 5&6) as well as the naked DNA (Lane 7). This result suggests that Fkh1 protein is not required for open chromatin within the RE region.

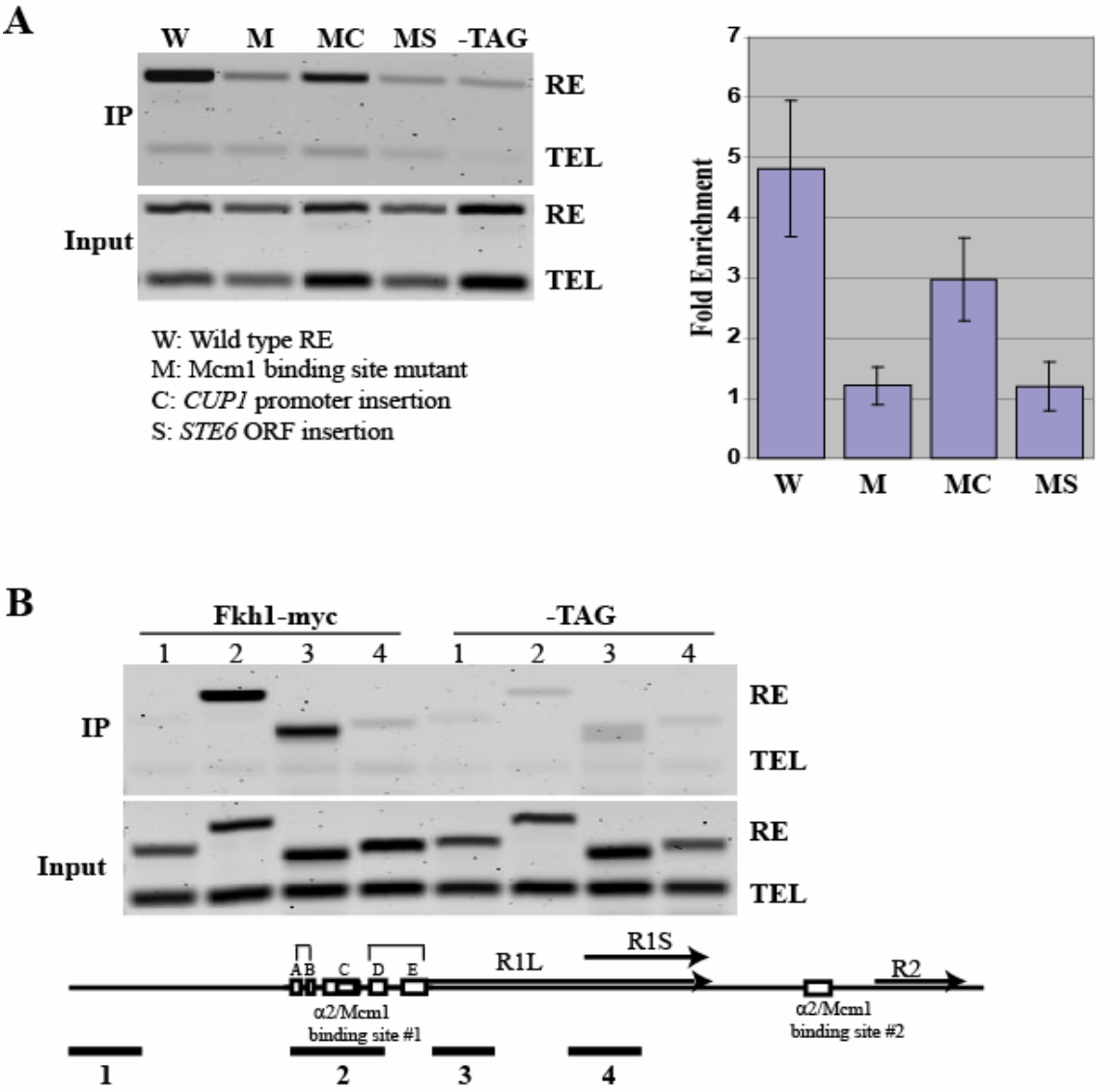

Figure 3-9: Analysis of Fkh1 binding to RE

A: ChIP analysis was performed in different strains. Primers amplified a portion of the RE domains comprising most of the Fkh1 binding sites. (B). Fold enrichment of RE DNA was calculated and plotted with standard deviations, right panel. B: Primers sets 1, 2, 3 and 4 were used to amplify the following regions of RE: 1: 28125-28400, 2: 29051-29363, 3: 29520-29742, 4: 30117-30355. The locations of the primers are shown schematically, below. Fkh1 binding sites occur between coordinates 29087 and 29548 in domains A, B, D and E, highlighted with brackets above. All samples were prepared from asynchronously growing cultures and 0.2 mM $CuSO_4$ was added to these cultures 45 minutes before crosslinking.

High resolution MNase I maps have demonstrated that Mcm1 protein was required for the open RE chromatin in **a** cells (Wu et al. 1998). We also observed that when Mcm1 binding was prevented (M), RE presented a more closed chromatin structure (Lanes 12&13) compared to that of the **a** cells (Lanes 5&6). The MNase I cutting pattern in these cells indicated the presence of nucleosomes. When *CUP1* promoter was inserted in cells containing the Mcm1 binding site mutation (MC), RE presented a more open chromatin structure (Lanes 16&17). A strong MNase I sensitive site (highlighted with a dot to the left of Lane 16) was present within the promoter, close to the UAS. Moreover, the cutting pattern at the region covering the A, B, C and D domains in MC cells (Lanes 16&17) was similar to that of the **a** cells (Lanes 5&6). Thus, the chromatin was more open compared to the cells containing the Mcm1 binding site mutation only (M, Lanes 12&13) or cells containing the control sequence insertion (MS, Lanes 20&21). This suggests that the insertion of *CUP1* promoter disrupts the repressive chromatin structure that is present at the RE when Mcm1 binding is prevented. Additionally, Mcm1 may enhance Fkh1 binding by opening the chromatin around the conserved domains, where the binding sites are located. This was supported by the data showing that in the absence of Mcm1, *CUP1* promoter opens the chromatin structure and also enhances Fkh1 binding. Since Fkh1 is not required for open chromatin, I propose that this protein functions downstream of the chromatin remodeling that is recruited to RE by Mcm1. A second map, focusing on the region downstream of the *CUP1* promoter insertion site was generated, shown in Appendix 1 Figure A-2.

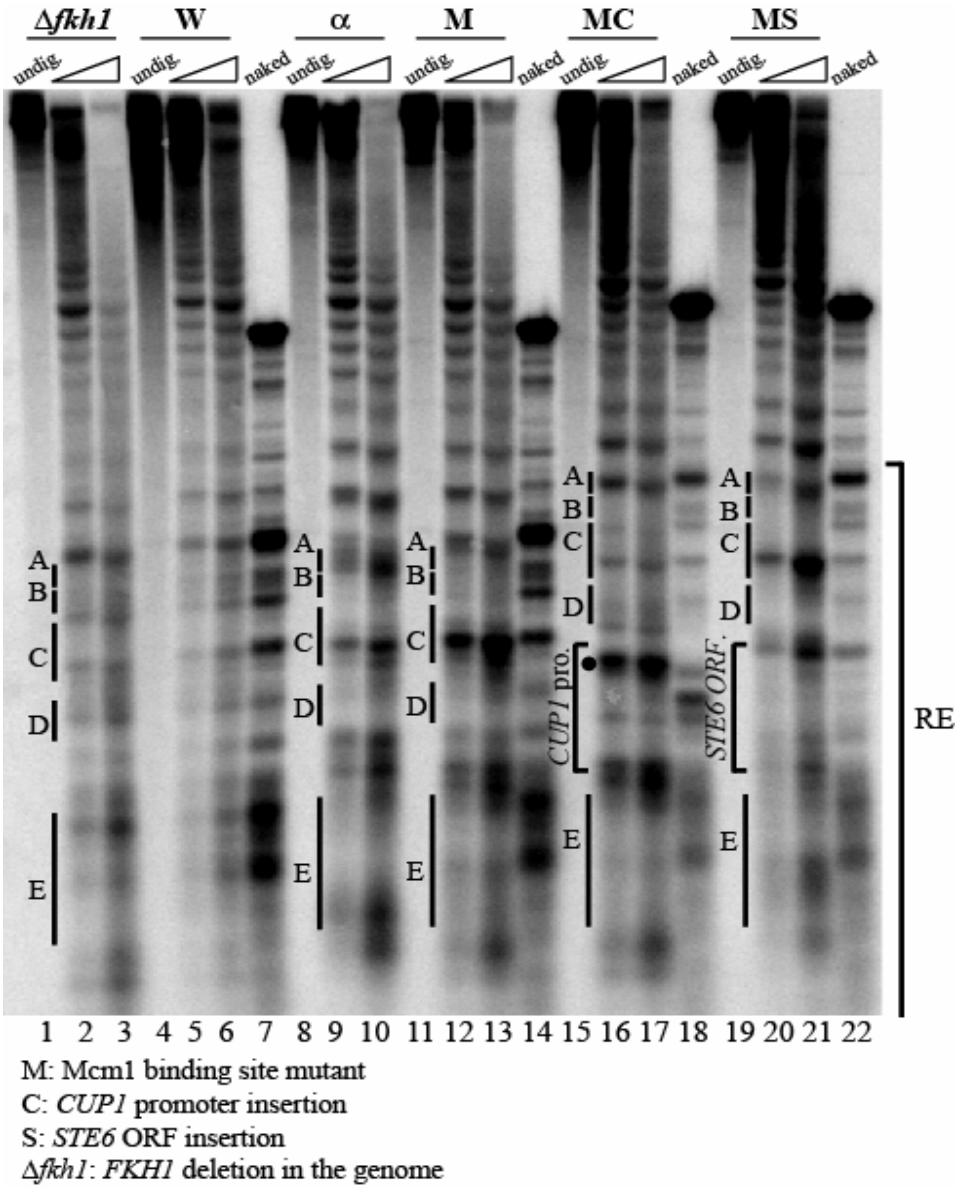

Figure 3-10: Chromatin structure around RE region mapped by MNase I

The chromatin structure around the conserved domains of RE was mapped by indirect end labeling. An undigested sample with two levels of MNase I digestion is shown for every sample. Naked DNA was used to monitor the sequence specificity of the enzyme. The nomenclature of the samples is explained below. The domains of RE are shown to the left of each sample and the insertions are indicated with brackets.

3.5 Discussion

Over the past 25 years, many details of yeast mating-type switching have been revealed. Although mating-type switching is commonly used as a model system to study homologous recombination in yeast, one aspect of switching remains to be a puzzle. This aspect is the directionality of recombination. During mating type switching, **a** cells choose *HML*α and α cells choose *HMR***a** for gene conversion. This process is called "donor preference" and is controlled by a small cis-acting element, RE. A number of **a**-cell specific non-coding RNAs are transcribed from the RE region (Szeto et al. 1997). However, the role of transcription from RE is not clear. In this chapter, I studied the properties of this transcription and its relation to Mcm1 protein and the binding of Fkh1.

3.5.1 The role of cell cycle in RE transcription and switching

Yeast mating-type switching is cell cycle regulated. *HO* gene is induced at the G1 phase. Consequently, switching starts during late G1 and is completed before the end of the S phase (Breeden and Nasmyth 1987). The transcription of the non-coding RNAs from RE is also cell cycle regulated. The significance of this regulation is not clear. Transcription starts at late G1 and continues through the S phase. Moreover, normal progression through G1/S is required for both proper switching and donor preference, suggesting that a factor(s) that is present at the late G1 and S phase is necessary. This factor may be a recombination protein that is required for gene conversion. This was supported by a recent study suggesting that the recombination machinery that is required for proper gene conversion is affected by the cell cycle (Wang et al. 2004). However,

whether these recombination proteins also affect the donor preference, is not known. Perhaps, the transcription from RE itself may be the cell cycle regulated factor that is required for correct donor preference. Further experiments are required to test this possibility.

3.5.2 The role of Mcm1 in RE activation

α2/Mcm1 operators are present at the promoter regions of **a**-specific genes. Although there are no genes within the RE locus, the **a**-specific non-coding RNAs are transcribed downstream of the two operators (Szeto et al. 1997). The results in this chapter showed that when Mcm1 binding site is mutated, the transcription of the non-coding RNAs is abolished. The same mutation also reduces donor preference (Wu et al. 1998). Additionally, a temperature sensitive Mcm1 affects donor preference. These results suggest that Mcm1 protein is required for donor preference and RE transcription. Moreover, in **a** cells, the open chromatin structure of RE requires Mcm1 binding (Wu et al. 1998). It was not clear whether the Mcm1 protein itself or its requirement for open chromatin was necessary for RE function. My results suggest that since the requirement for Mcm1 binding can be bypassed by inserting another promoter into the RE, Mcm1 activates RE by facilitating the formation of open chromatin structure and may not play a role in the downstream functions of RE. This is supported by data showing that the promoter insertion that bypasses the Mcm1 requirement, also induces more open chromatin (Figure 3-10).

I wanted to dissect the role of RNA synthesis and transcription activation in RE. However, this proved to be very difficult to address. In a number of experiments I tried to distinguish whether the transcription itself or the activation of transcription, involving chromatin remodeling around the conserved domains, was required for RE function. Although I was able to show that a promoter is required, I could not dissect the two possibilities. Since, reversing the direction of transcription from the *CUP1* promoter in cells containing the Mcm1 binding site mutation did not significantly reduce the donor preference, the sequence of the non-coding RNAs may not be important for RE function. Previous studies showed that the deletion of the TATA boxes from the inserted *CUP1* promoter reduces the level of transcription but does not prevent the chromatin remodeling activity of the promoter (Shen et al. 2001; Shen et al. 2002). I observed that the deletion of the TATA boxes in the inserted *CUP1* promoter increased the donor preference to high levels, suggesting that the chromatin remodeling activity of the promoter, rather than transcription itself is important for RE function. However, these do not address the question directly. Since the transcription start site of the non-coding RNAs resides in a conserved domain that is known to affect donor preference, the deletion of this region would not distinguish whether transcription or these sequences are required for RE function. Due to the fact that the transcription of the non-coding RNAs diminishes at higher temperatures (Data not shown), a temperature sensitive mutant of the RNA Polymerase II could not be used to test whether the polymerase itself is required for RE function.

Since the donor preference of **a** cells can be increased up to ~ 76% *HML* use without Mcm1 binding, I conclude that the role of Mcm1 in RE function is through

transcription activation, involving the recruitment of chromatin remodeling and/or transcription complexes. I propose that this activation helps recruit the necessary proteins to govern long range chromosomal interactions during mating-type switching.

3.5.3 Fkh1 binding at the RE

Other groups showed that several transcription factors, including Fkh1, bind to RE and regulate donor preference (Sun et al. 2002). The deletion of *FKH1* causes a significant reduction in the donor preference of **a** cells. Surprisingly, the transcription of the non-coding RNAs does not require Fkh1, suggesting that transcription by itself is not sufficient for donor preference. Since the deletion of the Fkh1 protein does not affect the overall chromatin structure of RE, and its binding depends on Mcm1, Fkh1 may act downstream of the chromatin remodeling activity recruited to RE by Mcm1 or *CUP1* promoter. This was supported by the data presenting that the insertion of the *CUP1* promoter enhances Fkh1 binding and also opens the chromatin structure around the conserved domains of RE. Additionally, the level of Fkh1 binding positively correlates with the level of donor preference. These results suggest that Mcm1 facilitates opening the chromatin around the Fkh1 binding sites and this enhances Fkh1 binding. The level of this binding then determines the level of donor preference.

Fkh1 is not required for RE transcription and a promoter cannot bypass its function, suggesting that the role of this protein in RE is different from its role in transcription activation. I also wanted to know whether Fkh1 was involved in elongation of transcription in the RE locus. In the RE region, Fkh1 binding is higher around its

binding sites. This is in contrast to Fkh1 binding at the *CLB2* gene, where it binds more towards the coding region (Morillon et al. 2003). Combined with the results showing that Fkh1 is not required for RE transcription or open chromatin, this suggests that the function of Fkh1 in RE maybe different from its role in both transcription activation and elongation.

The results presented in this chapter suggest that the activation of RE in **a** cells consists of a sequence of events that starts with binding of Mcm1p, which results in transcription activation and open chromatin structure. This is followed by Fkh1 binding, and the level of this binding positively correlates with the level of donor preference. After this, how Fkh1 governs long range chromosomal interactions is not known. Perhaps, Fkh1 serves as a marker for the active RE, where recombination proteins would recognize and be directed to *HML*. This model would predict that Fkh1 may directly interact with one or more proteins that are associated with the double strand break at the *MAT* locus.

The presence of transcription around active RE in **a** cells suggests a role for transcription in increasing the rate of homologous recombination. Another well studied system that connects transcription to recombination involves V(D)J recombination in immune cells. This type of recombination is also directional and the donor sequences that undergo recombination are marked by the presence of an active promoter upstream of the donor (reviewed in (Oltz 2001)). The rate of homologous recombination can also be increased by transcribing the donor sequences (Aguilera et al. 2000; Saxe et al. 2000). The increase in the rate of recombination around *HML* by RE transcription might be due to a similar mechanism. However, the donor of recombination (*HML*) in mating-type

switching is silenced and heterochromatic (reviewed in (Herskowitz 1989; Laurenson and Rine 1992; Haber 1998)). Moreover, the transcription happens ~ 16 kb away from the donor sequence. Therefore, I propose that RE might serve as an entry point for the recombination complex scanning the region for homology. The advantage of having RE as an entry point would be two fold. First, RE activity may be modulated easily according to the cell type by having the α2/Mcm1 operator to regulate transcription and the chromatin structure in a similar manner to **a**-specific genes. This is supported by the data presented in this study along with the previous studies (Szeto and Broach 1997; Szeto et al. 1997; Weiss and Simpson 1997; Wu et al. 1998). Second, if RE serves as an entry point, transcription and/or associated factors may increase the rate of recombination around the whole region, including *HML*. Increased recombination frequency around RE is consistent with the observation that the entire left arm of chromosome III is hot for recombination in **a** cells and the location and distance of RE from *HML* influences donor preference (Wu and Haber 1995).

Chapter 4

Summary

Yeast mating-type switching is commonly used as a model system to study homologous recombination. Although many details of this recombination are revealed, the mechanism of donor preference during switching is not clear. Donor preference is controlled by the Recombination Enhancer (RE). I set out to find how RE functions. First I addressed whether RE governs the primary chromatin structure of an extended region on left arm of chromosome III. In the course of this study, observations about the overall organization of primary chromatin in yeast were made. Second, I studied how RE is activated in **a** cells. In the following two sections, I will summarize my studies on global chromatin organization and then discuss the mechanism of RE activation.

4.1 Global organization of primary chromatin structure

I mapped the primary chromatin structure of ~45 kbp of the left arm of chromosome III by using DNase I (Chapter 2). This is the first study on chromatin structure spanning an extensive region of a yeast chromosome. The mapped region includes 30 ORFs, 6 potential replication origins, 2 known silencers at *HML*α, and 2 Matα2p/Mcm1p operators at the RE.

My analysis revealed that significant features of organized chromatin exist for the entire region and these features are largely static. In contrast to higher organisms, where

chromatin structure undergoes major changes when genes are repressed or transcribed, hypersensitive sites are present at the 5' ends of most genes and some replication origins in the left arm of yeast chromosome III. I observed organized arrays of positioned nucleosomes on ~25% of the genes examined. Occurrence of hypersensitive sites in the promoter region of nearly all the genes suggests that a basic chromatin structural feature exists at every promoter, regardless of its transcriptional status. A recent genome-wide study showed that nucleosomes are distributed heterogeneously over the genome (Lee et al. 2004). There are simply less nucleosomes present at the intergenic compared to the coding regions. This heterogeneous distribution of nucleosomes may contribute to the presence of DNase I hypersensitive sites at the promoter. However, the intergenic regions that are located at the 3' end of the genes are less hypersensitive to DNaseI cutting than 5' ends, thus the presence of less nucleosomes may not be the only contributor. Another contributor to the 5' DNase I hypersensitivity may be the chromatin remodeling that is associated with the promoters.

Several protein complexes remodel chromatin in yeast. A number of these complexes operate at the promoters of the genes and modify or remodel histones (reviewed in (Vignali et al. 2000)). Additionally, a histone variant, H2A.Z replaces H2A at the promoter regions of a number of genes (Larochelle and Gaudreau 2003). Although the stability of the nucleosomes that contain H2A.Z may be reduced, how H2A.Z affects the chromatin structure at the promoters is not known (Suto et al. 2000; Abbott et al. 2001; Park et al. 2004; Placek et al. 2005). Genome-wide studies showed that the binding of other proteins might be restricted to the intergenic regions. For instance, Rap1 protein binds to the promoter regions of a wide variety of the genes that are highly expressed

(Lieb et al. 2001). Another example is the cohesion complex in yeast. The cohesion proteins bind to the intergenic regions, preferably to the 3' ends of two convergent genes (Glynn et al. 2004; Lengronne et al. 2004). Whether binding of the cohesion complex would create or require a specific chromatin structure is not known. Other proteins were shown to affect chromatin structure at the 3' end of genes and have been associated with the termination of transcription (Alen et al. 2002).

My study contributes to the efforts made to understand how yeast organizes primary chromatin structure that is compatible with transcription. The genome is organized in such a way that a basic chromatin structural feature exists at the promoters and the coding regions generally harbor more nucleosomes that are loosely positioned. This basic chromatin structure organization may be required for transcriptional regulation of the genes. The identification of the structural elements, the proteins involved in its creation, and how it relates to further aspects of transcriptional control, will be important in understanding eukaryotic gene regulation.

4.2 The mechanism of RE function

RE regulates donor preference by activating recombination between *MAT* and *HML* in **a** cells. In the following two subsections I will discuss the mechanism of RE function.

4.2.1 Differences and changes in chromatin structure during mating-type switching

In **a** cells, a ~ 40 kbp region on the left arm of chromosome III, including RE and *HML*, is activated for homologous recombination. The same region is "cold" for recombination in α cells. This difference in recombination frequency was attributed to differences in the chromatin structure at this region (Wu and Haber 1995). In order to determine whether there are differences in the chromatin structure between **a** and α cells, I mapped this region at medium resolution with DNase I. Surprisingly, the chromatin structure of the left arm of chromosome III is the same for the two cell types, except in the RE. This suggests that previous expectations of chromatin structural differences underlying the differences in recombinational frequency between the two cell types are not true. Additionally, the chromatin structure of this region is largely static during mating-type switching. Small changes occur in *HML* during switching. One of the changes in the chromatin structure happens at the early stages of switching and is located around the site where the strand invasion is thought to occur. If this change marks the site for strand invasion of most of the cells, the timing of the appearance suggests that the correct donor for recombination is rapidly found, in less than 30 minutes. This is consistent with the observation that the recombining sequences rapidly associate after the HO break at the *MAT* locus (Simon et al. 2002). A unique change in the chromatin structure occurs at the RE during switching in **a**-cells. Other than *HML* and RE, the chromatin structure of the rest of the region remains static during switching.

Since the differences in chromatin structure are confined to RE in **a** and α cells and a unique change occurs at this region during switching, I propose that the RE

chromatin plays a local role that doesn't involve spreading chromatin changes to the surrounding regions. The differences and the changes that were observed at the RE region are around the two α2/Mcm1 operators. Therefore, I focused on understanding their function in RE activation.

4.2.2 RE is activated by a promoter

α2/Mcm1 operators are present at the promoter regions of **a-**specific genes. Although there are no genes at the RE locus, a number of **a**-cell specific non-coding RNAs are transcribed downstream of these operators (Szeto et al. 1997). The existence of cell type dependent transcription at RE was known since 1997. However, the nature of this transcription and how it relates to RE function is not clear. The data presented in this study suggest that the transcription of the non-coding RNAs is cell cycle regulated. Transcription starts at late G1 and continues through the S phase. Yeast mating-type switching is also cell cycle regulated by restricting the *HO* gene expression to the G1 phase. Consequently, switching starts during late G1 and is completed before the end of the S phase (Breeden and Nasmyth 1987). The significance of this regulation is not clear. Nevertheless, normal progression through G1/S is required for both proper switching and donor preference, suggesting that a cell-cycle regulated factor(s) is necessary. Perhaps RE transcription is that cell-cycle regulated factor. Further experiments are required to test this possibility.

A recent study showed that the recombination during mating-type switching may require a set of recombination proteins that are different from other homologous

recombination events and DNA synthesis (Wang et al. 2004). Mating-type switching may be affected by the presence of different recombination proteins at different stages of the cell cycle. This may be the reason why I observed that the recombination was slower when cells were forced to switch during G1. However, it is not known whether these recombination proteins also contribute to the donor preference during switching.

Mcm1 regulates RE activity (Wu et al. 1998). This was demonstrated by showing that a temperature sensitive Mcm1 affects donor preference. Additionally, a mutation that prevents Mcm1 binding also reduces donor preference. In Chapter 3, I presented that the same mutation also abolishes the transcription of the non-coding RNAs. The requirement for Mcm1 binding in donor preference and transcription can be bypassed by inserting a portion of the *CUP1* promoter into the RE. This suggests that the RE activation requires a promoter-related function of Mcm1. Since in **a** cells, open chromatin structure requires Mcm1 binding, this promoter-related function is likely to facilitate the chromatin remodeling at the RE (Wu et al. 1998).

Fkh1 binds and activates RE in **a** cells (Sun et al. 2002). The data presented in Chapter 3 suggest that Fkh1 requires Mcm1 binding at the RE. Moreover, the insertion of *CUP1* promoter, which increases donor preference, also increases Fkh1 binding in cells containing the Mcm1 binding site mutation. Since both Mcm1 and the *CUP1* promoter facilitate the formation of open chromatin at the RE and enhance Fkh1 binding, I propose that the mechanism of RE activation by Mcm1 is to recruit the chromatin remodeling complexes to increase accessibility to the Fkh1 binding sites. This is supported by data indicating that the deletion of Fkh1 protein does not affect the overall chromatin structure of the RE, suggesting that Fkh1 acts downstream of chromatin remodeling.

The results presented in this thesis suggest that the activation of RE in **a** cells consists of a sequence of events that starts with binding of Mcm1 which results in transcription activation and open chromatin structure around the conserved domains of RE. This is followed by Fkh1 binding. The level of Fkh1 binding positively correlates with the level of donor preference, suggesting that this protein governs the activity of the RE. However, the mechanism of Fkh1 function at the RE is not clear. The deletion of Fkh1 causes a reduction in donor preference, yet non-coding RNAs are transcribed. This result suggests that the transcription itself is not sufficient for donor preference. Additionally, a promoter insertion at the RE cannot bypass the requirement for Fkh1, suggesting that the function of this protein may be different from its role in transcription activation. The results presented in Chapter 3 suggest that the function of Fkh1 may be different from its role in transcription elongation as well. Then how does Fkh1 govern donor preference? Perhaps Fkh1 serves as a marker for the active RE, where recombination proteins would recognize and be directed to *HML*. This model would predict that Fkh1 may directly interact with one or more proteins that are associated with the double strand break at the *MAT* locus.

In a number of experiments I tried to distinguish whether the transcription itself or the activation of transcription, involving chromatin remodeling around the conserved domains, was required for RE function (Chapter 3). Although I was able to show that a promoter was required, I could not dissect the two possibilities. Since, inducing transcription in reverse direction did not significantly reduce the donor preference, the sequence of the non-coding RNAs may not be important for RE function. The level of transcription is not important for RE function as well. However, these do not address the

question directly. Deleting the transcription start site of the non-coding RNAs does not offer an answer since this site resides in a conserved domain that is known to affect donor preference. Therefore, the deletion of this region would not distinguish whether transcription or these sequences are required for RE function. Additionally, I could not use a temperature sensitive mutant of the RNA Polymerase II to test whether the polymerase was required for Fkh1 binding because the transcription of the non-coding RNAs diminishes at higher temperatures.

The presence of transcription around the active RE in **a** cells suggest a role for transcription in increasing the rate of homologous recombination. In yeast, the rate of homologous recombination can be increased by transcribing the donor sequences (Aguilera et al. 2000; Saxe et al. 2000). Another well studied system that connects transcription to recombination involves V(D)J recombination in immune cells, where the sequences that undergo recombination are marked by the presence of an active promoter and transcription upstream of the donor (reviewed in (Oltz 2001)). The increase in the rate of recombination around *HML* by RE transcription might be due to a similar mechanism. However, one difference between these systems and mating-type switching is that the transcription happens ~ 16 kb away and the donor is silenced. Therefore, I hypothesize that RE might serve as an entry point for the recombination complex instead of *HML*. The advantage of having RE as an entry point would be two fold. First, RE activity can be modulated easily according to the cell type by having the α2/Mcm1 operator to regulate transcription and the chromatin structure of RE in a similar manner to **a**-specific genes. In **a** cells, Mcm1 protein opens up the chromatin structure around the RE thus activates it. In α cells, RE is turned off by α2 and Mcm1, organizing a repressive

chromatin structure along the conserved domains. Second, if RE serves as an entry point, transcription and/or an associated factor may increase the rate of recombination around the whole region including *HML*. Increased recombination frequency around RE is consistent with the observation that the entire left arm of chromosome III is hot for recombination in **a** cells and the location and distance of RE from *HML* influences donor preference (Wu and Haber 1995). Since transcription activates the rate of recombination, and RE is transcribed in **a** cells, I propose that RE may serve as an entry point to the recombination machinery scanning for homology towards *HML*. If this is the case, the identification of the proteins that are involved in this mechanism of scanning will help us understand how transcription increases the rate of recombination.

Bibliography

Abbott, D. W., V. S. Ivanova, et al. (2001). "Characterization of the stability and folding of H2A.Z chromatin particles: implications for transcriptional activation." J Biol Chem **276**: 41945-9.

Aguilera, A., S. Chavez, et al. (2000). "Mitotic recombination in yeast: elements controlling its incidence." Yeast **16**(8): 731-54.

Alen, C., N. A. Kent, et al. (2002). "A role for chromatin remodeling in transcriptional termination by RNA polymerase II." Mol Cell **10**: 1441-52.

Allan, J., P. G. Hartman, et al. (1980). "The structure of histone H1 and its location in chromatin." Nature **288**: 675-9.

Anguita, E., C. A. Johnson, et al. (2001). "Identification of a conserved erythroid specific domain of histone acetylation across the alpha-globin gene cluster." Proc Natl Acad Sci U S A **98**(21): 12114-9.

Breeden, L. and K. Nasmyth (1987). "Cell cycle control of the yeast HO gene: cis- and trans-acting regulators." Cell **48**(3): 389-97.

Bressan, D. A., J. Vazquez, et al. (2004). "Mating type-dependent constraints on the mobility of the left arm of yeast chromosome III." J Cell Biol **164**: 361-71.

Bruhn, L. and G. F. Sprague, Jr. (1994). "MCM1 point mutants deficient in expression of alpha-specific genes: residues important for interaction with alpha 1." Mol Cell Biol **14**: 2534-44.

Chavez, S., T. Beilharz, et al. (2000). "A protein complex containing Tho2, Hpr1, Mft1 and a novel protein, Thp2, connects transcription elongation with mitotic recombination in Saccharomyces cerevisiae." Embo J **19**(21): 5824-34.

Clever, B., H. Interthal, et al. (1997). "Recombinational repair in yeast: functional interactions between Rad51 and Rad54 proteins." Embo J **16**: 2535-44.

Connolly, B. A., C. I. White, et al. (1988). "Physical monitoring of mating type switching in Saccharomyces cerevisiae." Mol.Cell.Biol. **8**: 2342-2349.

Dorigo, B., T. Schalch, et al. (2004). "Nucleosome arrays reveal the two-start organization of the chromatin fiber." Science **306**: 1571-3.

Dubey, D. D., L. R. Davis, et al. (1991). "Evidence suggesting that the ARS elements associated with silencers of the yeast mating-type locus HML do not function as chromosomal DNA replication origins." Mol.Cell.Biol. **11**: 5436-5355.

Elble, R. and B. K. Tye (1992). "Chromosome loss, hyperrecombination, and cell cycle arrest in a yeast mcm1 mutant." Mol Biol Cell **3**: 971-80.

Ercan, S. and R. T. Simpson (2004). "Global chromatin structure of 45,000 base pairs of chromosome III in a- and alpha-cell yeast and during mating-type switching." Mol Cell Biol **24**: 10026-35.

Gallardo, M. and A. Aguilera (2001). "A new hyperrecombination mutation identifies a novel yeast gene, THP1, connecting transcription elongation with mitotic recombination." Genetics **157**(1): 79-89.

Glynn, E. F., P. C. Megee, et al. (2004). "Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae." PLoS Biol **2**: E259.

Gross, D. S. and W. T. Garrard (1988). "Nuclease hypersensitive sites in chromatin." Annu Rev Biochem **57**: 159-97.

Haber, J. E. (1998). "Mating-type gene switching in Saccharomyces cerevisiae." Annu Rev Genet **32**: 561-99.

Hays, S. L., A. A. Firmenich, et al. (1995). "Complex formation in yeast double-strand break repair: participation of Rad51, Rad52, Rad55, and Rad57 proteins." Proc Natl Acad Sci U S A **92**: 6925-9.

Hebbes, T. R., A. L. Clayton, et al. (1994). "Core histone hyperacetylation comaps with generalized DNase I sensitivity in chicken beta globin chromosomal domain." EMBO J. **13**: 1823-1830.

Hecht, A. and M. Grunstein (1999). "Mapping DNA interaction sites of chromosomal proteins using immunoprecipitation and polymerase chain reaction." Methods Enzymol **304**: 399-414.

Hereford, L., S. Bromley, et al. (1982). "Periodic transcription of yeast histone genes." Cell **30**: 305-10.

Hereford, L. M., M. A. Osley, et al. (1981). "Cell-cycle regulation of yeast histone mRNA." Cell **24**: 367-75.

Herschbach, B. M., M. B. Arnaud, et al. (1994). "Transcriptional repression directed by the yeast alpha 2 protein in vitro." Nature **370**: 309-11.

Herskowitz, I. (1989). "A regulatory hierarchy for cell specialization in yeast." Nature **342**: 749-57.

Hochstrasser, M., M. J. Ellison, et al. (1991). "The short-lived MAT alpha2 transcriptional regulator is ubiquitinated in vivo." Proc.Natl.Acad.Sci.USA **88**: 4606-4610.

Holstege, F. C. P., E. G. Jennings, et al. (1998). "Dissecting the regulatory circuitry of a eukaryotic genome." Cell **95**(5): 717-728.

Huang, R. Y. and D. Kowalski (1996). "Multiple DNA elements in ARS305 determine replication origin activity in a yeast chromosome." Nucleic Acids Res **24**(5): 816-23.

Jaskelioff, M., S. Van Komen, et al. (2003). "Rad54p is a chromatin remodeling enzyme required for heteroduplex DNA joint formation with chromatin." J Biol Chem.

Jenuwein, T. and C. D. Allis (2001). "Translating the histone code." Science **293**: 1074-80.

Jin, Y., G. Binkowski, et al. (1997). "Ho endonuclease cleaves MAT DNA in vitro by an inefficient stoichiometric reaction mechanism." J Biol Chem **272**(11): 7352-9.

Kaplan, C. D., L. Laprade, et al. (2003). "Transcription elongation factors repress transcription initiation from cryptic sites." Science **301**(5636): 1096-9.

Keleher, C. A., M. J. Redd, et al. (1992). "Ssn6-Tup1 is a general repressor of transcription in yeast." Cell **68**: 709-19.

Klar, A. J., J. B. Hicks, et al. (1982). "Directionality of yeast mating-type interconversion." Cell **28**(3): 551-61.

Knezetic, J. A. and D. S. Luse (1986). "The presence of nucleosomes on a DNA template prevents initiation by RNA polymerase II in vitro." Cell **45**: 95-104.

Kornberg, R. D. and L. Stryer (1988). "Statistical distributions of nucleosomes: nonrandom locations by a stochastic mechanism." Nucleic.Acids.Res. **16**: 6677-6690.

Kostriken, R., J. N. Strathern, et al. (1983). "A site-specific endonuclease essential for mating-type switching in Saccharomyces cerevisiae." Cell **35**: 167-74.

Kostriken, R. and C. J. Wedeen (2001). "Engineered interphase chromosome loops guide intrachromosomal recombination." Embo J **20**: 2907-13.

Kumar, R., D. M. Reynolds, et al. (2000). "Forkhead transcription factors, Fkh1p and Fkh2p, collaborate with Mcm1p to control transcription required for M-phase." Curr Biol **10**: 896-906.

Larochelle, M. and L. Gaudreau (2003). "H2A.Z has a function reminiscent of an activator required for preferential binding to intergenic DNA." Embo J **22**: 4512-22.

Laurenson, P. and J. Rine (1992). "Silencers, Silencing, and Heritable Transcriptional States." Microbiol.Rev. **56**: 543-560.

Lee, C. K., Y. Shibata, et al. (2004). "Evidence for nucleosome depletion at active regulatory regions genome-wide." Nat Genet **36**(8): 900-5.

Lengronne, A., Y. Katou, et al. (2004). "Cohesin relocation from sites of chromosomal loading to places of convergent transcription." Nature **430**: 573-8.

Levy, A. and M. Noll (1981). "Chromatin fine structure of active and repressed genes." Nature **289**: 198-203.

Li, Y. C., T. H. Cheng, et al. (2001). "Establishment of transcriptional silencing in the absence of DNA replication." Science **291**(5504): 650-3.

Lieb, J. D., X. Liu, et al. (2001). "Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association." Nat Genet **28**(4): 327-34.

Litt, M. D., M. Simpson, et al. (2001). "Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci." Embo J **20**(9): 2224-35.

Longtine, M. S., A. McKenzie, 3rd, et al. (1998). "Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae." Yeast **14**: 953-61.

Lorch, Y., J. W. LaPointe, et al. (1987). "Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones." Cell **49**: 203-10.

Luger, K., A. W. Mader, et al. (1997). "Crystal structure of the nucleosome core particle at 2.8 angstrom resolution." Nature **389**: 251-260.

Lutter, L. C. (1989). "Digestion of nucleosomes with deoxyribonucleases I and II." Methods.Enzymol. **170**: 264-269.

Matsui, T. (1987). "Transcription of adenovirus 2 major late and peptide IX genes under conditions of in vitro nucleosome assembly." Mol Cell Biol **7**: 1401-8.

Miller, A. M. and K. A. Nasmyth (1984). "Role of DNA replication in the repression of silent mating type loci in yeast." Nature **312**: 247-51.

Moreira, J. M. and S. Holmberg (1998). "Nucleosome structure of the yeast CHA1 promoter: analysis of activation-dependent chromatin remodeling of an RNA-

polymerase-II-transcribed gene in TBP and RNA pol II mutants defective in vivo in response to acidic activators." Embo J **17**: 6028-38.

Morillon, A., J. O'Sullivan, et al. (2003). "Regulation of elongating RNA polymerase II by forkhead transcription factors in yeast." Science **300**: 492-5.

Murphy, M. R., M. Shimizu, et al. (1993). "DNA-protein interactions at the *S. cerevisiae* alpha2 operator *in vivo*." Nucleic.Acids.Res. **21**: 3295-3300.

Nickoloff, J. A., E. Y. Chen, et al. (1986). "A 24-base-pair DNA sequence from the MAT locus stimulates intergenic recombination in yeast." Proc Natl Acad Sci U S A **83**: 7831-5.

Ogawa, T., X. Yu, et al. (1993). "Similarity of the yeast RAD51 filament to the bacterial RecA filament." Science **259**: 1896-9.

Oliver, S. G., Q. J. M. van der Aart, et al. (1992). "The complete DNA sequence of yeast chromosome III." Nature **357**: 38-46.

Oltz, E. M. (2001). "Regulation of antigen receptor gene assembly in lymphocytes." Immunol Res **23**: 121-33.

Paques, F., W. Y. Leung, et al. (1998). "Expansions and contractions in a tandem repeat induced by double-strand break repair." Mol Cell Biol **18**: 2045-54.

Park, Y. J., P. N. Dyer, et al. (2004). "A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome." J Biol Chem **279**: 24274-82.

Placek, B. J., L. N. Harrison, et al. (2005). "The H2A.Z/H2B dimer is unstable compared to the dimer containing the major H2A isoform." Protein Sci **14**: 514-22.

Rattray, A. J. and L. S. Symington (1995). "Multiple pathways for homologous recombination in Saccharomyces cerevisiae." Genetics **139**: 45-56.

Ravindra, A., K. Weiss, et al. (1999). "High resolution structural analysis of chromatin at specific loci: Saccharomyces cerevisiae silent mating type locus HMRa." Mol.Cell.Biol.: in press.

Rine, J. and I. Herskowitz (1987). "Four genes responsible for a position effect on expression from HML and HMR in Saccharomyces cerevisiae." Genetics **116**(1): 9-22.

Roth, S. Y., M. Shimizu, et al. (1992). "Stable nucleosome positioning and complete repression by the yeast alpha 2 repressor are disrupted by amino-terminal mutations in histone H4." Genes Dev **6**(3): 411-25.

Sandell, L. L. and V. A. Zakian (1993). "Loss of a yeast telomere: arrest, recovery, and chromosome loss." Cell **75**: 729-39.

Sauer, R. T., D. L. Smith, et al. (1988). "Flexibility of the yeast alpha2 repressor enables it to occupy the ends of its operator, leaving the center free." Genes.Dev. **2**: 807-816.

Saxe, D., A. Datta, et al. (2000). "Stimulation of mitotic recombination events by high levels of RNA polymerase II transcription in yeast." Mol Cell Biol **20**(15): 5404-14.

Scherer, S. and R. W. Davis (1979). "Replacement of chromosome segments with altered DNA sequences constructed in vitro." Proc Natl Acad Sci U S A **76**: 4951-5.

Sharma, V. M., B. Li, et al. (2003). "SWI/SNF-dependent chromatin remodeling of RNR3 requires TAF(II)s and the general transcription machinery." Genes Dev **17**: 502-15.

Shen, C. H., B. P. Leblanc, et al. (2001). "Remodeling of yeast CUP1 chromatin involves activator-dependent repositioning of nucleosomes over the entire gene and flanking sequences." Mol Cell Biol **21**(2): 534-47.

Shen, C. H., B. P. Leblanc, et al. (2002). "Targeted histone acetylation at the yeast CUP1 promoter requires the transcriptional activator, the TATA boxes, and the putative histone acetylase encoded by SPT10." Mol Cell Biol **22**(18): 6406-16.

Shimizu, M., S. Y. Roth, et al. (1991). "Nucleosomes are positioned with base pair precision adjacent to the α2 operator in Saccharomyces cerevisiae." EMBO J. **10**: 3033-3041.

Shinohara, A., H. Ogawa, et al. (1992). "Rad51 protein involved in repair and recombination in S. cerevisiae is a RecA-like protein." Cell **69**: 457-70.

Shrader, T. E. and D. M. Crothers (1989). "Artificial nucleosome positioning sequences." Proc.Natl.Acad.Sci.U.S.A. **86**: 7418-7422.

Sikorski, R. S. and P. Hieter (1989). "A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae." Genetics **122**: 19-27.

Simon, P., P. Houston, et al. (2002). "Directional bias during mating type switching in Saccharomyces is independent of chromosomal architecture." Embo J **21**: 2282-91.

Simpson, R. T. (1990). "Nucleosome positioning can affect the function of a cis-acting DNA element in vivo." Nature **343**(6256): 387-9.

Simpson, R. T. (1998). "Chromatin structure and analysis of mechanisms of activators and repressors." Methods **15**: 283-294.

Simpson, R. T. (1999). "In vivo methods to analyze chromatin structure." Curr Opin Genet Dev **9**: 225-9.

Simpson, R. T. and D. W. Stafford (1983). "Structural features of a phased nucleosome core particle." Proc.Natl.Acad.Sci.USA **80**: 51-55.

Sippel, A. E., H. Saueressig, et al. (1996). "Identification of cis-acting elements as DNase I hypersensitive sites in lysozyme gene chromatin." Methods Enzymol **274**: 233-46.

Smith, D. L. and A. D. Johnson (1994). "Operator-constitutive mutations in a DNA sequence recognized by a yeast homeodomain." EMBO J. **13**: 2378-2387.

Spellman, P. T., G. Sherlock, et al. (1998). "Comprehensive identification of cell cycle-regulated genes of the yeast Saccharomyces cerevisiae by microarray hybridization." Mol.Biol.Cell **9**: 3273-3297.

Staynov, D. Z. and C. Crane-Robinson (1988). "Footprinting of linker histones H5 and H1 on the nucleosome." Embo J **7**: 3685-91.

Staynov, D. Z. and Y. G. Proykova (1998). "Quantitative analysis of DNase I digestion patterns of oligo- and polynucleosomes." J Mol Biol **279**(1): 59-71.

Strathern, J. N. and I. Herskowitz (1979). "Asymmetry and directionality in production of new cell types during clonal growth: the switching pattern of homothallic yeast." Cell **17**(2): 371-81.

Strathern, J. N., A. J. Klar, et al. (1982). "Homothallic switching of yeast mating type cassettes is initiated by a double-stranded cut in the MAT locus." Cell **31**(1): 183-92.

Sugawara, N., E. L. Ivanov, et al. (1995). "DNA structure-dependent requirements for yeast RAD genes in gene conversion." Nature **373**: 84-6.

Sugawara, N., X. Wang, et al. (2003). "In vivo roles of Rad52, Rad54, and Rad55 proteins in Rad51-mediated recombination." Mol Cell **12**: 209-19.

Sun, K., E. Coic, et al. (2002). "Saccharomyces forkhead protein Fkh1 regulates donor preference during mating-type switching through the recombination enhancer." Genes Dev **16**(16): 2085-96.

Sung, P. (1994). "Catalysis of ATP-dependent homologous DNA pairing and strand exchange by yeast RAD51 protein." Science **265**: 1241-3.

Sung, P. (1997). "Yeast Rad55 and Rad57 proteins form a heterodimer that functions with replication protein a to promote DNA strand exchange by Rad51 recombinase." Gene Develop. **11**(9): 1111-1121.

Suto, R. K., M. J. Clarkson, et al. (2000). "Crystal structure of a nucleosome core particle containing the variant histone H2A.Z." Nat Struct Biol **7**: 1121-4.

Szeto, L. and J. R. Broach (1997). "Role of α2 protein in donor locus selection during mating type interconversion." Mol.Cell.Biol. **17**: 751-759.

Szeto, L., M. K. Fafalios, et al. (1997). "Alpha2p controls donor preference during mating type interconversion in yeast by inactivating a recombinational enhancer of chromosome III." Genes Dev **11**(15): 1899-911.

Thoma, F. (1986). "Protein-DNA interactions and nuclease sensitive regions determine nucleosome positions on yeast plasmid chromatin." J.Mol.Biol. **190**: 177-190.

van Holde, K. E. (1988). Chromatin. New York, Springer-Verlag.

Venditti, S. and G. Camilloni (1994). "In vivo analysis of chromatin following nystatin-mediated import of active enzymes into Saccharomyces cerevisiae." Mol Gen Genet **242**: 100-4.

Vignali, M., A. H. Hassan, et al. (2000). "ATP-dependent chromatin-remodeling complexes." Mol Cell Biol **20**: 1899-910.

Vujcic, M., C. A. Miller, et al. (1999). "Activation of silent replication origins at autonomously replicating sequence elements near the HML locus in budding yeast." Mol Cell Biol **19**(9): 6098-109.

Wang, X., G. Ira, et al. (2004). "Role of DNA replication proteins in double-strand break-induced recombination in Saccharomyces cerevisiae." Mol Cell Biol **24**: 6891-9.

Wang, X. and R. T. Simpson (2001). "Chromatin structure mapping in Saccharomyces cerevisiae in vivo with DNase I." Nucleic Acids Res **29**(9): 1943-50.

Weiler, K. S. and J. R. Broach (1992). "Donor locus selection during Saccharomyces cerevisiae mating type interconversion responds to distant regulatory signals." Genetics **132**(4): 929-42.

Weintraub, H. and M. Groudine (1976). "Chromosomal subunits in active genes have an altered conformation." Science **193**: 848-856.

Weiss, K. and R. T. Simpson (1997). "Cell type-specific chromatin organization of the region that governs directionality of yeast mating type switching." Embo J **16**(14): 4352-60.

Weiss, K. and R. T. Simpson (1998). "High-resolution structural analysis of chromatin at specific loci: Saccharomyces cerevisiae silent mating type locus HMLalpha." Mol Cell Biol **18**(9): 5392-403.

White, C. I. and J. E. Haber (1990). "Intermediates of recombination during mating type switching in Saccharomyces cerevisiae." Embo J **9**(3): 663-73.

Winston, F., C. Dollard, et al. (1995). "Construction of a set of convenient Saccharomyces cerevisiae strains that are isogenic to S288C." Yeast **11**: 53-5.

Wolner, B., S. van Komen, et al. (2003). "Recruitment of the recombinational repair machinery to a DNA double-strand break in yeast." Mol Cell **12**: 221-32.

Workman, J. L. and R. G. Roeder (1987). "Binding of transcription factor TFIID to the major late promoter during in vitro nucleosome assembly potentiates subsequent initiation by RNA polymerase II." Cell **51**: 613-22.

Wu, C., P. M. Bingham, et al. (1979). "The chromatin structure of specific genes: I. Evidence for higher order domains of defined DNA sequence." Cell **16**: 797-806.

Wu, C. and W. Gilbert (1981). "Tissue specific exposure of chromatin structure at the 5' terminus of the rat preproinsulin gene." Proc.Natl.Acad.Sci.USA **78**: 1577-1580.

Wu, C., K. Weiss, et al. (1998). "Mcm1 regulates donor preference controlled by the recombination enhancer in Saccharomyces mating-type switching." Genes Dev **12**(11): 1726-37.

Wu, C., Y. C. Wong, et al. (1979). "The chromatin structure of specific genes: II. Disruption of chromatin structure during gene activity." Cell **16**(4): 807-14.

Wu, X. and J. E. Haber (1995). "MATa donor preference in yeast mating-type switching: activation of a large chromosomal region for recombination." Genes Dev **9**(15): 1922-32.

Wu, X. and J. E. Haber (1996). "A 700 bp cis-acting region controls mating-type dependent recombination along the entire left arm of yeast chromosome III." Cell **87**(2): 277-85.

Wu, X., J. K. Moore, et al. (1996). "Mechanism of MAT alpha donor preference during mating-type switching of Saccharomyces cerevisiae." Mol Cell Biol **16**(2): 657-68.

Wu, X., C. Wu, et al. (1997). "Rules of donor preference in saccharomyces mating-type gene switching revealed by a competition assay involving two types of recombination." Genetics **147**(2): 399-407.

Wyrick, J. J., F. C. Holstege, et al. (1999). "Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast." Nature **402**(6760): 418-21.

Yu, L. and R. H. Morse (1999). "Chromatin opening and transactivator potentiation by RAP1 in Saccharomyces cerevisiae." Mol Cell Biol **19**(8): 5279-88.

Zhang, Z. and J. C. Reese (2004). "Redundant mechanisms are used by Ssn6-Tup1 in repressing chromosomal gene transcription in Saccharomyces cerevisiae." J Biol Chem **279**: 39240-50.

Zhu, G., P. T. Spellman, et al. (2000). "Two yeast forkhead genes regulate the cell cycle and pseudohyphal growth." Nature **406**: 90-4.

Appendix A

Supplementary Data

A.1 *CUP1* promoter insertion failed to activate RE during G1 phase of the cell cycle

Figure 3-2 suggested that a factor is required for proper switching and correct donor preference during late G1 and S phase. Additionally, RE transcription starts during late G1 and continues through the S phase (Figure 3-1). I wanted to test whether RE transcription was the necessary factor for cells to choose the correct donor. For this, I planned to induce transcription by using the *CUP1* promoter inserted at RE and force the cells to switch during G1. This could test whether *CUP1* promoter would rescue the donor preference of cells switching during G1. Cells that were forced to switch during G1 exhibited low donor preference of 45% and the cells containing the *CUP1* promoter insertion showed 49% donor preference (Figure A-1A). However, the analysis of RE transcription showed that the cells containing the *CUP1* promoter at the RE could not induce transcription at normal levels during the arrest (Figure A-1B). Therefore, I cannot rule out the possibility that *CUP1* promoter somehow failed to activate RE during G1.

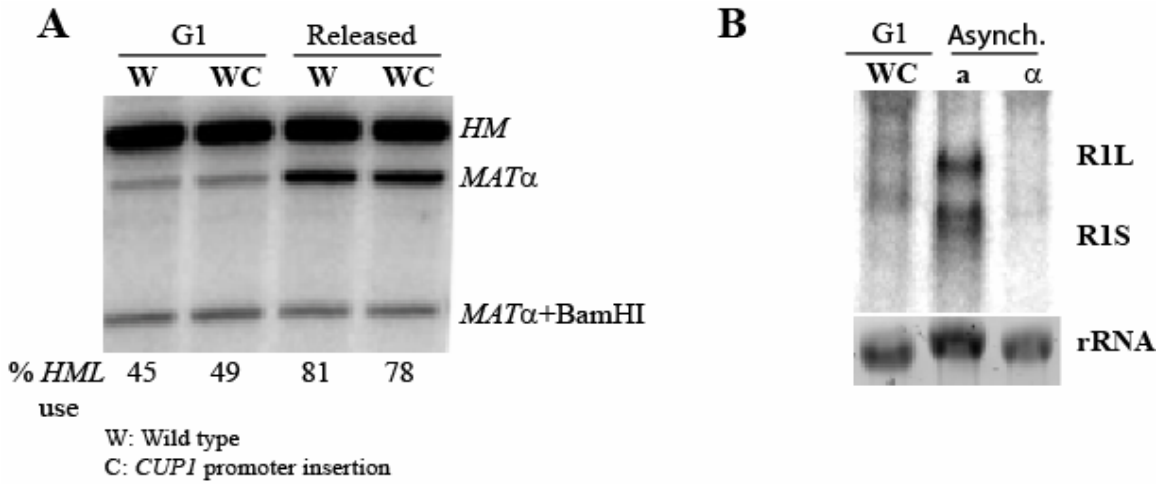

Figure A-1: *CUP1* promoter insertion failed to activate RE during G1 phase of the cell cycle

A: Donor preference analysis. Strains that were used in this experiment carry a BamHI marked *HML* in place of the *HMR***a**. HO induced switching from *MAT***a** to *MAT*α or *MAT*α+BamHI was monitored by Southern Blot analysis. DNA isolated from switched cells was digested with HindIII and BamHI. A specific probe was used to give fragments that correspond to *MAT*α and *MAT*α+BamHI, 4.4 kb and 3.1 kb respectively. Donor preference is shown as percent *HML* use at the bottom of each lane. Cells with *CUP1* insertion at RE (WC) or without this insertion (W) were forced to switch during G1 (G1) or released from arrest (Released) as explained in part A. B: Northern Blot analysis of RE transcription in cells from Part B along with a and α cells.

A.2 The chromatin structure of RE in cells containing *CUP1* promoter insertion

In order to determine how *CUP1* promoter insertion affects the chromatin structure of the region downstream of the insertion, another secondary enzyme was used to generate indirect end mapping focusing at this region Figure A-2.

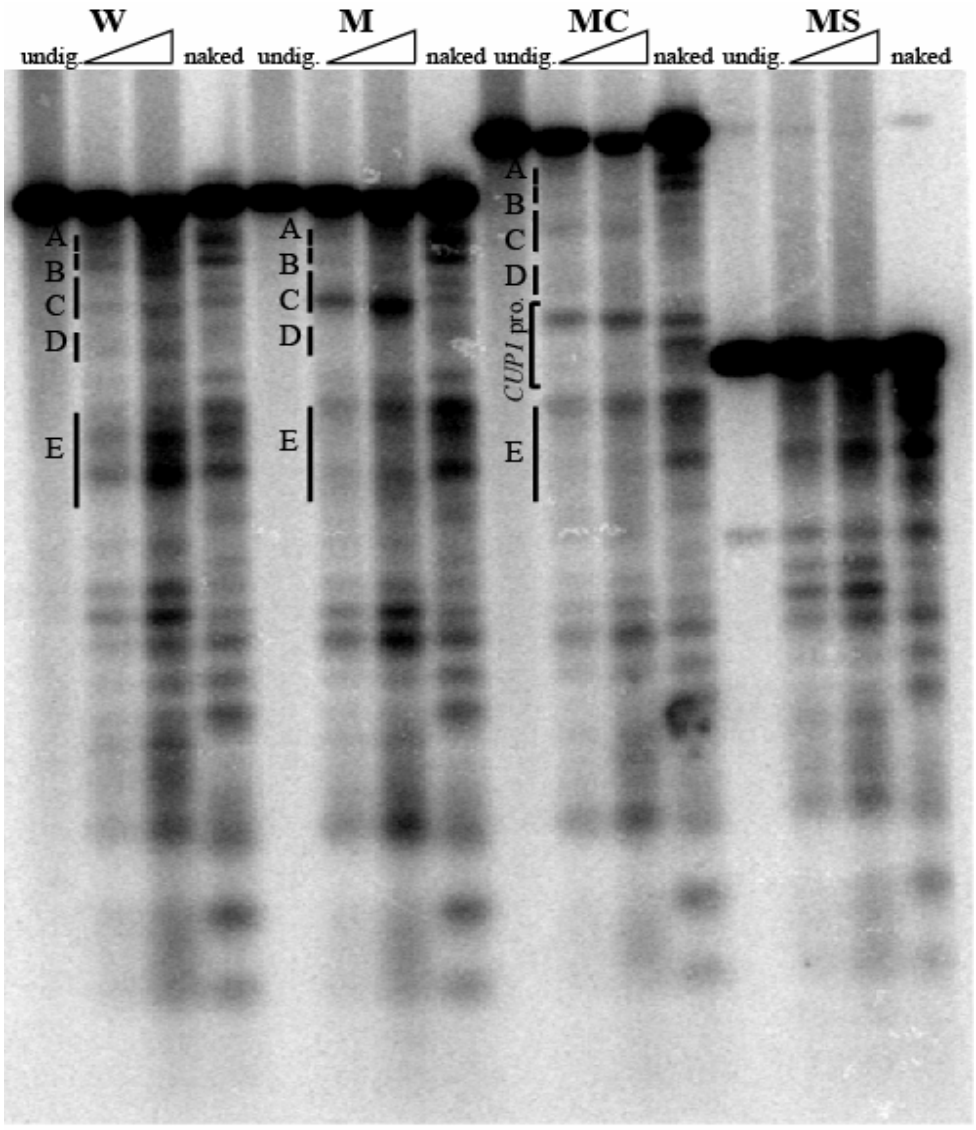

Figure A-2: MNase I map of the chromatin structure downstream of the conserved domains

The chromatin structure was mapped at low resolution by MNaseI. The domains of RE are shown to the left of each sample. This map was generated by using SpeI enzyme, which also cuts in the control sequence, thus producing a smaller parental fragment in MS cells.

Appendix B

Projects in Progress

B.1 Cell cycle regulation of chromatin organization at RE

Mating-type switching is regulated by cell cycle. HO endonuclease initiates gene conversion during late G1 by introducing a double strand break at the *MAT* locus. This double strand break is repaired before the DNA synthesis is completed (Connolly et al. 1988). Mating-type switching in yeast cells results in expression of a new set of transcription factors from the switched locus. For instance, when **a** cells undergo switching, the **a** allele present at the *MAT* locus changes to α allele, therefore α1 and α2 expression starts. α2, along with Mcm1, represses the transcription of **a**-specific genes by organizing a repressive chromatin domain over the coding region. Since the cells switch during G1 and S phase, I want to know when this repressive chromatin domain is formed. S phase is required for the establishment of repressive chromatin at the mating-type cassettes (Miller and Nasmyth 1984; Li et al. 2001).

In order to test whether RE and **a-**specific genes also require S phase for the reorganization of chromatin, the expression of α2 will be induced in cells that are synchronized at different stages of the cells cycle and the presence of repressive chromatin at the RE and **a-**specific genes will be monitored. For this, α2 gene was cloned downstream of *GAL1-10* promoter and inserted into the genome at the *ura3* locus. Galactose induction was performed in **a** cells containing the Gal1-10-α2 construct, (Gal-

α2) as well as in cells lacking this construct (YPH499**a**). The expression of α2 and an **a**-specific gene (*STE6*) was monitored by northern blot analysis (Figure B-1A). STE6 transcription decreased when α2 was induced. However, the level of STE6 transcript was low when Gal-α2 cells were grown in raffinose containing media, suggesting that leaky expression of α2 might occur (compare Gal-α2 (R) to YPH499 (R) (Figure B-1A).

In order to monitor the formation of a repressive chromatin domain over RE and **a-**specific genes, I decided to develop restriction enzyme accessibility assays for each. I used a method that allows permeabilizing cells with nystatin for *in vivo* analysis (Venditti and Camilloni 1994). At the RE region, DdeI accessibility marked the presence of different chromatin structure in **a** and α cells (Figure B-1B). In **a** cells, RE was accessible to DdeI and in α cells it was not accessible. Therefore, I can use this method to determine the presence of repressive chromatin at RE when α2 is expressed at different phases of the cell cycle. Next, I would like to optimize the α2 expression system and determine whether chromatin at RE and **a-**specific genes can be compacted at any stage of the cell cycle.

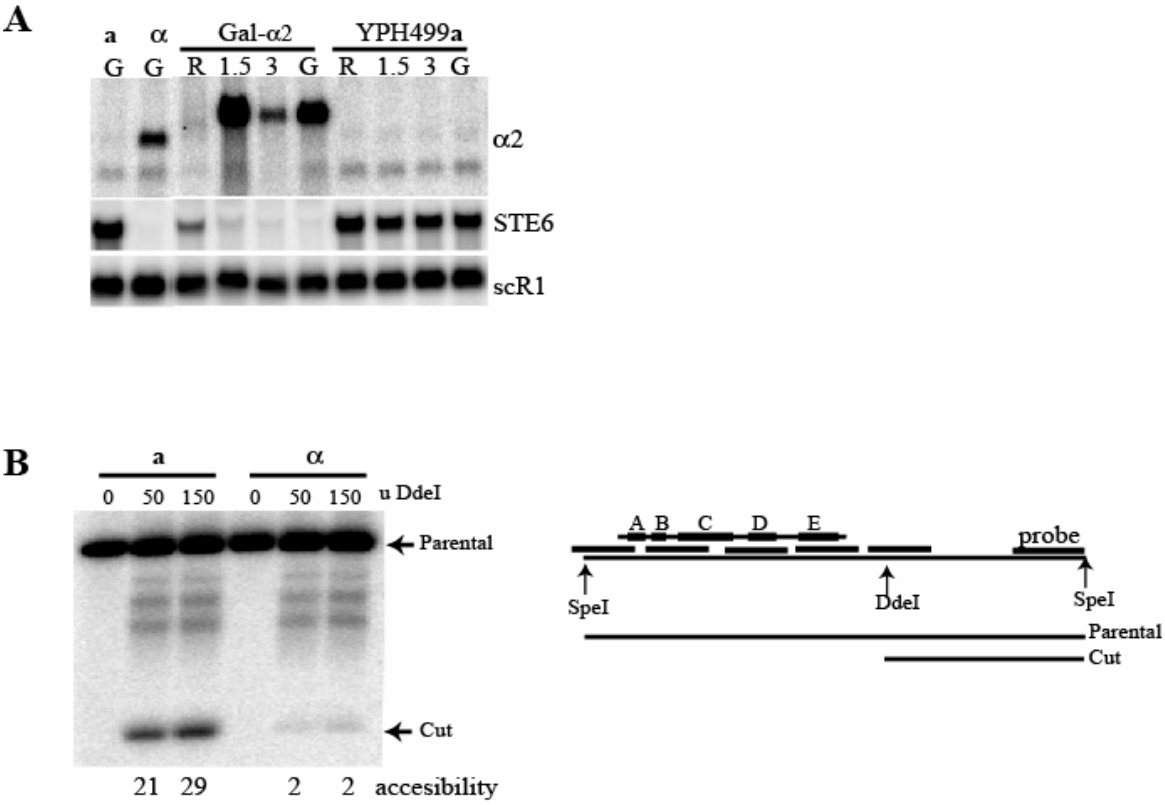

Figure B-1: Galactose inducible α2 expression system and RE accessibility assays

A: Northern blot analysis of α2 expression. **a** and α cells were grown in galactose containing media (G). Cells containing the Gal110-α2 construct (Gal-α2) or not (YPH499**a**) were grown in raffinose (R) or galactose (G) containing media, overnight. To the cells grown in raffinose, 2% galactose was added for 1.5 hrs (1.5) or 3 hours (3). RNA was purified and specific probes were used to analyze α2 and STE6 transcripts. scR1 was used as loading control. B: Restriction enzyme accessibility assay at RE in **a** and α cells. Spheroblasts were incubated in digestion buffer containing 100 ug/ml nystatin, increasing amounts of DdeI was used to digest chromatin. Amount of DdeI is indicated in units per reaction on top of each lane. DNA was isolated from spheroblasts and SpeI was used as secondary enzyme to generate the parental and cut fragments in the southern blots. The schematic representation of RE domains and approximate locations of first 5 nucleosomes are shown to the right of the gel. The fragments resulting from DdeI and SpeI cut, along with the location of the probe, are shown to the right of the gel.

B.2 Materials and Methods

B.2.1 Cloning and expression of α2 and the northern blot analysis of RNA

Alpha2 ORF was cloned into pESC-URA (Stratagene) into SpeI/XhoI site. ORF was amplified by PCR from Chr III, 12380-13018. KpnI/NaeI fragment from pESC-ura+alpha2 was sub-cloned into pRS406 (pRS406+alpha2). This plasmid was used to insert the construct into the genome at *ura3* locus in YPH499**a** cells (Sikorski and Hieter 1989). Galactose induction was performed by adding galactose to the raffinose containing media when cells grew to an OD_{600} of ~0.4. Total RNA was prepared from 20 mls of cells by glass bead cell disruption as described previously (Zhang and Reese 2004). For Northern blotting, 20 μg of each RNA sample was subjected to electrophoresis in a formaldehyde/agarose gel, followed by transfer to a nylon membrane, and hybridization to specific probes. Probes for Northern blot hybridization were prepared by amplification from genomic DNA by PCR. Probe for α2 signal: Chr III, 12957-13257, STE6: Chr XI, 45413-46230. Gel purification of the probes was followed by random primer labeling (Stratagene). An RNA polymerase III-transcribed transcript, scR1: Chr V, 441736-442428, was used as loading control. Blots were stripped for re-probing by boiling in 0.1% SDS.

B.2.2 Restriction enzyme accessibility assay

For determining DdeI accessibility, FY24 α and FY23**a** cells (Winston et al. 1995) were grown in standard yeast media (YP) with 2% dextrose at 30ºC to OD_{600} =1.0.

After harvesting, they were washed twice with 20 ml Sorbitol buffer (1.4M Sorbitol, 40 mM HEPES ph7.5, 0.5 mM $MgCl_2$). Cell pellet was weighed, resuspended in 4XW (weight in grams) ml of Sorbitol buffer containing 0.5 mg/ml Zymolyase (Seigaku) for ~30 min at 30^{o}C. After zymolyase treatment, nystatin was used as described previously (Venditti and Camilloni 1994). Permeabilized spheroblasts were incubated in digestion buffer (50 mM NaCl, 1.5 mM $CaCl_2$, 20 mM Tris.Cl, pH 8, 1 M sorbitol and 100 μg/ml nystatin) with 50 and 150 u/reaction of DdeI. DNA was isolated after Proteinase K and RNaseA treatment by phenol/chloroform extraction and ethanol precipitated. Southern blots were performed as described previously in Chapters 2 and 3. DNA from cells was purified, cut with SpeI, subjected to electrophoretic separation on agarose gels, transferred to Hybond-NX membrane (Amersham), cross linked with UV light and hybridized with a specific probe. The probe was gel purified and random primer labeled with [α-^{32}P]dATP. The blots was exposed to a PhosphorImager screen and analyzed using ImageQuant v.5.

VITA

Sevinc Ercan

Education

1995 – 1999	Bachelor of Science - Molecular Biology and Genetics Bilkent University, Ankara, Turkey

Publications

Ercan, S., Simpson R.T. 2004. Global Chromatin Structure of 45,000 Base Pairs of Chromosome III in **a**- and α-Cell Yeast and During Mating-Type Switching. *Molecular and Cellular Biology*. **24(22)**:10026-35.

Ercan, S., Carrozza, M. J. and J. L. Workman. 2004. Global Nucleosome Distribution and the Regulation of Transcription in Yeast. *Genome Biology.* **5(10):**243.

Academic Honors

1995 - 1999	Undergraduate Student Scholarship, Bilkent University, Ankara, Turkey
2000	2000 Paul Berg Prize, The Pennsylvania State University