Compaction of chromatin in *Saccharomyces cerevisiae*

Cheng-Fu Chang

Submitted in fulfilment of the requirements for the degree Master of Science in the

Department of Molecular and Cell Biology, University of Cape Town

2006

Cape Town

Declaration

The experimental work described in this report was carried out in the Department of Molecular and Cell Biology, University of Cape Town, from February 2004 to December 2005, under the supervision of Associate Professor Hugh G. Patterton. The results presented here are the original, unaided work of the author. Where use has been made of the work of others it is duly acknowledged in the text.

Cheng-Fu Chang

February 2006

Acknowledgement

My most sincere gratitude goes to Hugh G. Patterton, for his brilliance, expertise and patience. Much of this thesis also owes to the guidance and aid of Dr Georgia Schafer, Dr Chris McEnvoy, Sachin Somers and Tim Frows for their endless support and guidance through the years. Special thanks go to the Department of Biotechnology, University of Free State, for their generosity and hospitality in time of despair. I am forever indebt to every smile and kind word that was with me all the way. I thank the Council of Scientific and Industrial Research (CSIR) for the scholarship awarded. This research was funded by the Wellcome Trust and the National Research Foundation (NRF).

Abstract

The DNA of eukaryotes is packaged into chromatin that restricts access of transcription machineries and other DNA-binding complexes to their targets. It was proposed that transcription activity can be regulated through the formation of specific compact higher-order chromatin structures. However, the relationship between such compacted structures and transcription activity remains elusive.

This study investigated the link between the association of the yeast linker histone homologue, Hho1p, and the compaction of the yeast genome during stationary phase. The relative gene content of condensed chromatin, fractionated and isolated by sucrose gradient ultracentrifugation from stationary and exponential phase cultures was compared using genome-wide technologies. This study showed that condensed chromatin of stationary phase culture contained an enriched density of genes on all the chromosomes, indicating global compaction of the yeast genome during stationary phase. In addition, several genes of interest were identified that displayed transcriptional activity during exponential phase and were found to be enriched in condensed fraction of the chromatin.

In order to obtain a detailed view of the chromatin compaction status at the level of individual genes, a reagent to accurately measure chromatin compaction *in vivo* was designed. A novel recombinant protein composed of a fusion of MNase and LexA was developed in this study. Future work will involve the use of this protein at specific loci identified in this study.

Table of contents

Abbreviations

Å - angstroms

ATP – adenosine tri-phosphate

bp – base pairs

BSA – bovine serum albumin

°C - degrees Celsius

DNaseI – deoxyribonuclease I

DNA – deoxyribonucleic acid

DTT - dithiothreitol

EDTA - ethylenediaminetetraacetic acid

EGTA - ethylene glycol bis(2-aminoethyl ether)-N,N,N'N'-tetraacetic acid

ISWI – imitation of Switch

kb – kilobase pairs

$MgCl_2$ – magnesium chloride

MNase - micrococcal nuclease

µg - microgram

µl - microlitre

mg - milligram

ml - millilitre

mM - millimolar

min - minute

M - molar

NaCl – sodium chloride

ng - nanogram

NLS – nuclear localization signal

ORF – open reading frame

pg - picogram

PCR – polymerase chain reaction

rRNA – ribosomal ribonucleic acid

sec - second

SDS – sodium dodecyl sulphate

SWI/SNF – switching/sucrose non-fermenting

TE – Tris-EDTA

TAE – Tris-acetate EDTA

U - unit

UV – ultra violet

V – volt

Chapter 1

Introduction

1.1. Chromatin in eukaryotes

The packaging of DNA in chromatin offers an efficient method of storage of genetic material in the constrained nuclear space of eukaryotic organisms. However, it also determines the accessibility of the genetic molecule to transcription, replication, recombination and repair machineries. Thus the structure of chromatin plays a key role in the regulation of DNA function, including gene expression (Wolffe, 2001)

Chromatin in eukaryotes can take up two different states: the transcriptionally active euchromatin, or the transciptionally silent heterochromatin. Both euchromatin and heterochromatin have distinctive structural and biochemical properties (Richards and Elgin, 2002). Histone proteins play a major role in chromatin status, and many transcription regulatory processes are controlled through the association of DNA with histone proteins.

1.2. The nucleosome

The fundamental packaging unit of chromatin is the nucleosome. A nucleosome (Fig. 1) is formed by the association of two copies of each of four core histones, namely H3, H4, H2A and H2B, which are highly conserved amongst the different eukaryotic organisms (Luger *et al.*, 1997). The helix-loop-helix-loop-helix structure, also known as the histone fold, of the core histones allows the formation of H3-H4 and H2A-H2B

dimers through a hand-shake motif. The H3-H4 dimers associate to form a tetramer through a four-helix bundle interaction between the C-termini of the two copies of histone H3. Formation of similar four-helix bundle structures between the C-termini of H4 and H2B leads to the binding of the two H2A-H2B dimers to the H3-H4 tetramer, and the formation of the core histone octamer (Fig. 1.). Approximately 146bp of DNA is wrapped around the disc shaped core histone octamer in approximately 1.75 turns of a left-handed superhelix.

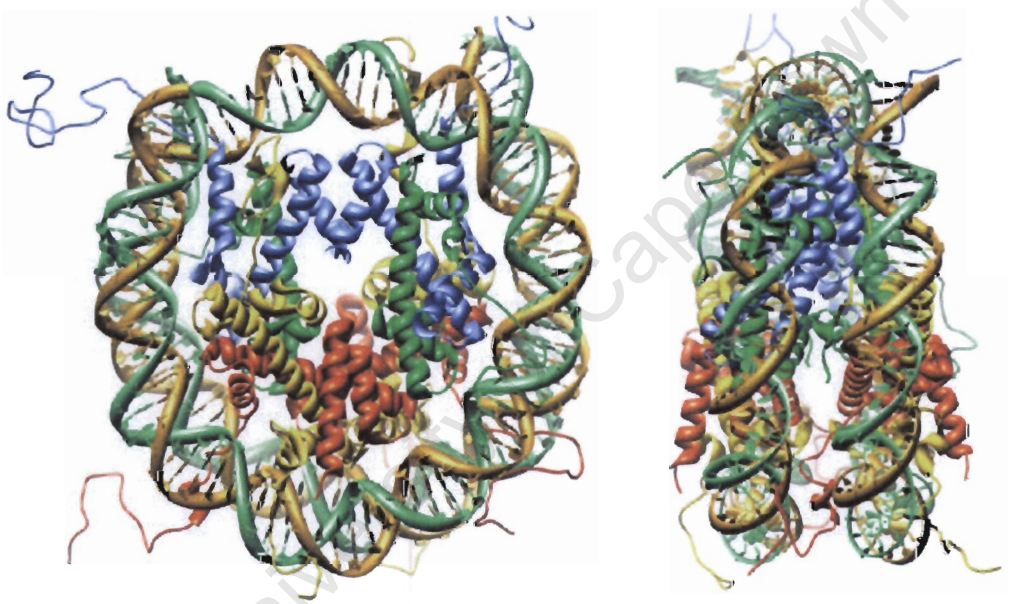


Fig. 1. Ribbon model representing crystal structure of a nucleosome. Histone H3 (blue) forms a heterodimer with histone H4 (green), while histone H2A (yellow) dimerizes with histone H2B (red). H3-H4 dimers assemble into a tetramer through a 4-helix bundle interaction of the α3 helices of the two H3 molecules. The octamer is formed through similar 4-helix bundle interaction between H4 and H2B, linking the two H2A-H2B heterodimers to the tetramer (image created by Tim Frows with UCSF Chimera (Peterson *et al.*, 2004); PDB coordinates are from 1KX5 (Richmond and Davey 2003)).

1.3. Regulation of nucleosome

As the fundamental structural unit, the regulation of nucleosome formation and stability determines the condensation status of chromatin. Two regulatory mechanisms have been intensely researched: the modification of the core histones and chromatin remodeling. Covalent modifications of the core histones alter the affinity of core histone for DNA. Chromatin remodeling complexes are protein complexes that regulate the association of the core histones and other proteins with the chromatin, apparently by mechanical alteration of nucleosome structure.

1.3.1. Covalent modification of core histones

The N-terminal tails of the core histones were clearly shown to be exposed on the outside surface of a nucleosome in the crystal structure of the nucleosome (Luger *et al.*, 1997), and are thus likely to be accessible to complexes involved in post-translational modifications. Intense research effort has revealed many covalent modifications that exert significant effects on the binding stability of a nucleosome and play major roles in the regulation of chromatin status (Jenuwein and Allis, 2001; Strahl and Allis, 2000).

1.3.1.1. Core histone acetylation

The patterns of covalent modifications are distinctively different between euchromatin and heterochromatin. An early correlation made between transcription and histone acetylation (Allfrey *et al.*, 1964) has put histone acetylation under the spotlight for many decades. The subsequent identifications of several histone acetyltransferases

and deacetylases were invaluable to the progress of the current understanding between histone acetylation and transcription activities. It is generally thought that hyperacetylation of histone proteins leads to weakening of the interaction between the DNA molecule and the core histones, through charge neutralization. This loosened interaction could result in the exposure of DNA to other proteins, such as transcription factors or activators. However, Kurdistani and Grunstein (2003) suggested the possibility that acetylation markers serve to create binding surfaces for recruitment of other proteins. In support of this view, studies in *S. cerevisiae* have revealed that heterochromatin protein Sir3p displayed a reduced affinity *in vitro* towards H4 acetylated at Lys 16 (Carmen *et al.*, 2002). *In vivo*, deacetylated H4 is present at Sir3p-bound heterochromatin regions (Braunstein *et al.*, 1996), suggesting that a deacetylated H4 tail provided an appropriate binding site for Sir3p.

Although the mode of action of histone acetylation in transcription is unresolved, the link between histone acetylation and transcription is better defined. Several protein complexes that associate with transcription machinery were found to display histone acetyltransferase (HATs) activity. The yeast TBP-associated factor 130 (also TAF230 in *Drosophila*, Taf250 in humans) subunit of the TFIID complex, a basic component involved in RNA polymerase II (Pol II) basal transcription, is a histone acetyltransferase (Struhl, 1998). The p300/CBP transcription co-activators, which are coupled to the Pol II holoenzyme, also contain histone acetyltransferase activity. Evidence linking core histone acetylation and transcription was also presented by mutation studies. In *S. cerevisiae*, the mutation of Lys 56 of histone H3, which is acetylated by the putative HAT Spt10p, resulted in altered expression of the histone genes (Feng *et al.*, 2005).

On the other hand, histone deacetylase activities correlated strongly with transcription repression. The deacetylation activity of Rpd3p, a yeast protein which complexes with the transcription co-repressor Sin3p and Ume6p, was found to play an important role in transcription repression (Kadosh and Struhl, 1998a). A mutant Rpd3p with abolished deacetylase activity but still capable of forming a complex with Sin3p, was found to be defective in transcription repression. Direct analysis of repressed promoter regions revealed promoter-specific deacetylation of histone H4 and H3 by Rpd3p (Kadosh and Struhl, 1997; Kadosh and Struhl, 1998b). Apart from acting in a localized fashion, a knock out study in *S. cerevisiae* gave an example of histone acetylation and deacetylation acting on a global scale (Vogelauer *et al.*, 2000).

1.3.1.2. Core histone methylation

Another intensively studied histone modification is methylation. Histone methylation was shown to be associated with both transcription activation and repression. *In vivo* studies on the effect of co-activator-associated arginine methyltransferase 1 (CARM1) activity on Arg17 of the H3 N-terminal tail, demonstrated a link between core histone methylation and active transcription of the estrogen-inducible *pS2* gene (Bauer *et al.*, 2002). In *S. cerevisiae*, Set1p H3 Lys 4 methylation was found to be necessary to specifically target the chromatin remodeling protein Isw1p *in vivo* (Santos-Rosa *et al.*, 2003). The same group later reported H3 Lys 4 methylation excluded the binding of silent information regulator Sir3p *in vitro* (Santos-Rosa *et al.*, 2004). These findings suggested that histone methylation promoted euchromatin formation by excluding heterochromatic proteins.

The connection between histone methylation and transcription repression has also been investigated. The identification of human SUV39H1 and murine Suv39h as SET-domain histone methyltransferases, lead to the discovery that H3 Lys 9 methylation served as an important marker for heterochromatin formation (Rea *et al.*, 2000), since H3 Lys 9 methylation prevented acetylation of the same residue. H3 methylation at Lys 9 by SUV39H1 histone methyltransferase also contributed towards heterochromatin formation by acting as binding site for the heterochromatin-associated HP1 protein *in vivo* (Lachner *et al.*, 2001).

Although several histone methyltransferases have been identified in different organisms, a mechanism to revert a methylated residue has been missing. The lack of evidence for a demethylase lead to the thinking that methylation was a permanent event, and that histone methylation could not contribute toward chromatin regulation in a "switch" fashion as other covalent modifications. Efforts by two groups have recently revealed plausible mechanisms that regulate histone methylation. Wang *et al.* (2004) showed that human peptidylarginine deiminase 4 (PAD4) was able to convert methylated arginine into citrulline *in vitro*, in a reaction referred as demethylimination. *In vivo*, PAD4 was shown to be directly responsible for a decrease in methylated arginine level in histone H3 and H4, and an increase in citrulline level, correlated to the decrease in methylation. In a similar study by Cuthbert *et al.* (2004), PAD4 demethylimination of arginine on histone H3 tail was found to prevent methylation of CARM1. Furthermore, both studies have shown that PAD4 had a repressive effect on hormone inducible *pS2* promoter *in vivo*. Together, these results presented a novel transcription regulatory mechanism through coupled histone methylation/demethylimination.

1.3.1.3. Core histone Phosphorylation

Phosphorylation of core histones has been observed to function in a concerted fashion with other covalent modifications. In *Drosophila*, H3 phosphorylation at Ser 10 is coupled with acetylation of Lys 14 on histone H4 during the equalization of transcription from the sex chromosomes, where the male X chromosome was reported to exhibit a two-fold increase in its transcription (Lucchesi, 1998). In addition, phosphorylation of the histone H2B Ser 33 residue by the carboxyl-terminal kinase domain of *Drosophila* TAFI, correlated with transcriptional activation that promotes cell cycle progression and development in *Drosophila* cells and embryo (Maile *et al.*, 2004). A study in *Tetrahymena thermophila* by Wei *et al.* (1999) presented a different picture for H3 Ser 10 phosphorylation in the regulation of chromatin structure. Through a mutagenesis experiment, Wei *et al.* showed that phosphorylation of H3 Ser 10 was required for proper condensation in the micronucleus. Chromosome loss and defects in chromosome segregation was observed in the micronuclei of mutants lacking phosphorylated histone H3 during mitosis and meiosis, respectively.

1.3.1.4. Core histone ubiquitylation

A covalent modification with less well-defined role in nucleosome dynamics is ubiquitylation. Studies on histone ubiquitylation showed a linkage between histone methylation and ubiquitylation in *S. cerevisiae*. Sun and Allis showed that disruption of Rad6p mediated ubiquitylation of H2B Lys 123 led to the abolishment of H3 Lys 4 methylation (Sun and Allis, 2002). Ubiquitylation of H2B tails thus seemed to provide an additional level of control to the regulation of other histone modifications. Further support for this view was provided by another investigation on the effect of an

ubiquitin-specific protease (Ubp10p) on gene silencing. Ubp10p targets and de-ubiquitylates histone H2B at telomeric regions, and is associated with a reduced level of H3 Lys 4 methylation. Disruption of *UBP10* resulted in increased H3 Lys 4 and Lys 79 methylation, as well as reduced association of Sir2p at telomeric regions (Emre *et al.*, 2005). This finding demonstrated the linkage between deubiquitylation by Ubp10p and heterochromatin maintenance at the telomeric regions.

1.3.1.5. Globular domain covalent modifications

Apart from covalent modifications of the histone tails, high resolution studies using mass spectroscopy techniques revealed a number of covalent modifications that occurred in the globular domains of the core histones (Zhang *et al.*, 2003). Several of these newly identified histone modifications were positioned on the surface of the histone globular domain, placing them in close proximity to the DNA molecule, which led to the suggestion that these modifications could directly regulate the interaction between DNA and the histones (Cosgrove *et al.*, 2004).

1.3.1.6. The "histone code" hypothesis

These covalent modifications of the core histones work in a concerted manner, and display interdependence to each other. The correlation between different patterns of histone modifications, associated with different chromatin regions, led to the "histone code" hypothesis. Interplay, such as the coupled H3 Ser 10 phosphorylation and H3 Lys 14 acetylation, the mutual exclusion between H3 Lys 9 methylation and H3 Ser 10 phosphorylation, and the dependence of H3 Lys 4 methylation on H2B Lys 123

ubiquitylation, were just some of the interactions that provided clues towards the existence of a "histone code" (Jenuwein and Allis, 2001; Margueron *et al.*, 2005; Strahl and Allis, 2000).

1.3.2. Chromatin remodeling complexes

In addition to the observed post-translational modifications of the core histones, nucleosome mobility is also affected through the involvement of protein complexes that utilize ATP as energy source to alter chromatin structure in a non-covalent fashion.

1.3.2.1. Different chromatin remodeling complex groups

ATP-dependent chromatin remodeling complexes can be grouped according to the ATPase component of the complex (Vignali *et al.*, 2000). The SWI2/SNF2 group of protein complexes, such as yeast SWI/SNF, yeast RSC, *Drosophila* Brahma complex and human SWI/SNF, share additional homology of a bromodomain in the C-terminus and two other conserved domains apart from the ATPase domain of the Swi2/Snf2 protein subunit. The ISWI group of chromatin remodeling complexes includes ATP-utilizing chromatin assembly and remodeling factor (ACF), nucleosome remodeling factor (NURF) and chromatin assembly complex (CHRAC). A third group of chromatin remodeling complexes, the Mi-2 group, contains similar complexes that display both chromatin remodeling and histone deacetylase activities.

1.3.2.2. Functioning of chromatin remodeling complexes

The identification of similar motifs in the NS3 RNA helicase and the different chromatin remodeling complexes resulted in the postulation that ATP-dependent chromatin remodelers act by altering local DNA topology (Flaus and Owen-Hughes, 2001). However, different chromatin remodeling complexes display different substrate affinities. While the SWI/SNF group is stimulated by both naked DNA and nucleosomes, the activity of ISWI and Mi-2 groups requires intact nucleosomes. The ISWI group further requires intact histone tails for both ATPase and nucleosome mobilization activities (Boyer *et al.*, 2000; Brehm *et al.*, 2000).

The best characterised interaction between ATPase-dependent chromatin remodeling and transcription comes from the study on the yeast *HO* gene (Krebs *et al.*, 1999). The chromatin immunoprecipitation study revealed the recruitment of SWI/SNF by Swi5 is necessary for the activity of *GCN5*-dependent HAT. This SWI/SNF dependent recruitment of *GCN5*-dependent HAT was observed at many genes expressed during late mitosis (Krebs *et al.*, 2000). Krebs *et al.* proposed a model where SWI/SNF dependent recruitment of *GCN5*-dependent HAT regulates late mitotic gene expression in a global fashion through disruption of condensed chromatin, hence increasing the accessibility of histone tails for stable *GCN5*-dependent HAT association.

1.4. Higher order chromatin structures

The formation of nucleosomes is the basic level of chromatin condensation. Early electron microscopic studies revealed condensation of nucleosomal arrays at increased ionic strengths. At low ionic concentration, the nucleosome array did not

condense into a defined structure. However, in the presence of linker histone H1, a non-essential histone protein that does not form part of the nucleosomal octamer, resulted in the formation of a more condensed structure, termed the 30nm fibre (Finch and Klug, 1976; Thoma *et al.*, 1979).

1.5. Linker histone

The linker histone family of proteins is generally larger than the core histones, and are not involved in the formation of nucleosome cores. The association of a linker histone with a core leads to the formation of a chromatasome. Evidence of linker histone association with nucleosome cores was provided by nuclease footprinting studies where an additional 20bp of DNA was protected from micrococcal nuclease (MNase) cleavage upon the addition of linker histone (Simpson, 1978; Wolffe, 1998).

1.5.1. Linker histone diversity

Unlike the core histones, linker histones are poorly conserved between organisms. In vertebrates, different variants of linker histones have been found to be associated with different developmental stages (Khochbin and Wolffe, 1994). The *X. laevis* linker histone B4 is a classical example of a cleavage stage linker histone present in early embryogenesis, and the linker histone H5 is specifically found in terminally differentiated avian erythrocytes.

The structure of the histone H1 family is also poorly conserved amongst different organisms. The generic structure of metazoan linker histone H1 is a short N-terminus tail, a central globular domain, and a lysine rich C-terminus tail (Harvey and Downs,

2004). The putative linker histone homolog in *Saccharomyces cerevisiae* contains two homologous globular domains joined by a short, lysine-rich inter-globular region. Another example of the structural diversity of the linker histones is provided by that of *Tetrahymena*, where both the macronuclear and micronuclear linker histone contains no obvious globular region (Hayashi *et al.*, 1987; Wu *et al.*, 1986).

1.5.2. Functioning of linker histone

It is generally accepted that the linker histones contribute toward packaging of DNA by partially neutralising the negative charge of the linker DNA regions between the nucleosomes with the highly positively charged C-terminal tail. The importance of the highly charged C-terminal for linker histone binding was demonstrated *in vivo* study (Hendzel *et al.*, 2004). However, chromatin condensation was shown to also be dependent on the core histone N-terminal tails (Carruthers and Hansen, 2000). It was suggested that the linker histones play a supportive role in higher order structure formation, stabilizing condensed chromatin. In addition to electrostatic neutralisation, it was also postulated that the linker histones may be involved in direct protein-protein interactions (Hayes and Hansen, 2001).

1.6. Chromatin dynamics in transcription regulation

1.6.1. Linker histone is not a global repressor of transcription

The condensation of chromatin, especially in the promoter regions, prevents the access of transcription activators to their targets. Studies into nucleosome dynamics have contributed to our understanding of the regulation of transcription at the

nucleosome level. The influence of higher-order structure, particularly the regulation of chromatin compaction on gene expression is poorly understood. Clues on the contribution of the higher-order structure to transcription regulation have been provided by the effect of linker histones on transcription. Linker histones were originally thought of as a global repressor of transcription (Croston *et al.*, 1991). However, a later study clearly showed that linker histones participated in localized transcription regulation. In the protozoan *Tetrahymena thermophila*, removal of linker histone was not lethal and did not result in any phenotype, other than enlarged nuclei (Shen *et al.*, 1995). Examination of transcription revealed that general transcription was not affected by the absence of histone H1. However, the expression of a starvation-induced gene, *Cyp*, decreased about 10-fold in a histone H1 knockout mutant (Shen and Gorovsky, 1996). Paradoxically, absence of histone H1 also resulted in an increased basal expression of another stationary phase gene, *ngoA*, in exponential phase cells. These findings demonstrate linker histones act on a gene-specific level, and can be involved both in repression and induction, rather than simply performing a global repressive role. It has been reported that *S. cerevisiae* lacking the linker histone, Hho1p, did not show a global up-regulation of transcription. Instead, a small amount of genes displayed a minor reduction in their transcription level in the absence of Hho1p (Hellauer *et al.*, 2001).

1.6.2. Linker histone, nucleosome positioning and transcription

As a general architectural component of the chromatin, linker histones must work in concert with other elements to exert transcription regulatory effect in a specific manner. In the *Xenopus laevis* embryo, the oocyte and somatic 5S rDNA genes are differentially expressed. The transition of the expression from the oocyte 5S rDNA

gene to the somatic type correlates with the replacement of the cleavage stage linker histone B4 with the somatic histone H1. The somatic histone H1 selectively blocks transcription factor TFIIIA from binding to the oocyte 5S rDNA gene, but not the somatic 5S rDNA gene (Sera and Wolffe, 1998). The association of somatic histone H1 with the oocyte 5S rDNA gene positions the nucleosome such that all elements of the promoter was sequestered within the nucleosome, while histone H1 association at the somatic 5S rDNA gene repositioned the nucleosome such that the promoter region was still available to transcription factor binding. Based on the equal affinity of histone H1 for both the 5S rDNA genes, the authors concluded that the selective repression of oocyte 5S rDNA was dependent on the positioning of the nucleosome, mediated by histone H1, to specific nucleotide sequences. The AT-rich sequence at the 3' end of the oocyte 5S rDNA gene was a key determinant in the positioning of the nucleosome, and hence the selective repression. The selective repression of the oocyte 5S rDNA presents a model of how a linker histone was able to regulate transcription through the assembly of particular chromatin structures.

However, another *in vitro* study employing nucleosome mapping of the *Xenopus* oocyte and somatic 5S rDNA presented a different model for the mechanism of selective repression of the oocyte 5S rDNA by histone H1. Nucleosome positioning at the two genes was found to be significantly different. The authors proposed a different model that both H1 and TFIIIA binding was dependent on the nucleosome position, and that TFIIIA binding out-competed H1 binding at the somatic type gene, while H1 preferentially bound to the nucleosomes at the oocyte type gene. The binding of each element would exclude the binding of the other (Panetta *et al.*, 1998).

Although different models for the mechanism have been proposed, both models highlight the role of chromatin dynamics, in the form of nucleosome positioning, in the regulation of gene expression. A knockout study in mice has provided further indirect evidence supporting the importance of histone H1 mediated chromatin dynamics in gene expression regulation. After sequential inactivation of three linker histone H1 subtypes (H1c, H1d and H1e), mutant mice embryos suffered a broad range of defects, indicating that histone H1 was essential for normal mouse development (Fan *et al.*, 2003). The level of H1 in triple-knockout mutant embryos was found to be 50% of the normal level. Examination of thymus nuclei of the triple-knockout mutant did not reveal enlarged nuclei compared to wild-type. However, a reduction in nucleosome repeat length was observed in the same study, and the magnitude of the reduction appeared related to the level of H1 present.

1.6.3. Linker histone induced synergism activates transcription

The stabilizing effect of histone H1 on promoter nucleosomes could also affect transcription regulation by influencing the affinity of activators for their targets. The mouse mammary tumor virus (MMTV) promoter is activated through the synergism between progesterone receptor (PR) and nuclear factor1 (NF1), where the binding of PR to hormone responsive elements and recruitment of ATP-dependent chromatin remodeling complexes, aids the binding of NF1. The association of NF1 to the promoter enhanced the binding of other factors by maintaining an open chromatin structure, which resulted in the full activation of the promoter (Croce *et al.*, 1999). The association of H1 repressed the basal activity as well as PR or NF1 activation of the promoter. However, simultaneous presence of PR and NF1 at H1 bound chromatin of MMTV promoter resulted in increased synergism of PR and NF1, as well as

transcription from the MMTV promoter. Analysis of the translation and rotational setting of the promoter nucleosomes revealed that in the presence of H1, a higher portion of the nucleosomes adopted a dominant position, suggesting the stabilizing effect of H1 on a particular setting of the promoter nucleosomes (Koop *et al.*, 2003).

1.6.4. Chromatin compaction by non-histone proteins

Apart from linker histone H1, several other proteins have been reported to be capable of regulating transcription through formation of distinct chromatin structures. The human MeCP2 is a DNA-binding protein that can selectively bind methylated CpG dinucleotides. It has been proposed that MeCP2 acted as a specific DNA methylation-dependent transcriptional repressor. The MeCP2 was shown to bind reconstituted nucleosomal arrays and induce extensive compaction of the nucleosome arrays into condensed secondary structures. This MeCP2 mediated compaction was also shown to be independent of salt presence (Georgel *et al.*, 2003). In *Drosophila melanogaster*, Polycomb Repressive Complex 1 (PRC1) of the Polycomb group (PcG) proteins was able to induce similar extensive compaction of nucleosomal arrays *in vitro*. The induced compaction was found to be dependent on PRC1 core complex-nucleosome interaction, but the histone tails were not essential (Francis *et al.*, 2004). The ability of MeCP2 and PRC1 to mediate extensive compaction of chromatin suggests possible alternative agents that can influence transcription by acting as chromatin architectural components.

1.6.5. Linking chromatin compaction and transcription activity

The investigation of global chromatin architecture has also yielded valuable insight into the relationship between chromatin dynamics and transcription regulation. Chromatin of different conformations can be separated according to density. Based on this method, Gilbert *et al.* (2004) was able to analyze the distribution of condensed and decondensed chromatin on a genome-wide scale. The authors discovered that open chromatin fibres included high gene density regions, but no correlation was found between open chromatin and gene expression, as inactive genes could also be present at gene-dense regions, while an active gene could be situated in a low gene density region within a condensed chromatin fibre.

1.7. Project Aim

The picture presented by previous investigation into the relationship between chromatin dynamics and transcription regulation is an intriguing one. The influence of general structural components of chromatin architecture on transcription activity has been observed on many occasions, thus unravelling the tie between chromatin dynamics and transcription activities would be a huge advancement in the acquisition of the complete understanding of gene expression regulation. In order to facilitate investigations into chromatin state and its correlation to transcription, an efficient molecular tool would be invaluable. A recombinant protein capable of detecting chromatin compaction status, without any treatments that might disrupt the native chromatin structure, is desired. This study proposes the design and cloning of such a recombinant protein.

Previous studies within the laboratory have shown that the yeast linker histone, Hho1p, is associated with the chromatin during stationary phase growth, and the dissociation of Hho1p occurs in a random fashion (Schafer and Patterton, unpublished data). The involvement of Hho1p in regulating gene expression of cells exiting stationary phase appears to be absent. In attempt to elucidate a possible role of Hho1p in transcription regulation, this study also aims to identify specific regions of interest in the *S. cerevisiae* genome in terms of chromatin status correlated to supposed gene expression activity, employing a similar strategy as Gilbert *et al*. The recombinant protein would then offer an opportunity to gaze into the local chromatin compaction status at high resolution.

Chapter 2

Materials and methods

2.1. Yeast culture strains and culturing conditions

The *Saccharomyces cerevisiae* W303 strain (*MATa/MATaADE2/ade2 CAN1/can1-100 CYH2/cyh2 his3-11,15/his3-11,15 LEU1/leu1-c LEU2/leu2-3,112 trp1-1:URA3:trp1-3Δ /trp1-1 ura3-1/ura3-1*) was used in this study for the isolation of soluble chromatin. Pre-inoculums were prepared by inoculating a single colony into 5ml YPD [1% (w/v) yeast extract, 2% (w/v) peptone, 2% (w/v) D-(+) glucose] medium and grown overnight at 30°C with orbital shaking (200rpm). All cultures were inoculated with a 1ml volume of pre-inoculum, and grown at 30°C with orbital shaking (200rpm).

2.2. Phenol/chloroform DNA purification

Contaminating protein was removed from DNA samples by addition of an equal volume of phenol/chloroform/isoamylalcohol (25:24:1 [v:v:v]). The mixture was vortexed using a Vortex Genie (Scientific industries, Inc.) at maximum speed for 10min. The mixture was then centrifuged at 16100g for 10min using a bench-top micro-centrifuge (Eppendorf). The aqueous phase was transferred into the same volume of chloroform/isoamylalcohol (24:1 [v:v]) solution, and vortexed for 5 min at maximum speed. The mixture was centrifuged at 16100g for 5 min and the aqueous phase collected.

2. 3. Ethanol precipitation

A volume of ice-cold absolute ethanol (-20 C), two times that of the sample volume, as well as a volume of 3M sodium acetate, a tenth to that of the final volume, was added to a DNA sample. The mixture was incubated at -70°C for at least 30min, followed by centrifugation at 4 C for 10min at full speed on a bench-top micro-centrifuge. The supernatant was decanted and the pellet was resuspended in 70% (v/v) ice-cold (-20°C) ethanol. The mixture was incubated at -20°C for 10min and the DNA pellet collected by centrifugation at 4°C for 10min at full speed on a micro-centrifuge. The supernatant was discarded and the DNA pellet dried using a SpeediVac rotatory evaporator (Sorvall).

2.4. Agarose gel electrophoresis

Standard agarose (Sigma) was dissolved in 1 × TAE (40mM Tris pH8.3, 20mM acetic acid, 1mM EDTA) for all agarose gels. DNA samples were mixed with 1/6 volume of 6 × loading buffer (Promega) before applying the sample to the gel. Electrophoresis took place in 1 × TAE buffer at 100V for at least 30min. Agarose gels were stained in ethidium bromide (0.75µg/ml) for 10min, and de-stained in distilled water for 5 min. The DNA in the stained agarose gel was visualised with an UV transilluminator..

2.5. Genomic screening of distribution of genes within condensed chromatin

2.5.1. Yeast chromatin preparation

Exponential phase and stationary phase culture of *S .cerevisiae* were isolated after cultures reached an OD_{600} of approximately 0.5 or a culturing time of 6 days, respectively. Volumes of 5 l and 400ml of liquid YPD medium were used for culturing exponential phase and stationary phase yeast, respectively, since these volumes contained approximately equal cell densities after growth for the stated periods. Cells were collected by centrifugation at 3000g using a JA-17 rotor (Beckman) for 5min and washed twice in S buffer (1.4M sorbitol, 40mM HEPES, pH 7.5, 0.5mM $MgCl_2$, 0.5% (v/w) β-mercaptoethanol, 1mM PMSF).

The cells were resuspended in S buffer and incubated at 30°C for 10min. Treated cells were pelleted in pre-weighed 250ml capacity polypropylene centrifuge tubes (Beckman) at 3000g, in a JA-17 rotor (Beckman), and the wet weight of the pellet noted. The pellet was resuspended in S buffer containing 5 mM DTT (Fluka) and 20mg of zymolyase T100 (Sigma), in a volume (in millilitres) of 4 times the wet weight of the pellet (in grams). Spheroplasting was allowed to proceed at 30°C for 60min. Spheroplasts were collected by centrifugation using a JA-17 rotor at 2000g for 5min at 4°C, and washed twice in ice-cold (4°C) S buffer. The spheroplasts were resuspended in a volume (in millilitres) of ice-cold (4°C) F buffer (18% [w/v] Ficoll 400, 20mM PIPES, pH6.5, 0.5mM $MgCl_2$, 1mM PMSF) that was three times the wet weight of the pellet (in grams). The resuspended spheroplasts were lysed with at least 10 passes of a tight dounce on ice. The lysate was layered onto an equal volume of GF buffer (20% [w/v] glycerol, 7% [w/v] Ficoll 400, 20mM PIPES, pH 6.5, 0.5mM

$MgCl_2$, 1mM PMSF) and spun in a JA-20 rotor (Beckman) for 30min at 20000g at 4°C. Crude nuclei were resuspended in one volume of F buffer and vortexed at 4°C for 5min. Cell debris was removed by centrifugation at 3000g for 15min at 4°C using a JA-20 rotor, and the supernatant collected. Nuclei were pelleted by centrifugation of the supernatant at 20000g for 25min at 4°C in a JA-20 rotor.

The nuclear pellet was resuspended in 2ml of D buffer (10mM HEPES, pH 7.5, 0.5mM $MgCl_2$, 0.1mM $CaCl_2$, 1mM PMSF) containing 100µg/ml RNAseA (Sigma) and incubated at 37°C for 15min. Chromatin was digested by addition of 10 U/ml micrococcal nuclease (MNase; Sigma). Digestion was allowed to proceed for 15 min and stopped by addition of 1/10 volume ice-cold (4°C) 100mM EDTA, and chilled on ice. The nuclei containing digested chromatin were pelleted using a bench-top micro-centrifuge at 14000g for 5min. Digested chromatin was solubilized by resuspending the pellet in 2ml of Y-PER reagent (Pierce), and incubating at room temperature for 30min with gentle agitation. The soluble chromatin was collected as the supernatant fraction after centrifugation at 12000g for 2 min using a bench-top micro-centrifuge.

2.5.2. Sucrose gradient ultracentrifugation

The isolated soluble chromatin was immediately loaded onto a 6-40% (w/v) continuous sucrose gradient containing $TEEP_{80}$ (10mM Tris, pH7.6, 0.1mM EDTA, 0.1mM EGTA, 0.25mM PMSF, 80mM NaCl). The sample was centrifuged at 83000g for 7 hrs at 4°C using a SW28 rotor (Beckman). Fractions were collected (1ml aliquots) by upward displacement using 50% (w/v) sucrose solution in $TEEP_{80}$ buffer. The collected aliquots as well as a sample of the total input soluble chromatin were

digested overnight with proteinase K (1 mg/ml) in 1% (w/v) SDS, and further purified by twice extracting with phenol/chloroform. DNA was collected by ethanol precipitation and resuspended in TE (10mM Tris, pH 8.0, 1mM EDTA) buffer. The DNA complement from the different sucrose fractions were analysed on 1% (w/v) agarose gels. Gradient fractions encompassing a range of DNA fragments (<1kb to >10kb) were identified.

2.5.3. Purification of condensed and decondensed DNA

Selected gradient fractions containing best representation of differently sized DNA fragments were separated on a 0.7% (w/v) agarose gel. The regions containing DNA fragments of size 10kb ±1kb (designated "10kb" in the text) and 4kb ±1kb (designated "4kb" in the text) were excised on a blue fluorescence light box that was able to induce fluorescence of ethidium bromide chelated to DNA without the damaging effects of a UV source, and DNA was further purified with a QiaQuick Gel Purification Kit (Qiagen).

2.5.4. Quantitative PCR

The concentration of the DNA fragments containing the *DNF1*, *ROX1* and *SNZ1* genes was quantified by quantitative PCR in the sucrose gradient input samples as well as the samples recovered from the sucrose gradient following gel purification. The standard curves were constructed using the PCR products of the three genes at concentrations that ranged from 10ng/µl to 0.1pg/µl. Reactions were optimised using SensiMix(dT) (Quantace Ltd.) and fluorescence probe SYBR® Green I was added following a protocol provided by the manufacturer. The reaction mixes were first

incubated at 37°C for 10min, followed by 10min enzyme activation at 95°C. The cycling parameters were set to 95°C for 15sec, 50°C for 30sec, and 72°C for 30sec. The kinetic PCR procedure was performed on a RG3000 instrument (Corbett Research) and analysed using Rotor-Gene 6 software (Corbett Research). Quantifications for DNA samples were performed in triplicate for all three genes.

2.5.5. Microarray probe generation

DNA from the total input and 4kb region of the fractionated sucrose gradient samples were labeled with Cy3 and Cy5 ester dye respectively using an indirect labelling method. Aminoallyl-dUTP was incorporated into probes through a random priming PCR method. The reaction mix for the first round of random amplification consisted of 3µl DNA template (10ng), 2µM of primer A (5'-GTTTCCCAGTCACGATCNNNNNNNNN-3'), 75µg/µl of BSA, 0.3mM dNTP mix, 1× Super-Therm polymerase reaction buffer (Roche Applied Science) and 2mM $MgCl_2$. Ten units of Super-Therm Taq polymerase (Roche Applied Science) was added to a final reaction volume of 20µl. The reaction mix was incubated at 94°C for 2min and rapidly cooled to 10°C on a PCR thermocycler (Eppendorf). Primer annealing was achieved by incubating the reaction mix at 10°C for 5min and slowly raising the temperature to 37°C over an 8min period. The temperature was held at 37°C for 30min before repeating the cycle. The PCR product was purified by a QIAquick PCR Purification Kit (Qiagen) and diluted to volume of 60µl. The reaction mix for the second round amplification consisted of 15µl of the first round PCR product, 1× Super-therm polymerase reaction buffer, 2mM $MgCl_2$, 1mM modified dNTP mix (aminoallyl-dUTP:dTTP [mole:mole] of 3:2) and 1µM of primer B (5'-GTTTCCCAGTCACGATC-3'). Nuclease free water (Qiagen) and 5U of

Super-therm Taq polymerase was added to a final reaction volume of 100µl. The reaction mix was incubated at 94°C for 2min and then cycled 35 times at 94°C for 30sec, 50°C for 30sec and 72°C for 2min (DeRisi *et al.*, 1997). The PCR products were purified by using a QIAquick PCR Purification Kit (Qiagen) and eluted with two aliquots of 30µl of nuclease-free water. The concentration of DNA was determined using a NanoDrop-1000 spectrophotometer (NanoDrop Technologies).

Approximately 2µg of modified dNTP incorporated probes was used in each coupling reaction. Input reference probe DNA was labeled with Cy3 dye (Amersham Pharmacia) and 4kb region DNA probe was labeled with Cy5 dye (Amersham Pharmacia). The probes were pre-treated with 5µl of 0.1M sodium carbonate buffer, pH 9.0 and incubated at 37°C for 10min. The ester Cy dyes were added in excess for each labelling reaction. The reaction mixture was incubated at room temperature in the dark for one hour, with mixing every 15min. The unincorporated ester Cy dyes were removed by using a QIAquick PCR Purification Kit, and eluted with two volumes of 30µl of nuclease free water (Qiagen). The fluorescence value of labeled probe DNA was determined using a NanoDrop-1000 spectrophotometer. The Cy3 labeled probes were measured at 550nm and Cy5 labeled probes were measured at 650nm. The frequency of incorporation of dye esters was calculated for each dye coupling reaction (Hegde *et al.*, 2000). The frequency of incorporation is equal to pmol of dye divided by pmol of DNA, which was calculated using the following equations:

$$\text{pmol of DNA} = \frac{2 \times \text{DNA (ng)}}{324.5}$$

$$\text{pmol of Cy3} = \frac{\text{Abs550} \times \text{volume (µl)}}{0.15} \qquad \text{pmol of Cy5} = \frac{\text{Abs650} \times \text{volume (µl)}}{0.25}$$

2.5.6. Microarray hybridization and scanning

Cy3 and Cy5 labeled probes were combined and dried down to approximately 5µl using a SpeediVac rotary evaporator (Sorvall). The hybridization mix containing 52µl DIG Easy Hyb solution (Roche Applied Science), 2µl of yeast tRNA (10mg/ml[Life Technologies]) and 1µl mouse COT1 DNA[10mg/ml(Life Technologies)] was incubated at 95°C for 2min followed by the addition of the combined probes. The final hybridization mix was again incubated at 95°C for 2min and then directly applied to microarray slides spotted with intergenic yeast DNA (Bradley Cairns, University of Utah). Hybridization was allowed to occur in the dark for 16 to 18 hours at 37°C, under sufficient humidity. After the hybridization, the microarray slides were first washed in the dark using low stringency buffer [2× SSC (300mM NaCl, 30mM sodium citrate), 0.5% (w/v) SDS] at 42°C with gentle shaking for 4min. The microarray slides were transferred into medium stringency buffer (0.5× SSC) and washed for 4min at room temperature with gentle shaking. The microarray slides were then transferred into high stringency buffer (0.05× SSC) and washed for 4min at room temperature with gentle shaking. The slides were dipped into absolute ethanol five times and then dried by centrifugation at 1000rpm using a swing-out rotor (Sorvall) (Hegde *et al.*, 2000). Slides were scanned on a GenePix 4000A (Axon Instruments) and initial quantification performed using GenePix Pro.6.0 software (Axon Instruments).

2.5.7. Microarray normalization method

The raw fluorescence signals of the duplicate spots on each microarray slide were averaged. In cases where the standard deviation of the values was greater than 50% of

the mean value, the average value was tagged "false", but not excluded for downstream analysis. The mean values for the *DNF1*, *ROX1* and *SNZ1* genes were determined and the average of the values for the three genes taken as the "slide representative average". Slide representative averages were determined for both Cy3 (input reference) and Cy5 (4kb) labeled DNA hybridized to a same slide. The average abundance (in pg/µl) of the *DNF1*, *ROX1* and *SNZ1* genes in the 4kb region DNA sample and input reference were determined by quantitative PCR (see above).. The ratio between the average for the three genes on the microarray slide and the average value for the three genes determined in the same samples by quantitative PCR was defined as the normalization factor for the slide. The background corrected values of all genes on the microarray slide were adjusted with this normalization factor. This ensured that the values analysed for the microarrays were related to the real concentrations of the genes in the original samples used for microarray slide probe preparation, and also allowed direct comparison between microarray slides, since all slides were adjusted to the real DNA concentrations. The normalisation method makes the assumption that the microarray values of *DNF1*, *ROX1* and *SNZ1*were correct in proportion to the rest of the microarray values. The ratio between the normalized microarray values of the 4kb region DNA and the input reference DNA was calculated for each slide. This ratio was independently determined using triplicate slides. The average of the normalized mean of the triplicate was used for comparison of the compaction state of chromatin in stationary phase and in exponential growth phase..

2.6. Construction of a recombinant MNase-LexA fusion protein

2.6.1. Bacteria strain and culturing conditions

The bacteria *Escherichia coli DH5α* strain (F^-, φ80d*lacZ*ΔM15, Δ(*lacZYA-argF*)U169, *deoR*, *recA*1, *endA*1, *hsdR*17(rk^-, mk^+), *phoA*, *supE*44, λ^-, *thi*-1, *gyrA*96, *relA*1) was used for all cloning experiments in this study. The cultures were grown in 5ml liquid LB [1% (w/v) tryptone, 0.5% (w/v) yeast extract, 0.5% (w/v) NaCl] medium at 37°C with orbital shaking (150rpm).

2.6.2. Plasmids and oligonucleotide primers

The construction of recombinant fusion protein was carried out using pYES2 (Invitrogen) as host vector. The final fragment encoding the fusion protein was inserted into plasmid pET20b-(+) (Novagen). All plasmid extractions and purifications were performed using the QIASpin Mini Prep Kit (Qiagen). The oligonucleotide primers used for the construction of recombinant fusion protein are listed in Table 2.1.

2.6.3. Polymerase chain reactions

The reaction mixes consisted of 100ng of DNA template (50ng for plasmid DNA), 1× *Pfu* DNA polymerase buffer (Promega), 0.5mM dNTP mix, 1μM of each of the forward and reverse primers, and 5U of *Pfu* DNA polymerase (Promega). The final reaction volumes were 50 μl, made up with nuclease-free water (Qiagen). The reaction was incubated at 95°C for 5min, followed by 25 cycles of 95°C for 30sec,

50°C for 1min, 72°C for 2min. At the end of the 25 cycles the reaction was incubated for 7min at 72°C to allow strand completion.

Table 2.1. Sequence of oligonucleotide primers used in this study

Primer	Sequence [a]
lacUV5-F	5′-GGGGTCTAGAAAATTATTGTCTAAC-3′
lacUV5-R	5′- GGGGGAATTCCGTCACTCATTAGGC-3′
lexA-F	5′- GGGGGGATCCATGAAAGCGTTAACG -3′
lexA-R	5′- GGGGTCTAGATTACAGCCAGTCGCCG-3′
MNase-F	5′-GGGGGAGCTCATGCCAAAGAAGAGAAAGGTTCCCGGGTTAATTAACGC -3′
MNase-R	5′- GGGGGGATCCTTGACCTGAATCAGC-3′
SEQ-R	5′- ACATGATGCGGCCCTC-3′
MNL-F	5′- GGGGCATATGATGCCAAAGAAGAG-3′
MNL-R	5′- GGGGCTCGAGCGTCACTCATTAGG-3′
Mut-F	5′- CGACTGGCTGAATCTAGAAA-3′
Mut-R	5′- TTTCTAGATTCAGCCAGTCG-3′

[a] Underlined sequences represent restrictions sites introduced by the oligonucleotide primers.

2.6.4. Cloning of *lexA – lacUV5* fragment

The *lac-UV5* promoter was amplified from pSUN1 (Wang and Simpson, 2001), using oligonucleotides lacUV5-F and lacUV5-R primers. This primer set introduced an XbaI restriction site at the 5' end and an EcoRI restriction site at the 3'end of the amplified product. The PCR product was purified by direct loading of the reaction mix onto a 12% non-denaturing polyacrylamide gel (29:1 [w:w] acrylamide:bisacrylamide) and electrophoresed at 100V for 50min in 1× TAE buffer. The electrophoresed DNA was visualised on a UV transilluminator following staining in 0.75µg/ml ethidium bromide. The band corresponding to the approximately 100bp *lac-UV5* promoter was excised from the gel, and the gel slice cut into approximate $1mm^3$ pieces. Gel fragments were placed in two volume of elution buffer (0.5M ammonium acetate, 1mM EDTA, pH 8.0) and incubated overnight at 37°C with agitation. The elution mixture was collected by centrifugation, and the DNA pelleted by an ethanol precipitation. The *lac-UV5* promoter DNA fragment was resuspended in distilled water and stored at -20°C.

The ORF encoding the LexA protein was amplified from pLexA (a kind gift of Takashi Ito, University of Tokyo) by PCR with primer LexA-F and Lex-R. The primers introduced a BamHI restriction site at the 5'end and an XbaI restriction site at the 3'end of *lexA* amplified fragment. The desired *lexA* DNA fragment was purified by agarose gel electrophoresis. The desired band of size corresponding to that of *lexA* sequence was excised from the gel and purified using a QIAquick Gel Extraction Kit (Qiagen). The *lac-UV5* and *lexA* DNA fragments were digested with XbaI (Promega) using the condition recommended by the supplier, and purified using a QIAquick PCR Purification Kit (Qiagen).

The XbaI digested *lac-UV5* DNA and *lexA* DNA fragments were ligated in a 10µl reaction containing equimolar amounts of the two fragments and 5U of T4 DNA ligase (Promega) at 4°C for 16 hours. The ligated *lexA-lac-UV5* fragment was purified by ethanol precipitation, and resuspended in distilled water. The ligated *lexA-lac-UV5* fragment was amplified by PCR with primers LexA F and lacUV5R and more than 2µg of the fragment was acquired. The amplified *lexA-lac-UV5* fragment was digested first with BamHI (New England Biolab.), followed by EcoRI (Promega) using conditions recommended by the manufacturers. The host vector pYES2 (Invitrogen) was digested with identical enzymes, and the BamHI/EcoRI double-digested *lexA-lac-UV5* fragment was ligated into the pYES2 vector in a 10µl reaction with 5U of T4 DNA ligase (Promega), at 4°C overnight. The amount of vector and insert fragments was adjusted to achieve a vector to insert ratio of 1:5. The ligation mixture was transformed into *E.coli DH5α* strain with a MicroPulser Electroporation Apparatus (Bio-Rad) and cultured on LB agar medium containing 100µg/ml ampicillin. Positive clones containing the correct *lexA-lac-UV5* insert was identified with restriction enzyme digest and by partial nucleotide sequencing method. The recovered positive plasmid was denoted pYES2-L/L and the strain stored as a glycerol stock at -70°C.

2.6.5. Cloning of the SV40 nuclear localization signal and the active N-domain of micrococcal nuclease (MNase)

A 21bp nuclear localization signal (5'-ATGCCAAAGAAGAGAAAGGTT-3') was included in primer MNase-F (Table 2.1). Primer MNase-F also contained a SacI restriction site at its 5'end, upstream of the nuclear localization signal. The sequence

of the catalytic active N-domain of MNase was amplified from pFA6a-MN-KanMX6 (a kind gift from Ulrich Laemmli, University of Geneva) using primer MNase-F and MNase-R. The reaction mix was incubated at 95 C for 5min followed by two cycles of 95 C for 30sec, 37 C for 1min and 72 C for 2min. The reaction was then cycled 25 times at 95 C for 30sec, 55 C for 1min and 72 C for 2min, followed by a final incubation at 72 C for 7min. Primer MNase-R introduced a BamHI restriction site at the 3'end of the PCR product. The PCR product was gel purified and subsequently digested with restriction enzyme SacI (Promega) and BamHI (New England Biolab) using the conditions recommended by the manufacturer. The digested fragment was ligated into SacI/BamHI-cleaved plasmid pYES2-L/L with T4 DNA ligase (Promega). The ligation product was transformed into *E. coli DH5α* and the transformants screened for the correct plasmid by restriction enzyme digestion and partial nucleotide sequencing, using the standard T7 promoter forward primer and primer SEQ-R. The positive plasmid was denoted pMNL and the strain was stored as glycerol stock at -70 C. Handling of all sequence data was performed using DNAssist software (Patterton and Graves, 2000).

2.6.6. Correction of point mutations

Partial nucleotide sequencing of the full NLS-MNase-LexA-*lacUV5* sequence revealed the presence of point mutations that would have rendered the resulting protein inactive. The mutation was corrected as follows: sense and anti-sense oligonucleotides of 20bp were constructed such that 10bp flanked either side of the point mutation that needed to be changed. The point mutation occurred in next to the XbaI restriction site and the XbaI restriction sites were included in oligonucleotides Mut-F and Mut-R (Table 2.1). Two fragments were amplified from pMNL by PCR

using primers MNase-F and Mut-R, and primers Mut-F and lacUV5-R. The amplified fragments were digested with XbaI (Promega) according to recommendations of the supplier, and equimolar amount of the amplified fragments were ligated in a 10µl reaction using T4 DNA ligase at 4°C overnight. The re-generated construct was amplified using PCR with primer MNL-F and MNL-R (Table 2.1). The MNL-F and MNL-R introduced an NdeI and an XhoI restriction site at 5`- and 3`- end of the product respectively. The amplified fragment was digested with NdeI (Promega) and XhoI (Promega) according to recommendation of the supplier. The vector pET20b-(+) (Novagen) was also digested with identical enzymes, and the NdeI/XhoI double-digested amplified fragment ligated in the pET20b-(+) vector in a 10µl reaction at a vector to insert ratio of 1:5 (mole:mole), using T4 DNA ligase at 4°C overnight. The ligation mix was transformed into *E.coli DH5α*. The positive clones were screened by restriction digest and the corrected sequence was verified by partial nucleotide sequencing. The verified positive plasmid was denoted pCFC-MNL and the strain was stored as glycerol stock at -70°C.

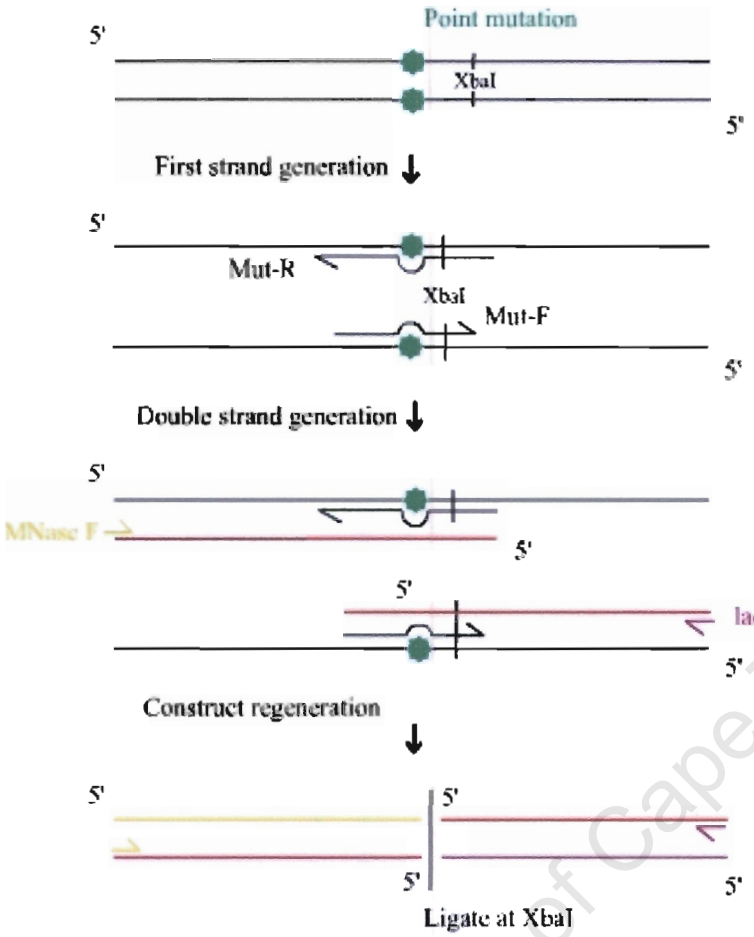


Fig. 2.1. Removal of point mutation from recombinant construct by site-directed mutagenesis. Two fragment of pMNL were amplified by PCR with oligonucleotides Mut-R and MNase-F, and oligonucleotides Mut-F and lacUV5-R. The fragments were digested with XbaI restriction site and the construct was regenerated by ligation of the two amplified fragments.

Chapter 3

A genome-wide study of chromatin compaction in *Saccharomyces cerevisiae*

3.1. Introduction

A clear link between linker histone binding and salt dependent chromatin compaction was previously shown (Thoma *et al.*, 1979). In this laboratory it was shown that yeast linker histone Hho1p displayed a significant increase in binding to chromatin during the transition to stationary phase (Schafer and Patterton, unpublished data). In addition, compaction of chromatin during stationary phase required Hho1p (Schafer and Patterton, unpublished data). The study of yeast in exponential phase and in stationary phase therefore provided an ideal system in which to study chromatin compaction, the relation of this compaction to the genomic distribution of Hho1p, and the relationship between this compaction and DNA function, particularly gene expression. We were interested in addressing this relationship on a genome-wide basis, and therefore aimed to investigate the condensation state of chromatin in exponential and stationary phase yeast cells at all gene promoter regions.

3.2. Fractionation of soluble yeast chromatin by sucrose gradient centrifugation

Chromatin fragments in different states of compaction could be separated by density, since an increase in compaction is associated with an increase in density (Ghirlando *et al.*, 2004). To allow separation of chromatin fragments by density and therefore compaction state, it is critical to retain the structure and compaction state that the

fragments possessed inside the living yeast cell. We therefore verified that the isolated chromatin fragments retained an intact nucleosomal structure by re-digestion of the fragments with MNase. Clear nucleosomal monomer and lower oligomer bands that were absent in the original digest were created by re-digestion of the isolated fragments (Fig. 3.1). This showed that the isolated chromatin fragments were still packaged into nucleosome cores, protecting 147bp from nucleolytic cleavage. Since linker histone H1/H5 quantitatively dissociated from chromatin only at 600mM NaCl, and since the extractions here were performed at 0.5mM $MgCl_2$, a concentration where Klug and colleagues previously demonstrated compaction in the presence of H1 (Thoma *et al.*, 1979), it appeared likely that Hho1p would remain bound to the chromatin.

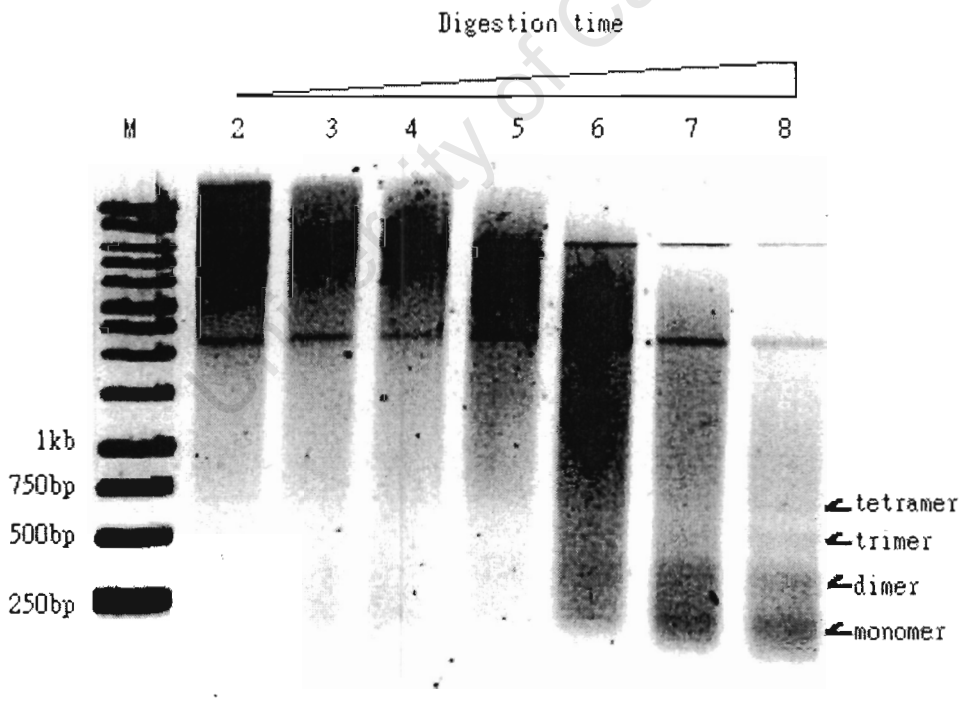


Fig. 3.1. The isolated chromatin fragments retained a nucleosomal structure. Soluble chromatin was re-digested with 1 U of MNase for 0sec, 30sec, 1min, 2min, 5min, 10min, 15min (lanes 2 to 8). The arrows indicate the position of nucleosome core-length fragments.

Soluble chromatin was prepared from exponential phase and stationary phase yeast cells, separated on a sucrose gradient, fractions collected, and the purified DNA from each fraction resolved on an agarose gel (see Fig. 3.2.). When considering a single fraction, the small DNA fragments were expected to be highly compacted, thereby sedimenting to a high density region on a sucrose gradient. The larger DNA fragments present in the same fraction correspond to decondensed chromatin where the larger mass of the DNA fragments allowed their sedimentation to the same density fraction. Fig. 3.2 shows a representative gel of the sucrose gradient fractions of chromatin isolated from a 6 day stationary phase culture. It was previously shown in our group that chromatin assumed a more condensed conformation during stationary phase in a Hho1p-dependent manner (Schafer and Patterton, unpublished data). The banding pattern in Fig. 3.2 indicates that the soluble chromatins fractionates according to density. The bulk of each fraction (most intense region of the fractions) serves as a good indicator of the average molecular size of DNA required to reach a certain density under a particular condensation state. The relatively small (4±1kb) fragments presented in the late fractions (fractions 13, 14 and 15) was expected to be composed of highly condensed chromatin (Ghirlando *et al.*, 2004).

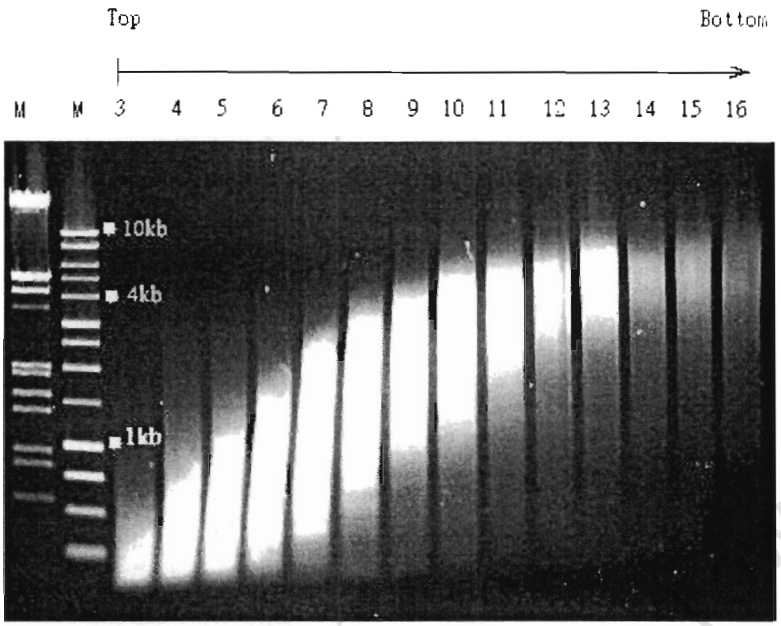


Fig. 3.2. Sucrose gradient fractionation of soluble chromatin extracted from a stationary phase yeast culture. Marker (lane 1). EcoRI/HindIII digested λ DNA marker. Marker (lane 2). 1kb DNA ladder (Promega). (Lanes 3 – 16). Collected sucrose gradient fractions 2 to 15.

A different banding pattern was observed for the sucrose gradient fractions of soluble chromatin extracted from exponential cultures. During exponential growth transcription level is highly elevated (Wodicka *et al.*, 1997), and chromatin was expected to be decondensed at most regions to allow access by the transcription machinery. Unlike the distribution of the chromatin fragments from the stationary phase cultures, the density distribution of the exponential phase culture was enriched in smaller DNA fragments that resolved at the top of the gradient (Fig. 3.3). Since the chromatin from exponential phase cultures was digested with identical amounts of MNase, this finding suggested that yeast chromatin in exponential phase was more sensitive to MNase cleavage, due to its less condensed state. It was previously observed that the chromatin in purified nuclei from exponential cultures was digested

to smaller fragments more quickly than that of stationary phase cultures, in agreement with the observed compaction differences between the chromatin in the two growth phases (Schafer and Patterton, unpublished data). Although the major proportion of the total DNA appeared in the upper fractions of the sucrose gradient, the presence of condensed chromatin was still detectable in bottom fractions, albeit in significantly less amounts.

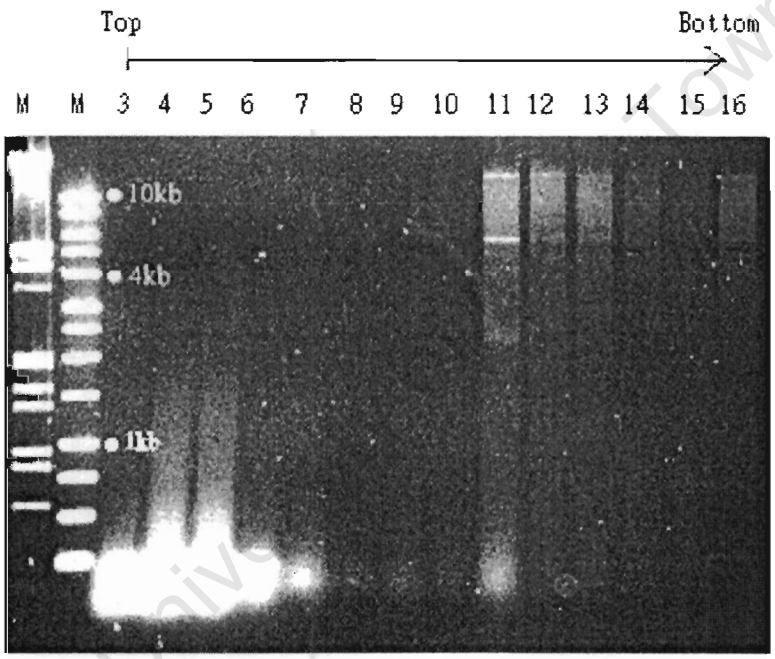


Fig. 3.3. Sucrose gradient fractionation of soluble chromatin extracted from an exponential phase yeast culture. Marker (lane 1). EcoRI/HindIII digested λ DNA marker. Marker (lane 2). 1kb DNA ladder (Promega). (Lanes 3–16). Collected sucrose gradient fractions 2 to 15.

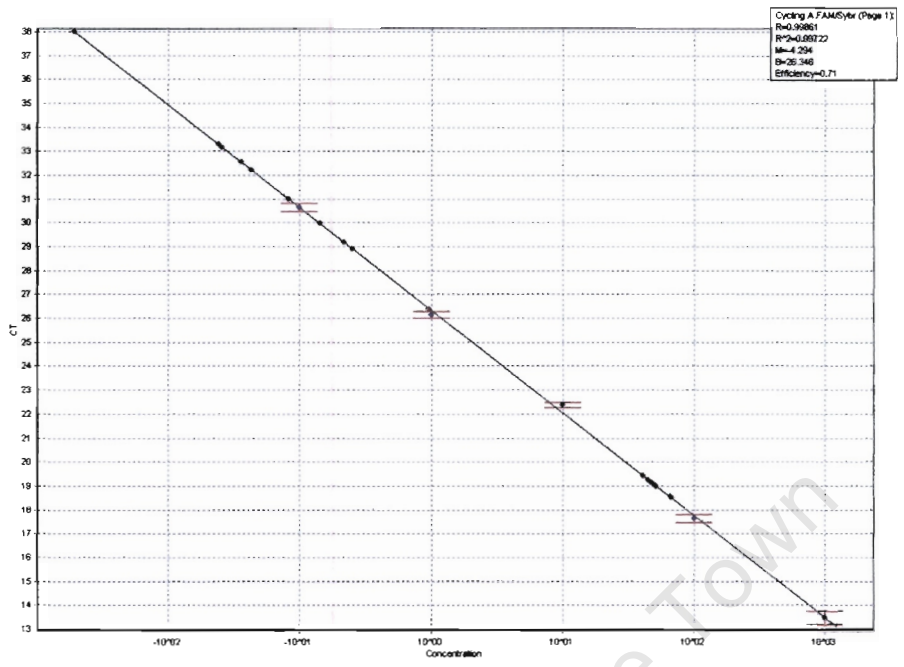
3.3. DNA quantification by kinetic PCR

Effective high throughput genome-wide screening for the presence of genes in a DNA sample can be achieved by the use of microarray technologies. A difficulty with

microarray techniques is the choice of proper normalization of the detected fluorescence signals to allow meaningful comparisons between slides. To overcome this, we designed a normalization method by relating the microarray fluorescence values of a slide to the absolute amounts of DNA used in the preparation of the Cy3/Cy5-labeled probe for that slide. This normalization method was used to adjust the microarray values to reflect the amount of DNA associated with the particular microarray slide.

The focus of the study was placed on the condensed regions of chromatin (4kb±1kb) isolated from the sucrose gradient (selected fractions contained best representation of differently sized DNA fragments). The total input DNA as well as the isolated 4kb region from a sucrose gradient fraction, representing condensed chromatin, were quantified for the abundance of 3 genes, namely *DNF1, ROX1* and *SNZ1*. *ROX1* encodes a heme-dependant repressor protein whose expression decreased after yeast cells entered diauxic shift. *SNZ1* is a stationary phase-induced gene, which encodes a protein involved in vitamin B biosynthesis. The expression of *DNF1* gene, which encodes a non-essential ATPase, is approximately constant during growth from exponential phase to stationary phase (data from *Saccharomyces* Genome Database, http://www.yeastgenome.org/). The abundance of the 3 genes in the different DNA samples was determined in triplicate by kinetic PCRs. Standard curves, in the range from 0.1pg/µl to 10ng/µl covered the range of concentrations determined for the 3 genes in the different DNA samples (Fig. 3.4).

a.



b.

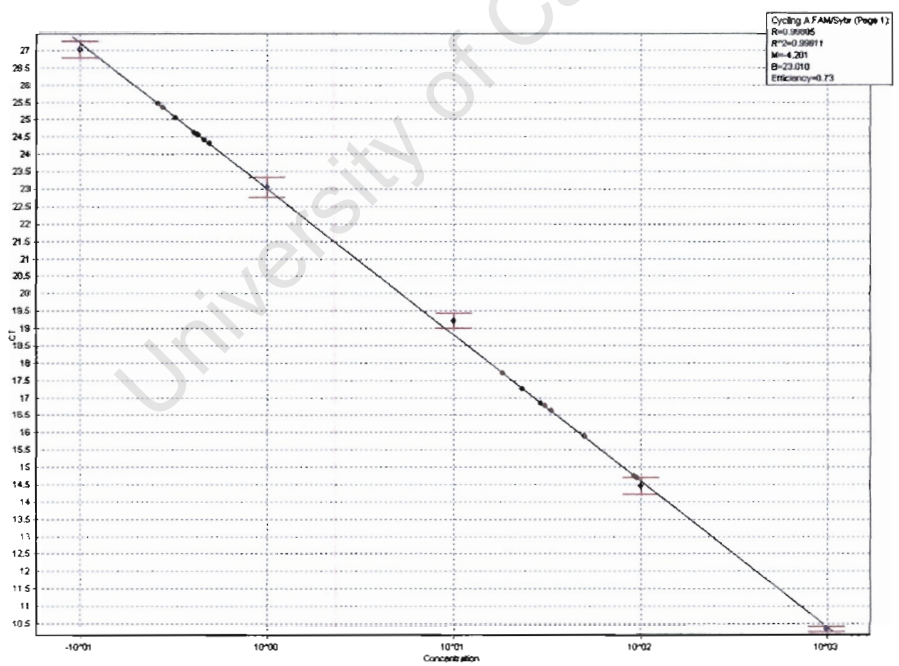
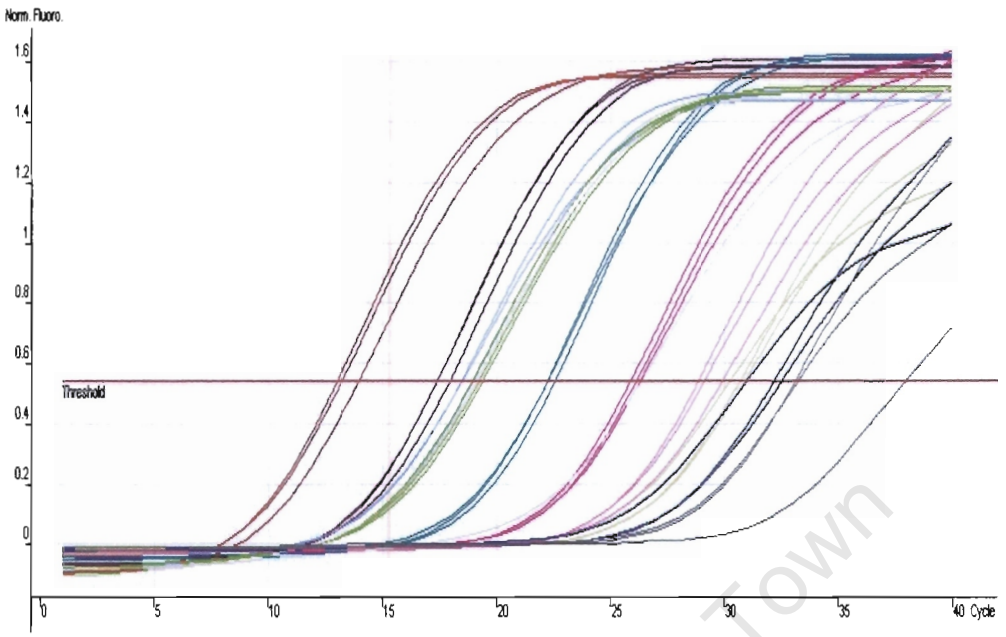


Fig. 3.4. Representative standard curves used for kinetic PCR quantification. a. The standard curve constructed for *ROX1* quantification in chromatin obtained from an exponential phase culture. **b.** Standard curve constructed for *ROX1* quantification in chromatin obtained from a stationary phase culture.

The raw fluorescence signals shown in Fig. 3.5 represent the data captured for the sucrose gradient input and the condensed (4kb) chromatin sample by kinetic PCR. The threshold value was automatically calculated for all the kinetic PCRs. The presence of the correct amplified product and absence of fluorescence artefacts such as primer dimers was confirmed by a melting curve analysis following the kinetic PCR, and involved the clear identification of a peak in the first derivative of the melting curve at approximately 85°C (Fig. 3.6). The appearance of a second peak at approximately 76°C was most likely due to the amplification of either contaminating DNA, or extended primer pairs. The second peak was observed mostly at low template concentrations (<1pg/µl). This finding suggested that at low template concentrations, the probability of primer dimer extension was greatly enhanced. The presence of primer extensions also resulted in shifting of the peaks since the PCR products melted over a broader range of temperatures (Fig. 3.6, b.).

a.



b.

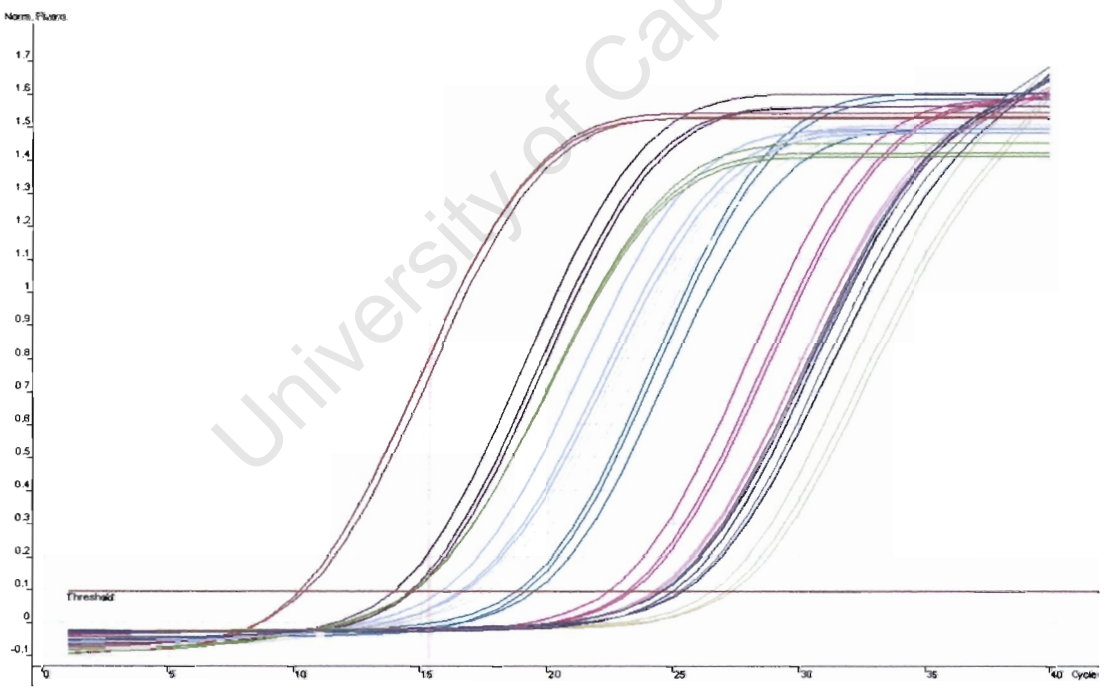
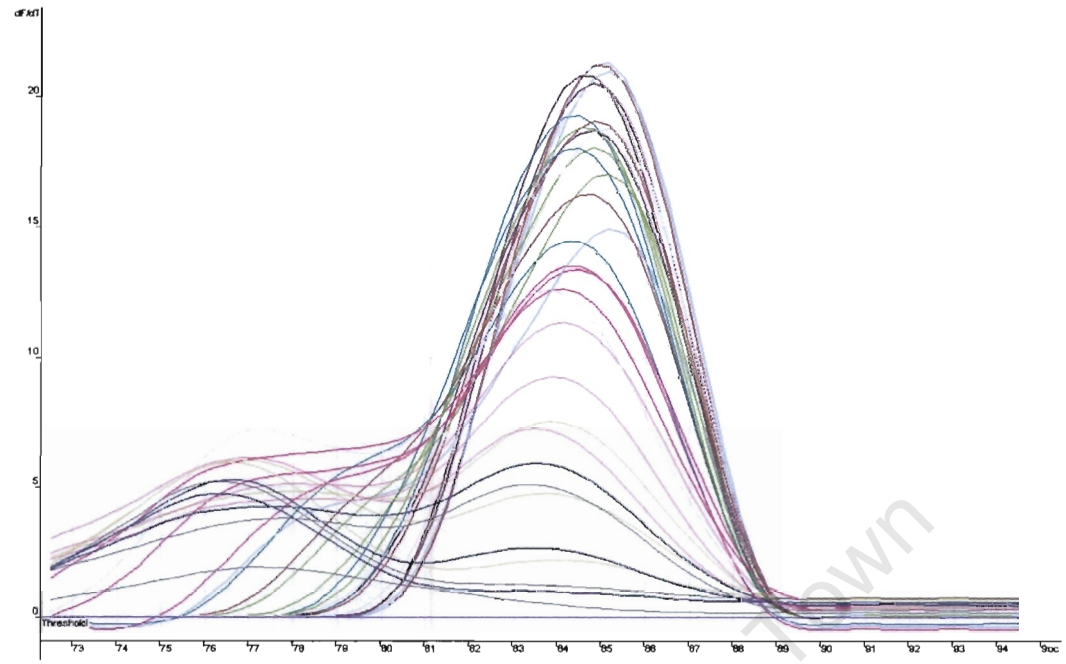


Fig. 3.5. Representative raw fluorescence data obtained by kinetic PCR. a. Raw fluorescence data of *ROX1* quantification in chromatin obtained from exponential phase culture. **b**. Raw fluorescence data of *ROX1* quantification in chromatin obtained from stationary phase culture.

a.



b.

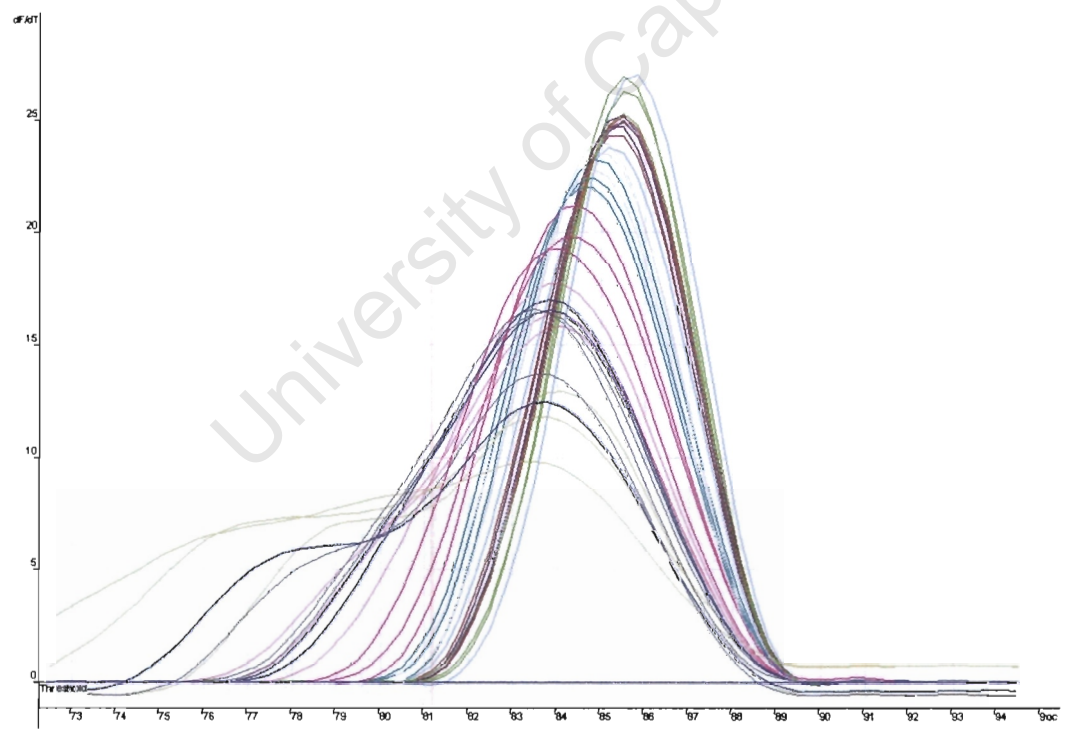


Fig. 3.6. Representative melting curve data obtained from kinetic PCR. a. Melting curve of *ROX1* quantification in chromatin obtained from exponential phase culture. **b**. Melting curve of *ROX1* quantification in chromatin obtained from stationary phase culture.

The concentrations of the 3 genes present in each of the extracted DNA samples were determined in triplicate (Table 3.1). The *ROX1* gene was observed to display a slight enrichment in the condensed chromatin of stationary phase cultures at similar input concentration levels (Table 3.1.). This initial finding supported the thinking that transcriptionally inactive genes were packaged into condensed structure, since the expression of *ROX1* is down-regulated after the diauxic shift. Similar deductions could not be made for the *DNF1* and *SNZ1* gene, since the level of input DNA between the exponential and stationary phase cultures differed significantly and two reactions for *SNZ1* quantification failed to produce any reliable fluorescence signal.

Table 3.1. Concentration of *DNF1, ROX1* and *SNZ1* in input and condensed chromatin determined by kinetic PCR. Each value reported was the average concentration calculated from 3 reactions performed using the same DNA template. An identical analysis was performed on each of three independently fractionated preparations.

	DNF1 (pg/µl)		*ROX1* (pg/µl)		*SNZ1* (pg/µl)	
	Exponential	Stationary	Exponential	Stationary	Exponential	Stationary
Input ref. 1	0.112± 0.035	30.2± 12.3	61.4± 8.70	37.9± 10.5	41.9± 16.4	34.6± 8.22
Sample 1	0.065± 0.036	0.138± 0.020	0.206± 0.057	0.458± 0.035	5.41	8.96± 0.760
Input ref. 2	0.118± 0.018	27.4± 31.9	45.9± 5.63	94.4± 1.24	43.9± 12.6	55.8± 7.71
Sample 2	0.080± 0.028	0.089± 0.025	0.054± 0.025	0.331± 0.075	ND*	11.8± 1.50
Input ref. 3	0.138± 0.01	4.81± 2.36	32.7± 27.5	23.7± 5.5	47.9± 15.8	25.9± 4.24
Sample 3	1.343± 0.313	0.08± 0.04	0.017± 0.013	0.377± 0.085	ND*	13.2± 6.52
Input avg.	0.123± 0.014	20.8± 13.9	46.7± 14.4	52± 37.4	44.6± 3.06	38.8± 15.4
Sample avg.	0.496± 0.734	0.102± 0.031	0.0923± 0.1	0.389± 0.064	5.41	11.32± 2.16

*None detected: these samples failed to produce a reliable fluorescence signal

3.4. Microarray analysis

3.4.1. The yeast genome is more condensed in stationary phase

It is well established that transcription and protein synthesis levels drop dramatically in *S. cerevisiae* during stationary phase growth (Pringle *et al*.,1991) It was also observed in this laboratory that yeast chromatin was more compacted in a Hho1p-dependent manner in stationary phase. In order to analyze the degree of chromatin compaction in exponential phase and in stationary phase on a genome-wide basis, cyanine-labeled probe was prepared from the total input DNA and from the isolated 4-kb region of sucrose density fractionated samples. Independent analyses involving triplicate DNA preparations and triplicate microarray slide hybridizations were performed. The compaction state of the chromatin region in which an intergenic region is located is indicated by the proportional abundance of such an intergenic region in DNA contained in condensed chromatin *versus* the total input DNA. A comparison of the sample:input ratio showed that condensed chromatin extracted from exponential phase cultures included less intergenic regions than that of stationary phase cultures (Fig. 3.7). This finding indicated that there is an enrichment of intergenic regions in condensed chromatin extracted from stationary phase cultures. This result also suggested that the majority of the yeast chromosomes are decondensed during exponential phase growth. This result agrees well with the general thinking that active transcription during exponential phase is associated with decondensed chromatin.

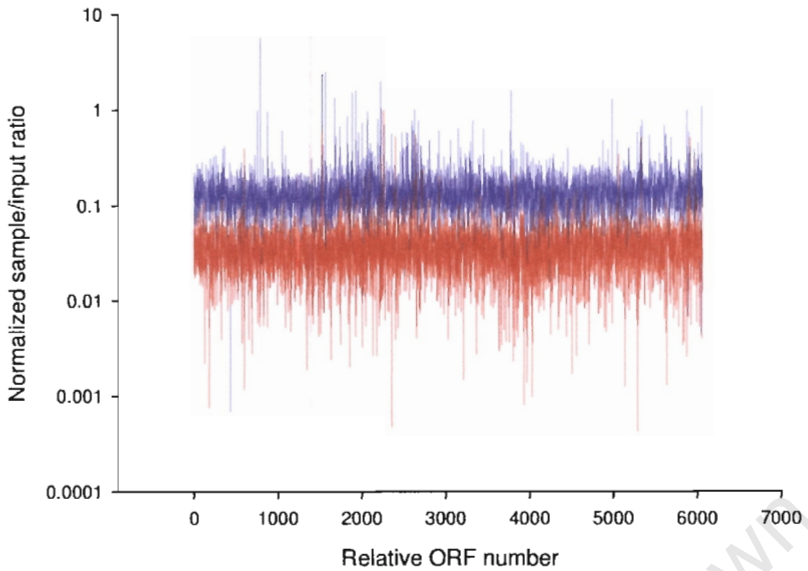


Fig. 3.7. A histogram of the normalized 4kb:input rations of all intergenic regions **in *S. cerevisiae* chromatin in exponential phase and stationary phase.** The blue columns represent stationary phase enrichment ratios and the red columns represent exponential phase enrichment ratios. Each column represents an averaged ratio obtained from three independent experiments. Intergenic regions are consecutively arranged from the left-hand side of chromosome I to the right-hand side of chromosome XVI.

Data from a previous chip-on-ChIP study in this laboratory indicated a genome-wide association of Hho1p to yeast chromosomes during stationary phase entry (McEvoy and Patterton, unpublished data). The increase in genome-wide condensation of chromatin in stationary phase suggested that this extensive binding resulted in compaction of the entire yeast genome. This suggestion is in agreement with the observed dependence of chromatin compaction in stationary phase cells on the presence of Hho1p (Schafer and Patterton, unpublished data).

The comparisons of the relative intergenic region abundance in exponential and stationary phase condensed chromatin revealed a number of genes that appeared to be more compacted during exponential phase than in stationary phase (Table 3.2), unexpected for many of these genes as they encoded products function associated with exponential cell growth. The *SRD1* gene encodes a protein involved in processing of pre-rRNA to rRNA and is transcribed at high levels during exponential phase growth (data from *Saccharomyces* Genome Database). However, it was observed that the *SRD1* gene was enriched in condensed chromatin in exponential phase. In contrast, a gene encoding a nuclear protein that was also involved in rRNA processing is the *NOP7* gene. The ORF of the *NOP7* gene was observed to be enriched in condensed chromatin in exponential phase when it was expressed at basal levels. The *NOP4* gene, encoding another nuclear protein involved in rRNA processing, was also found to be more compacted during exponential phase. The occurrence of these genes encoding products involved in rRNA processing within compacted chromatin during exponential phase suggested that the presence of a gene in a condensed region of chromatin did not necessarily preclude transcriptional induction. These genes are located on different chromosomes (chromosome III, chromosome VII and chromosome XVI), suggesting that coincidental enrichment of the genes in the condensed chromatin of exponential phase cultures was unlikely. The *URA4* gene was observed to be more condensed with the chromatin of exponential phase culture. The *APL6* gene, which is transcriptionally silenced after diauxic shift, displayed low presence in the condensed stationary phase chromatin, which suggested that silenced *APL6* gene does not become packaged even when its expression was silenced. This data indicated that repression of transcription does not always correlate with the condensation of the chromatin that packages a particular gene promoter. Although contrary to the general thinking that transcription activity is often coupled

with open chromatin, the data in this study suggested possible novel function of chromatin compaction in transcription regulation.

Table 3.2. List of ORFs that displayed a higher abundance in condensed chromatin in exponential phase than in stationary phase.

ORF name	*Gene*	*Function*
YAR019C	*CDC15*	Protein kinase of the Mitotic Exit Network
YBR041W	*FAT1*	Fatty acid transporter
YDL100C	*GET3*	ATPase, subunit of the GET complex
YDR334W	*SWR1*	Swi2/Snf2-related ATPase, component of the SWR1 complex
YDR529C	*QCR7*	Subunit 7 of the ubiquinol cytochrome-c reductase complex
YER025W	*GCD11*	Gamma subunit of the translation initiation factor eIF2
YER110C	*KAP123*	Karyopherin beta, mediates nuclear import of ribosomal proteins prior to assembly into ribosomes and import of histones H3 and H4
YGL148W	*ARO2*	Bifunctional chorismate synthase and flavin reductase
YGL222C	*EDC1*	RNA-binding protein, activates mRNA decapping
YGL240W	*DOC1*	Processivity factor required for the ubiquitination activity of the anaphase promoting complex
YGR009C	*SEC9*	t-SNARE protein important for fusion of secretary vesicles with the plasma membrane
YHR072W	*ERG7*	Lanosterol synthase
YHR079C	*IRE1*	Serine-threonine kinase and endoribonuclease
YHR091C	*MSR1*	Mitochondrial arginyl-tRNA synthetase
YHR120W	*MSH1*	DNA-binding protein of the mitochondria involved in repair of mitochondrial DNA
YHR196W	*UTP9*	Nucleolar protein, component of the small subunit (SSU) processome involved in processing of pre-18S rRNA
YKL041W	*VPS24*	One of four subunits of the endosomal sorting complex required for transport III
YLL003W	*SFI1*	Centrin (Cdc31p)-binding protein required for spindle pole body (SPB) duplication
YLR119W	*SRN2*	Component of the ESCRT-I complex, which is involved in ubiquitin-dependent sorting of proteins into the endosome

YLR425W	*TUS1*	Guanine nucleotide exchange factor (GEF) that functions to modulate Rhop1 activity as part of the cell integrity signaling pathway
YOR193W	*PEX27*	Peripheral peroxisomal membrane protein involved in controlling peroxisome size and number
YPL043W	*NOP4*	Nucleolar protein, essential for processing and maturation of 27S pre-rRNA and large ribosomal subunit biogenesis
YPL161C	*BEM4*	Protein involved in establishment of cell polarity and bud emergence
YPL177C	*CUP9*	Homeodomain-containing transcriptional repressor of PTR2, which encodes a major peptide transporter
YBR040W	*FIG1*	Integral membrane protein required for efficient mating
YCR018C	*SRD1*	Protein involved in the processing of pre-rRNA to mature rRNA
YDL230W	*PTP1*	Phosphotyrosine-specific protein phosphatase that dephosphorylates a broad range of substrates in vivo, including Fpr3p
YDR135C	*YCF1*	Vacuolar glutathione S-conjugate transporter of the ATP-binding cassette family
YEL071W	*DLD3*	D-lactate dehydrogenase
YFR028C	*CDC14*	Protein phosphatase required for mitotic exit
YFR034C	*PHO4*	Basic helix-loop-helix (bHLH) transcription factor
YGL205W	*POX1*	Fatty-acyl coenzyme A oxidase, involved in the fatty acid beta-oxidation pathway
YGR082W	*TOM20*	Component of the TOM (translocase of outer membrane) complex responsible for recognition and initial import steps for all mitochondrially directed proteins
YGR103W	*NOP7*	Nucleolar protein involved in rRNA processing and 60S ribosomal subunit biogenesis
YGR261C	*APL6*	Beta3-like subunit of the yeast AP-3 complex
YHR206W	*SKN7*	Nuclear response regulator and transcription factor, part of a branched two-component signaling system
YLR420W	*URA4*	Dihydroorotase, catalyzes the third enzymatic step in the de novo biosynthesis of pyrimidines
YML104C	*MDM1*	Intermediate filament protein, required for nuclear and mitochondrial transmission to daughter buds
YMR239C	*RNT1*	RNAase III; cleaves a stem-loop structure at the 3' end of U2 snRNA
YNR044W	*AGA1*	Anchorage subunit of a-agglutinin of a-cells
YOL013C	*HRD1*	Ubiquitin-protein ligase required for endoplasmic reticulum-associated degradation (ERAD) of misfolded proteins

We were interested in determining whether there was a correlation between the compaction state of each gene in Table 3.2 in stationary phase and in exponential phase. To address this issue, we performed a scatter analysis on the condensation state of each gene in exponential phase and in stationary phase (Fig. 3.8). Referring to Fig. 3.8, it appears that the condensation state of a gene in stationary phase has little predictive value for the degree of compaction of that gene in exponential phase. In fact, the two data sets have a correlation coefficient of 0.036, indicating negligible dependency between the data set pairs. This is not unexpected, since the entire genome is compacted in stationary phase (Fig. 3.7), and it is unlikely that the functional requirement for compaction of a genomic region during stationary phase will remain in exponential phase, where gene expression and chromatin decondensation is likely to be related to the particular carbon source and general environment.

An interesting feature of Fig. 3.8 is the "+" appearance of the data pairs. This does not mean that genes that were within the average condensed range in exponential phase became highly condensed or decondensed during stationary phase. An analysis of the standard deviation of the stationary phase compaction values at 3σ for the ranges <-0.1, -0.1-0.1 and >0.1 on the abscissa showed that the standard deviation of the extreme ranges (horizontal arms of the "+") were larger than that of the central range which contained most of the data points. This showed that the scarcity of data points with high and low compaction values in stationary phase in regions that were highly condensed or decondensed in exponential phase was probably due to the smaller data population in the extreme regions.

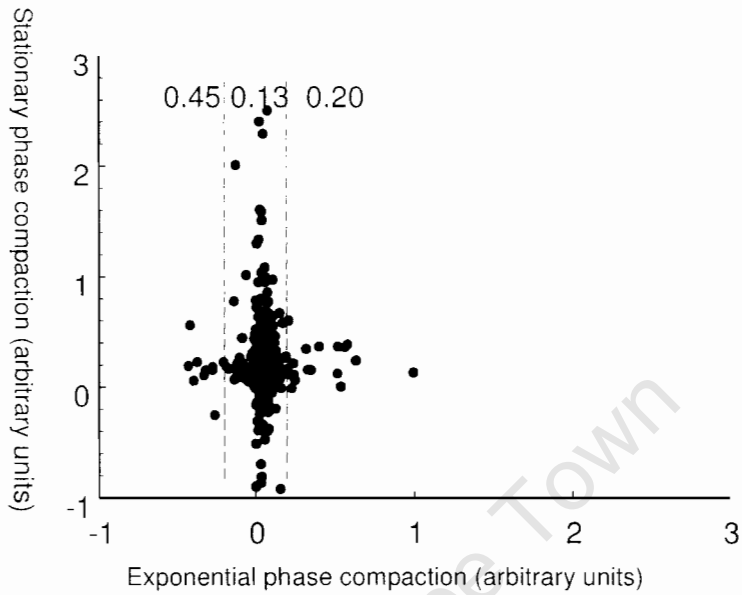


Fig. 3.8. Comparison of the genome-wide compaction state of *S. cerevisiae* genes during exponential phase and stationary phase. The compaction ratio of each gene in exponential phase was plotted against the value of the gene in stationary phase. The standard deviation of the stationary phase compaction values at 3σ for the abscissa ranges <-0.1, -0.1-0.1 and >0.1 are shown.

3.4.2. There is no relation between gene density and condensation state in *S. cerevisiae.*

The genome-wide study performed by Gilbert *et al.* (2004) indicated that open chromatin corresponded to regions of high gene density. We performed a similar investigation and analyzed the correlation between promoter density and chromatin compaction of the entire yeast genome using consecutive 100kb windows (Fig. 3.9). Referring to this figure, there appears to be little correlation between gene density and condensation status. This is in contrast to the finding of Gilbert *et al.*, which reported

that gene-dense regions of the human genome were in a more open configuration. We saw a clear difference in the condensation state of the yeast genome in exponential and in stationary phase, but this appeared to be a quality of the whole genome, as opposed to distinct regions. The data reported by Gilbert *et al.* may be a reflection of the presence large satellite sequences in the human genome which tend to be packaged into heterochromatin. However, the relationship between gene density and chromatin compaction remains complex and illusive as Gilbert *et al.* also reported variance in structure amongst the different human satellite sequences.

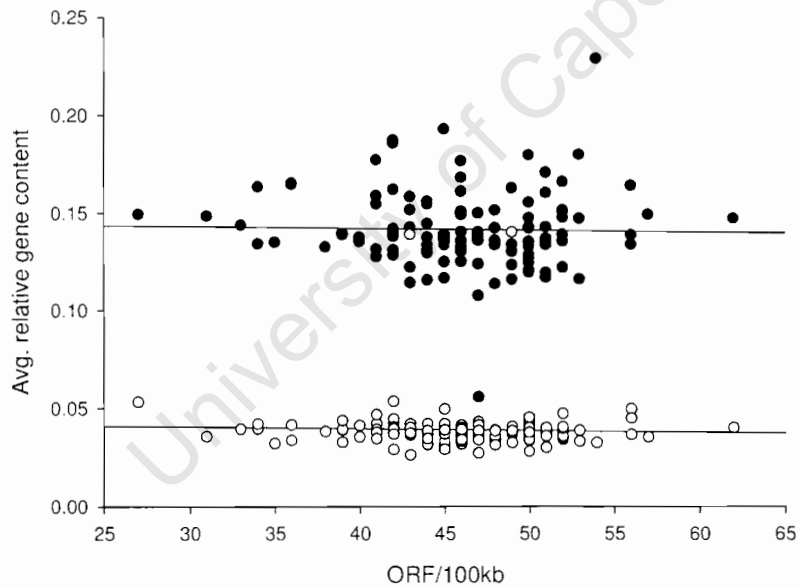


Fig. 3.9. A simple linear regression plot of average gene content of condensed chromatin to the density of ORFs over 100-kb repeats. The closed circles represent condensed chromatin of stationary phase cultures and the open circles represent condensed chromatin of exponential phase cultures.

3.4.3. The Hho1p-dependent condensation of the yeast genome appeared to occur on a global scale.

We finally looked at the relation between compaction of genes in the 10kb region that was expected to contain less condensed chromatin, versus the 4kb region of the same sucrose gradient fraction (Fig. 3.10). The linear regression plot comparison of the decondensed chromatin microarray values against condensed chromatin microarray value revealed a correlation, showing that genes that are enriched in the 4kb fraction is also enriched in the 10kb fraction. This is expected for a genome that is condensed or decondensed on a genome-wide scale, as opposed to local regions of decondensation. This data therefore suggests that Hho1p-dependent compaction of the *S. cerevisiae* genome occurs globally in stationary phase. The linker histone appeared to have little role in compaction of local regions of chromatin.

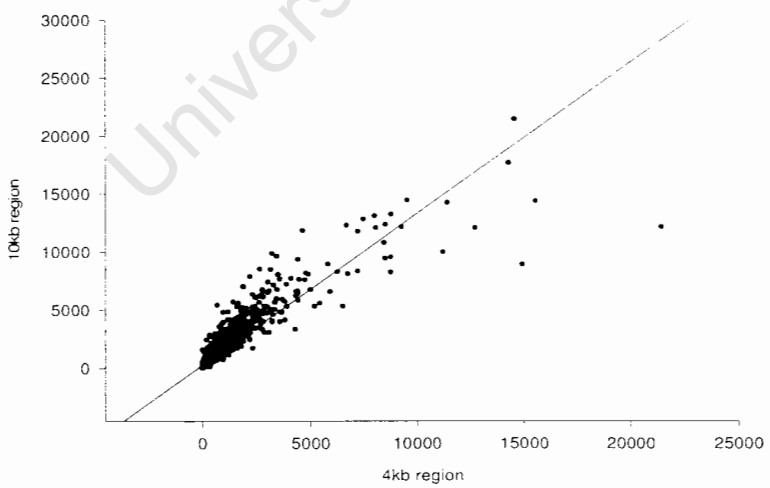


Fig. 3.10. A linear regression plot comparison of the relative gene content between decondensed (10kb) chromatin and condensed (4kb) chromatin of exponential phase cultures.

Chapter 4

Construction of a recombinant MNase-LexA fusion protein

4.1. Concept of a recombinant MNase-LexA fusion protein and cloning strategy

4.1.1. Design of recombinant fusion protein

An approach to determine the degree of compaction of a condensed chromatin fragments is to compare the total contour length of the DNA present within a defined space in the condensed chromatin. Based on this concept, a way to measure the contour length of DNA was devised in this study. A recombinant protein was designed to generate "marker signals" which would indicate the extent accessible to an anchored protein, relative to the anchoring site. DNA strand breakage is one way to generate such clear marker signals that would define accessible range, and at low frequency (single hit kinetics) is expected not to disrupt the compaction status of a chromatin. The degree of compaction of the chromatin fragment would then be determined by the average size of DNA fragment generated relative to the protein anchoring site (Fig. 4.1.).

In addition to the ability to generate DNA strand breakage, the recombinant protein is also required to effective at anchoring. In a previous yeast one-hybrid system, the bacterial LexA operon was shown to be capable of stably anchoring M.SssI DNA methyltransferases that were fused to a LexA protein (Feng *et al.*, 2004). The short consensus LexA operator sequence (5'-CTGTATGTACATACAG-3') can easily be introduced via PCR-based homologous recombination to any region of interest in the

S. cerevisiae genome, and the recombinant protein targeted in this fashion.

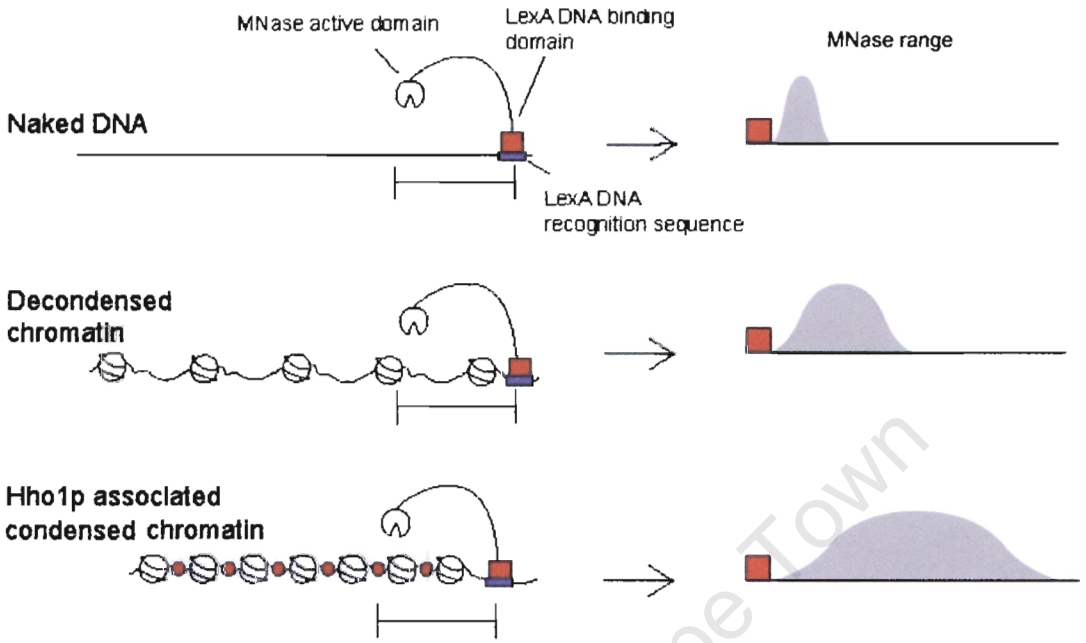


Fig. 4.1. Conception of a recombinant fusion protein to measure chromatin compaction. The contour distance accessible to an anchored protein is expected to be greater in condensed chromatin than in decondensed chromatin. It is assumed that the distribution of the protein-DNA contact frequency, proportional to the cleavage frequency of the anchored protein, would assume a Gaussian curve.

4.1.2. Cloning strategy

4.1.2.1. Dealing with leaky expression of the nuclease during sub-cloning

The nuclease activity of the recombinant protein presented an obstacle to construction and propagation of the encoding plasmid, since any expression of the construct during the cloning stages would kill the host cell. A mechanism was needed to suppress gratuitous expression of the recombinant protein during the cloning stages. Wang and

Simpson faced a similar problem in their study when they performed the sub-cloning of DNaseI (Wang and Simpson, 2001). They successfully solved the problem by placing a bacterial *lac-UV5* promoter in anti-sense direction at the 3' end of the recombinant construct. This was expected to produce low levels of anti-sense RNA to the encoded nuclease, which diminished the effect of leaky expression of the construct, allowing successful sub-cloning of the nuclease (Wang and Simpson, 2001). The same strategy was applied in this study to prevent any expression of the construct during sub-cloning of the recombinant protein.

4.1.2.2. Sub-domain order of the fusion protein and steric effects

From the crystal structure of MNase it was clear that the active site of MNase lies close to the N-terminal part of the peptide (Cotton *et al.*, 1979). Fusion of the LexA domain to N-terminus of MNase active domain could therefore result in disruption of active site structure and reduced activity. From the NMR structure of the LexA protein, it was reported that the N-terminal residues of the LexA protein formed a helix-turn-helix motif, and that this motif was involved in DNA binding (Fogh *et al.*, 1994). A model was also proposed where the C-terminal residues of LexA was involved in LexA dimerization (Dumoulin *et al.*, 1993). In another study, fusion proteins were successfully constructed by fusion of peptides to the N-terminus of LexA (Nehme *et al.*, 2004). From these considerations it appeared reasonable to place the LexA protein at the C-terminus of the fusion protein, minimizing any possible structural disruption to the MNase active site, as well as retaining the efficient DNA-binding property of the LexA protein.

A nuclear localization signal was needed to target the recombinant protein to the nucleus. The SV40 large T- antigen nuclear localization signal was successfully attached to DNaseI in an *in vivo* study by Wang and Simpson (2001). It was decided to use the same nuclear localization signal.

4.2. Recombinant MNase-LexA construction

The MNase-LexA fusion protein was constructed by sequential sub-cloning of the constituent domains. The correct orientation of the construct was ensured by the utilization of incompatible restriction sites for every ligation procedure. The *lacUV5* promoter (64bp) needed to be inserted first to limit leaky expression during the construction of the recombinant fusion protein. The *lexA* gene (606bp) was ligated to the *lacUV5* promoter at an XbaI restriction site (Fig. 4.2). The plasmid pYES2-L/L was constructed by the insertion of the *lexA-lacUV5* fragment into vector pYES2 at the BamHI and EcoRI restriction sites. The pYES2-L/L was transformed into *E. coli DH5α* and positive clones carrying the correct plasmid were identified by the release of an approximately 700bp product with a BamHI/EcoRI double digestion (Fig. 4.3). The plasmid pMNL was constructed by insertion of the MNase N-domain amplified from pFA6a-MN-KanMX6 into pYES2-L/L at the SacI and BamHI restriction site. Fig. 4.4. shows the restriction digest banding patterns of pMNL generated by a combination of different restriction enzymes. The complete construct inserted into pYES2 had an expected size of 1189bp.

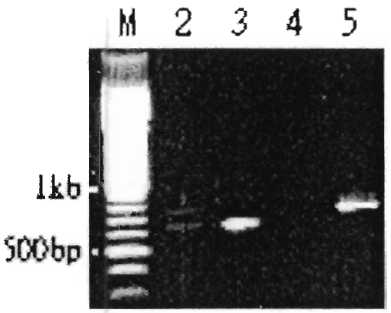


Fig. 4.2. The ligation of *lacUV5* promoter to *lexA*. Marker (lane M). 100bp DNA step ladder (Promega). (Lane 2). ligation mixture of the *lexA-lacUV5* fragment. The bottom band corresponds to a *lexA* fragment and the top band is the *lexA-lacUV5* fragment. (Lane 3). The *lexA* PCR product. (Lane 5). The *lexA-lacUV5* PCR product.

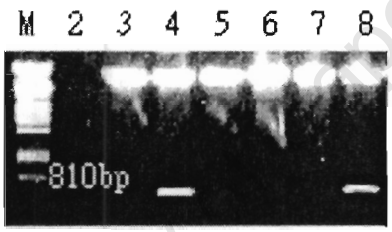


Fig. 4.3. The screening of positive pYES2-*lexA-lacUV5* plasmid. Marker (lane M). λ PstI DNA marker. The positive plasmids were identified in lane 4 and 8 by the release of a band corresponding to approximately 700bp after BamHI/EcoRI double digestion.

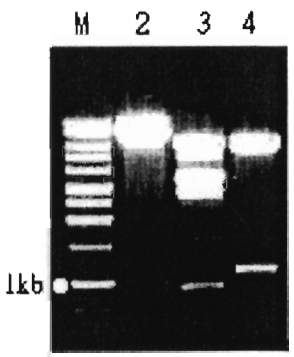


Fig. 4.4. Identification of positive pMNL using restriction enzyme digest. Marker (lane M). 1kb ladder marker (Promega). (Lane 2). pMNL digested with SacI. (Lane 3). pMNL digested with EcoRI. (Lane 4). pMNL double digested with SacI/XbaI.

The restriction digest of pMNL by SacI yielded a product of approximately 7kb, confirming the insertion of a full construct (lane 2, Fig. 4.2.). The many bands observed for the EcoRI-digested pMNL (lane3, Fig. 4.2.) resulted from the star activity of the restriction enzyme in Multi-Core buffer (Promega). The release of an approximately 900-bp fragment by EcoRI-digested pMNL verified the presence of an insert corresponding to the full-size construct.

An EcoRI restriction site was situated 260bp downstream from the start of the recombinant construct, in the MNase N-domain. It was thus expected that a fragment of 924bp would be released by EcoRI digestion of pMNL (Fig. 4.5). Lane 4 of Fig. 4.4 showed that a single band was released by SacI/XbaI double digestion of pMNL. There are two XbaI restriction sites in pMNL, one within the fusion region between the *lacUV5* promoter component and the *lexA* component, the other outside the construct at the end of the multiple cloning sites (MCS) of the pYES2 vector (Fig. 4.5). The SacI/XbaI double digested pMNL was therefore expected to release two fragments from the cleavage of three restriction sites. The single band representing fragments of approximately 1100bp corresponded to the fragment released by the digestion of the SacI site and the XbaI site situated within the construct. The other fragment of smaller size (approximately 100bp) was not observed, as the band had migrated off the gel.

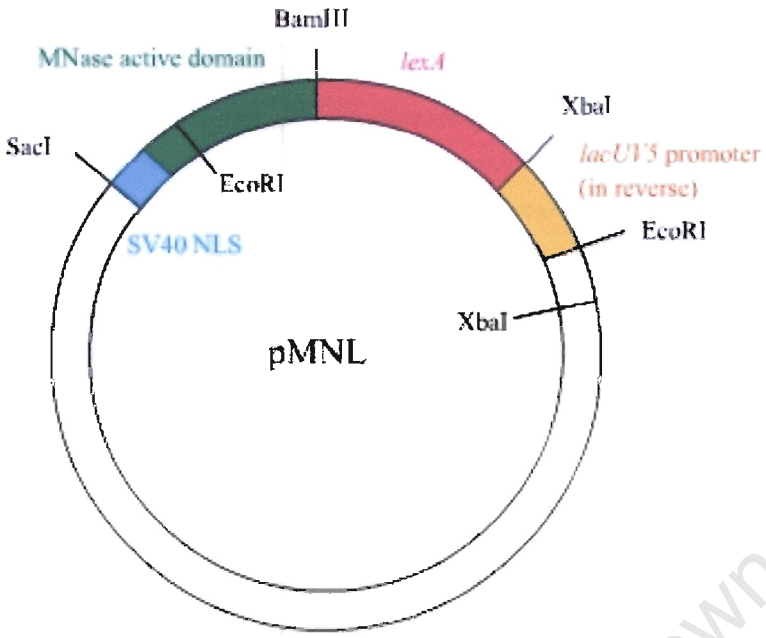


Fig. 4.5. A map showing the restriction sites utilized during the construction and screening of pMNL. The 1189bp full construct of the recombinant protein was inserted between the SacI and EcoRI restriction site of pYES2.

Unambiguous confirmation that the correct fusion construct was sub-cloned, came from partial nucleotide sequencing. Fig. 4.6. shows the alignment of the constructed fusion gene sequence to the theoretical, predicted sequence.

```
pMNL_SeqR          754    CGAGTTTAAACCAATTGTCGTAGATCTTCGTCAGCAGAGCTTCACCATTGAAG 806
Predicted construct 002   CGAGTTTAAACCAATTGTCGTAGATCTTCGTCAGCAGAGCTTCACCATTGAAG 1054
pMNL_SeqR           807   GGCTGGCGGTTGGGGTTATTCGCAACGGCGACTGGCTGTAATCTAGAAAATTAT 859
Predicted construct 1055  GGCTGGCGGTTGGGGTTATTCGCAACGGCGACTGGCTGAAATCTAGAAAATTAT 1107
pMNL_SeqR          860    TGTCTAACAATAATTCCACACATTATACGAGCCGGAAGCATAAAGTGTAAAGC 912
Predicted construct 108   TGTCTAACTATAATTCCACACATTATACGAGCCGGAAGCATAAAGTGTAAAGC 1160
pMNL_SeqR          913    CTGGGGTGCCTAATGAGTGACGGAATTCTGCAGAT 947
Predicted construct 1161  CTGGGGTGCCTAATGAGTGACGGAATTC       1188
```

Fig. 4.6. Sequence alignment of expected and predicted fusion protein ORF. The positive construct sequence was generated by joining the sequencing results acquired from the forward and reverse sequencing reaction at the position corresponding to position 524 of the expected sequence. Mismatches between the two sequences are highlighted in green.

4.3. Correction of point mutations accumulated during sub-cloning

Sequencing of the final construct revealed five regions of mismatches to the expected sequence. The sequence at the start of the final construct (the first 6bp) was observed to be completely different to that of the expected sequence. This observation was ignored as the sequence was part of the MNase-F primer. The close proximity of the upstream T7 promoter used for the priming of the sequencing reaction may have contributed towards the unstable signals, as signals acquired close to the priming regions of a sequencing reaction are often inaccurate.

At position 483 of the expected sequence (Fig. 4.6.), a TAA sequence, which codes for a UAA stop codon, was replaced by a BamHI restriction site in the actual construct. The sequence of the BamHI restriction site codes for a glycine (GGA) and a serine (UCC) residue (Fig. 4.6; Fig. 4.9). Three single nucleotide mutations were

observed in the full construct at the positions that corresponded to positions 942, 992 and 1116 of the expected construct. The insertion of an additional T at the position corresponding to position 992 of the expected sequence caused a frame-shift, which would render the encoded protein unusable (Fig. 4.3). The insertion was removed by site-directed mutagenesis with oligonucleotide primers Mut-F and Mut-R (see Table 2.1). The other two point mutations observed at positions corresponding to position 942 and 1116 of the expected construct did not alter the encoded amino acid residue, since the mutations were fortuitously situated on the third position of a glutamate codon (GAA → GAG) and a threonine codon (ACT → ACA), respectively (Fig. 4.6; Fig. 4.9). The codon usage frequency of the glutamate codon was shown to decrease from 68% to 32%, while the threonine codon usage frequency increased from 15% to 47% (Nakamura *et al.*, 1999). However, these two point mutations were ignored, since the mutations had no effect on the recombinant fusion protein sequence, and were not expected to have a significant effect on protein synthesis efficiency.

4.4. Sub-cloning of the corrected fusion protein ORF

The plasmid pCFCMNL was constructed by inserting the corrected full ORF fragment into vector pET20b-(+), between the NdeI and XhoI restriction sites. An XhoI site was introduced into the 3' end of the construct by replacing the EcoRI site on the lacUV5-R primer (MNL-R, Table 2.1), while an NdeI restriction sites was introduced to the 5'-end of the corrected construct by replacement of the SacI restriction site on the MNase-F primer (MNL-F, Table 2.1). Positive transformants were identified by the release of a fragment of approximately 1200bp following NdeI/XhoI double digestion (Lane 5, Fig. 4.7). The *lacUV5* promoter was also inserted into the expression vector along with the full construct.

Partial nucleotide sequencing of the positive pCFCMNL further confirmed the removal of the point mutation (Fig. 4.8). The translated amino acid sequence was observed to perfectly align to that of the expected sequence, with the expected addition of a glycine and a serine residue at the BamHI restriction site between the MNase and LexA encoding sequence fragments (Fig. 4.9).

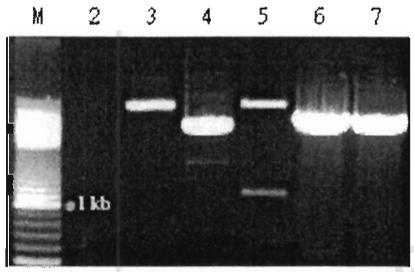


Fig. 4.7. Restriction digestion screening of the positive fusion construct inserted into pET20b-(+). Marker (lane M). 100bp step ladder (Promega). The positive plasmid carrying the corrected full construct was identified by the release of an approximately 1200bp fragment (lane 5).

```
pCFCMNL T7          1       AGAAATAATTCTCTGTTTAACTTTAAGAAGGAGATATACATATGATGCCAAAG 53
Predicted construct 1                                                 ATGCCAAAG 9
pCFCMNL T7          54      AAGAGAAAGGTTCCCGGGTTAATTAACGCAACTTCAACTAAAAAATTACATAA 106
Predicted construct 10      AAGAGAAAGGTTCCCGGGTTAATTAACGCAACTTCAACTAAAAAATTACATAA 62
pCFCMNL T7          107     AGAACCTGCGACTTTAATTAAAGCGATTGATGGTGATACGGTTAAATTAATGT 159
Predicted construct 63      AGAACCTGCGACTTTAATTAAAGCGATTGATGGTGATACGGTTAAATTAATGT 115
pCFCMNL T7          160     ACAAAGGTCAACCAATGACATTCAGACTATTATTGGTCGACACACCTGAAACA 212
Predicted construct 116     ACAAAGGTCAACCAATGACATTCAGACTATTATTGGTCGACACACCTGAAACA 168
pCFCMNL T7          213     AAGCATCCTAAAAAAGGTGTAGAGAAATATGGTCCTGAAGCAAGTGCATTTAC 265
Predicted construct 169     AAGCATCCTAAAAAAGGTGTAGAGAAATATGGTCCTGAAGCAAGTGCATTTAC 221
pCFCMNL T7          266     CAAAAAAATGGTAGAAAATGCAAAGAAAATTGAAGTCGAATTCGACAAAGGTC 318
Predicted construct 222     CAAAAAAATGGTAGAAAATGCAAAGAAAATTGAAGTCGAATTCGACAAAGGTC 274
pCFCMNL T7          319     AAAGAACTGATAAATATGGACGTGGGCTAGCGTATATTTATGCTGATGGAAAA 371
Predicted construct 275     AAAGAACTGATAAATATGGACGTGGGCTAGCGTATATTTATGCTGATGGAAAA 327
pCFCMNL T7          372     ATGGTAAACGAAGCTTTAGTTCGTCAAGGCTTGGCTAAAGTTGCTTATGTTTA 424
Predicted construct 328     ATGGTAAACGAAGCTTTAGTTCGTCAAGGCTTGGCTAAAGTTGCTTATGTTTA 380
pCFCMNL T7          425     CAAACCTAACAATACACATGAACAACATTTAAGAAAAAGTGAAGCACAAGCGA 477
Predicted construct 381     CAAACCTAACAATACACATGAACAACATTTAAGAAAAAGTGAAGCACAAGCGA 433
pCFCMNL T7          478     AAAAAGAGAAATTAAATATTTGGAGCGAAGACAACGCTGATTCAGGTCAAGGA 530
Predicted construct 434     AAAAAGAGAAATTAAATATTTGGAGCGAAGACAACGCTGATTCAGGTCAA--- 483
pCFCMNL T7            531       TCCATGAAAGCGTTAACGGCCAGG     554
Predicted construct 484         TAAATGAAAGCGTTAACGGCCAGG     508
pCFCMNL T7terminator  1         CAACAAGAGGTGTTTGATCTCATCCGTGATCACATCAGCCAGA 43
Predicted construct     509     CAACAAGAGGTGTTTGATCTCATCCGTGATCACATCAGCCAGA 550
pCFCMNL T7terminator  44        CAGGTATGCCGCCGACGCGTGCGGAAATCGCGCAGCGTTTGGGGTTCCGT 93
Predicted construct     551     CAGGTATGCCGCCGACGCGTGCGGAAATCGCGCAGCGTTTGGGGTTCCGT 600
pCFCMNL T7terminator  94        TCCCCAAACGCGGCTGAAGAACATCTGAAGGCGCTGGCACGCAAAGGCGT 143
Predicted construct     601     TCCCCAAACGCGGCTGAAGAACATCTGAAGGCGCTGGCACGCAAAGGCGT 650
pCFCMNL T7terminator  144       TATTGAAATTGTTTCCGGCGCATCACGCGGGATTCGTCTGTTGCAGGAAG 193
Predicted construct     651     TATTGAAATTGTTTCCGGCGCATCACGCGGGATTCGTCTGTTGCAGGAAG 700
pCFCMNL T7terminator  194       AGGAAGAAGGGTTGCCGCTGGTAGGTCGTGTGGCTGCCGGTGAACCACTT 243
Predicted construct     701     AGGAAGAAGGGTTGCCGCTGGTAGGTCGTGTGGCTGCCGGTGAACCACTT 750
pCFCMNL T7terminator  244       CTGGCGCAACAGCATATTGAAGGTCATTATCAGGTCGATCCTTCCTTATT 293
Predicted construct     751     CTGGCGCAACAGCATATTGAAGGTCATTATCAGGTCGATCCTTCCTTATT 800
pCFCMNL T7terminator  294       CAAGCCGAATGCTGATTTCCTGCTGCGCGTCAGCGGGATGTCGATGAAAG 343
Predicted construct     801     CAAGCCGAATGCTGATTTCCTGCTGCGCGTCAGCGGGATGTCGATGAAAG 850
pCFCMNL T7terminator  344       ATATCGGCATTATGGATGGTGACTTGCTGGCAGTGCATAAAACTCAGGAT 393
Predicted construct     851     ATATCGGCATTATGGATGGTGACTTGCTGGCAGTGCATAAAACTCAGGAT 900
```

```
pCFCMNL T7terminator  394    GTACGTAACGGTCAGGTCGTTGTCGCACGTATTGATGACGAGGTTACCGT 443
Predicted construct    901   GTACGTAACGGTCAGGTCGTTGTCGCACGTATTGATGACGAAGTTACCGT 950
pCFCMNL T7terminator  444    TAAGCGCCTGAAAAAACAGGGCAATAAAGTCGAACTGTTGCCAGAAAATA 493
Predicted construct    951   TAAGCGCCTGAAAAAACAGGGCAATAAAGTCGAACTGTTGCCAGAAAATA 1000
pCFCMNL T7terminator  494    GCGAGTTTAAACCAATTGTCGTAGATCTTCGTCAGCAGAGCTTCACCATT 543
Predicted construct    1001  GCGAGTTTAAACCAATTGTCGTAGATCTTCGTCAGCAGAGCTTCACCATT 1050
pCFCMNL T7terminator  544    GAAGGGCTGGCGGTTGGGGTTATTCGCAACGGCGACTGGCTGAATCTAGA 593
Predicted construct    1051  GAAGGGCTGGCGGTTGGGGTTATTCGCAACGGCGACTGGCTGAATCTAGA 1100
pCFCMNL T7terminator  594    AAATTATTGTCTAACAATAATTCCACACATTATACGAGCCGGAAGCATAA 643
Predicted construct    1101  AAATTATTGTCTAACTATAATTCCACACATTATACGAGCCGGAAGCATAA 1150
pCFCMNL T7terminator  644    AGTGTAAAGCCTGGGGTGCCTAATGAGTGACGCTCGAGCACCACCACCAC 693
Predicted construct    1151  AGTGTAAAGCCTGGGGTGCCTAATGAGTGACGCTCGAG        1188
```

Fig. 4.8. Sequence alignment of expected and predicted recombinant construct inserted into pET20b-(+). The point insertion was removed from the recombinant construct at position corresponding to construct at position 992 of the expected sequence. The recombinant construct sequence was derived by joining sequence acquired from the forward and reverse sequencing reactions at position 508 of the expected sequence.

```
Actual construct      1   MPKKRKVPGLINATSTKKLHKEPATLIKAIDGDTVKLMYKGQPMTFRLLLVDTPE 55
Predicted construct   1   MPKKRKVPGLINATSTKKLHKEPATLIKAIDGDTVKLMYKGQPMTFRLLLVDTPE 55
Actual construct      56  TKHPKKGVEKYGPEASAFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGKM 110
Predicted construct   56  TKHPKKGVEKYGPEASAFTKKMVENAKKIEVEFDKGQRTDKYGRGLAYIYADGKM 110
Actual construct      111 VNEALVRQGLAKVAYVYKPNNTHEQHLRKSEAQAKKEKLNIWSEDNADSGQGSMK 165
Predicted construct   111 VNEALVRQGLAKVAYVYKPNNTHEQHLRKSEAQAKKEKLNIWSEDNADSGQ--MK 163
Actual construct      166 ALTARQQEVFDLIRDHISQTGMPPTRAEIAQRLGFRSPNAAEEHLKALARKGVIE 220
Predicted construct   164 ALTARQQEVFDLIRDHISQTGMPPTRAEIAQRLGFRSPNAAEEHLKALARKGVIE 218
Actual construct      221 IVSGASRGIRLLQEEEEGLPLVGRVAAGEPLLAQQHIEGHYQVDPSLFKPNADFL 275
Predicted construct   219 IVSGASRGIRLLQEEEEGLPLVGRVAAGEPLLAQQHIEGHYQVDPSLFKPNADFL 273
Actual construct      276 LRVSGMSMKDIGIMDGDLLAVHKTQDVRNGQVVVARIDDEVTVKRLKKQGNKVEL 330
Predicted construct   274 LRVSGMSMKDIGIMDGDLLAVHKTQDVRNGQVVVARIDDEVTVKRLKKQGNKVEL 328
Actual construct      331 LPENSEFKPIVVDLRQQSFTIEGLAVGVIRNGDWLNLENYCLTIIPHIIRAGSIK 385
Predicted construct   329 LPENSEFKPIVVDLRQQSFTIEGLAVGVIRNGDWLNLENYCLTIIPHIIRAGSIK 383
Actual construct      386 CKAWGA 391
Predicted construct   384 CKAWGA 389
```

Fig. 4.9. Amino acid sequence alignment of expected and predicted construct. The glycine and serine (highlighted in green) inserted in the actual construct resulted from a BamHI restriction site joining the MNase component to the LexA component of the recombinant protein

4.5. Future work

The DNA fragment encoding the recombinant MNase-LexA fusion protein was shown to have the correct sequence. The recombinant construct is therefore ready to be expressed, and the feasibility of this approach to investigate chromatin compaction at specific target sites in living yeast cells tested. The ability of the recombinant MNase-LexA fusion to act accurately is dependent on proper anchoring of the fusion protein to LexA operator sequence, and effective cleavage activity of the MNase active domain. Regularly-spaced reconstituted nucleosomal arrays and identical, free

DNA provide ideal substrates for the testing the ability of the fusion protein to sense compaction. Due to the toxicity of the fusion protein to host cells following induction, expression and initial study of the fusion protein may be problematic using whole cell expression systems. This can be solved by using cell-free *in vitro* transcription/translation systems. Cloning of the construct into vector pET20b-(+) allows for easy expression of the fusion construct using a bacterial cell-free *in vitro* expression system, since pET20-(+) contains the necessary elements for gene expression in bacteria. The fusion protein could be tested on a reconstituted nucleosomal array and the corresponding free DNA fragment containing LexA operator sequences. Additional controls to test for any sequence-specific activity of the MNase domain of the recombinant fusion protein would also be conducted, using reconstituted nucleosomal array and the corresponding free DNA lacking the LexA operator sequence. An analysis of the range of DNA fragments generated by the active MNase-LexA fusion protein, prepared in a cell-free transcription/translation mix, would provide proof of concept for this approach to study chromatin compaction. This would then allow induced expression of the fusion protein *in vivo* to study the compaction status of chromatin in living yeast cells at specific loci where the LexA operator sequences had been inserted. Testing and optimization of this approach remains to be undertaken.

Chapter 5

Discussion

5.1. Compaction of chromatin in Saccharomyces cerevisiae

The yeast linker histone homolog, Hho1p, has been shown to be non-essential for cell viability and mutant yeast lacking Hho1p does not display any traceable phenotype (Patterton *et al.*, 1998). The role of Hho1p as a pure architectural component of the yeast chromosomes has been recently suggested and it was clear that the yeast genome was more decondensed in the absence of Hho1p during the stationary phase growth, and Hho1p bound chromatin assumed a more compact structure (Schafer and Patterton, unpublished data). The microarray data from this study indicated that the yeast genome is generally decondensed during exponential phase growth, and a higher level of compaction of the yeast genome occurs during stationary phase growth. In a previous study by Lohr and Hereford (1979) it was shown that chromatin of exponential phase *S. cerevisiae* was equally sensitive towards DNaseI, and this uniform sensitivity decreased in a uniform fashion when the yeast enters stationary phase. The uniform decrease of DNaseI sensitivity of the *S. cerevisiae* genome may be a result of restricted DNaseI access to the yeast genome by the global chromatin compaction induced by Hho1p association to the genome. The data of this study supported the postulation of a link between yeast genome condensation and Hho1p binding where Hho1p could contribute directly towards the genome condensation. The investigation into the relationship between the compaction states of genes at different growth phases did not reveal any clear correlation. Since the entire genome is compacted during stationary phase, it was unlikely that the genome would assume a

compaction structure during exponential growth when active transcription needs to take place.

The establishment of Hho1p-dependent compaction of the yeast genome during stationary phase growth lead to the thinking of whether the compaction occurred in a random non-specific manner, or clusters of genes in close proximity of each other would aid the compaction. The previous study by Gilbert *et al.* (2004) reported that high gene density areas on a chromosome corresponded to regions of open chromatin fibre, and no correlation could be made between open chromatin of high gene density to the expression of the genes within the open chromatin. We performed a similar analysis by comparing the average gene content within condensed chromatin to the gene density at the 100kb resolution. Unlike the study performed by Gilbert *et al.*, data from this study could not establish a clear correlation between chromatin compaction and gene density (Fig. 3.9). This data further suggested that Hho1p-dependent chromatin compaction occurs as a global event, rather than confined to distinct regions on the genome.

The attempt to relate the genes presented in the compacted fraction of exponential phase chromatin to their proposed expression revealed several genes whose compaction state displayed a contrary relationship to the general thinking that chromatin condensation is associated with transcription activity. Specifically, a group of genes (*UTP9, NOP4, NOP7, SRD1*) that encode for products involved in rRNA processing was of particular interest, as their transcription activities occur prior to diauxic shift, yet their ORF was enriched in the condensed fraction of the chromatin. From the study by Koop *et al.* (2003), linker histone H1 was shown to induced nucleosome positioning at the MMTV promoter that was able to enhance synergism

between two transcription factors, and this synergism between PR and NF1 resulted in full activation of the MMTV promoter. The finding from this study suggested a similar mechanism might be in place to regulate the expression of genes encoding product involved in rRNA processing.

The microarray data from a *HHO1* deletion study on *S. cerevisiae* reported that the expression of the majority of genes were mostly unaffected by the absence of Hho1p, and a modest reduction in mRNA level was observed for certain genes (Hellauer *et al.*, 2001). The data from the *HHO1* deletion study indicated that for the list of genes whose promoters were identified to be more compacted during exponential phase growth (Table 3.2), five out of the 49 genes showed a constant expression level in the absence of Hho1p, while the rest displayed minor decreases to their mRNA level. The modest decrease in the mRNA level of these genes suggested that Hho1p-dependent chromatin compaction might indeed exert an effect on the expression of genes, but not to a great extent.

5.2. Novel recombinant protein capable of chromatin compaction detection

The recombinant fusion protein constructed in this study provides a complete new approach to the measurement of chromatin compaction as it offers a mean to directly measure chromatin compaction. The strategy of employing tethered enzymes fused to DNA-binding proteins has resulted in the successful development of several molecular tools that are invaluable to *in vivo* investigations of DNA-protein interactions. van Steensel and Henikoff (2000) successfully demonstrated that the attachment of the *Escherichia coli* DNA adenine methyltransferase (Dam) to the DNA-binding domain of Gal4 in *Drosophila* allowed for targeted DNA methylation at

the region of Gal4 binding sequence. They further identified several known and novel loci that interact with *Drosophila* heterochromatin protein 1 employing the tethered Dam. The demonstration of the high sensitivity, as well as providing high resolution, of a cytosine-5 DNA methyltransferase, M.*Cvi*PI, fused to Pho4 further emphasized the value of tethered methyltransferase in the detection of DNA-protein interaction (Carvin *et al.*, 2003). The generation of DNA breakage as marker signals by MNase active domain possesses several advantages over other marker signals. Proper MNase action is solely dependent on the Ca^{2+} concentration and does not involve complex activation pathways which might exert any undesirable effect on the structure of the targeted chromatin region. The ability of the recombinant protein to be easily and accurately targeted to specific genes allow for future high resolution study of chromatin compaction.

The proof of concept of the recombinant protein involves the confirmation of the two main properties, the ability for the recombinant protein to bind properly to LexA operons, and the ability to cleave DNA. The efficiency of the recombinant protein to perform both functions can be concurrently tested on reconstituted nucleosomal array containing LexA operon sequences. The visualisation of the outcome of the recombinant protein action can be accurately achieved with the aid of end labelling experiments. From the *in vitro* testing, the generation of longer linear chromatin fragments is expected from reconstituted nucleosomal array treated with the recombinant protein than that from naked DNA containing LexA operons, which is a result of different degree of compaction of the DNA fragments.

The recombinant protein designed is essentially a nuclease with anchoring ability that is targeted to the nucleus by a nuclear localization signal. The nuclease activity of the

recombinant protein rendered the recombinant protein toxic to the host cell, and thus presents a challenge for its *in vivo* application. For future *in vivo* application of the recombinant protein, expression of the recombinant protein needs to be placed under an effective control, such as a highly stringent inducible promoter. The induction method of choice also needs to be well considered, as the induction of the recombinant protein must also not alter either the physiological state of the cell or the local chromatin environment that is under investigation.

This study attempted at bridging the knowledge of Hho1p binding and the compaction of the yeast genome. Although the data suggested possible involvement of Hho1p-dependent compaction of the chromatin in transcription regulation of specific genes, a clear correlation between gene activity and chromatin condensation state could not be established. The attempt to elucidate a relationship between compaction of chromatin and gene densities did not reveal any clear insights into the effect of possible gene clusters that might aid chromatin compaction. The data also showed that in general the yeast genome is decondensed during exponential growth and condensed during stationary growth, in addition there exist several genes whose compaction state of the promoter regions does not correspond to their expected transcription activities. Further detailed study of these genes, with the aid of the recombinant protein developed in this study, would yield added worthwhile insights in establishing a solid link between chromatin compaction and transcription activity.

Reference List

Allfrey,V.G., Faulkner,R., and Mirsky,A.E. (1964). Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. Proc. Natl. Acad. Sci. USA *51*, 786-794.

Bauer,U.-M., Daujat,S., Nielsen,S.J., Nightingale,K., and Kouzarides,T. (2002). Methylation at arginine 17 of histone H3 is linked to gene activation. EMBO reports *3*, 39-44.

Boyer,L.A., Logie,C., Bonte,E., Becker,P.B., Wade,P.A., Wolffe,A.P., Wu,C., Imbalzano,A.N., and Peterson,C.L. (2000). Functional delineation of three groups of the ATP-dependent family of chromatin remodeling enzymes. J. Biol. Chem. *275*, 18864-18870.

Braunstein,M., Sobel,R.E., Allis,C.D., Turner,B.M., and Broach,J.R. (1996). Efficient transcriptional silencing in Saccharomyces cerevisiae requires a heterochromatin histone acetylation pattern. Mol. Cell. Biol. *16*, 4349-4356.

Brehm,A., Langst,G., Kehle,J., Clapier,C.R., Imhof,A., Eberharter,A., Muller,J., and Becker,P.B. (2000). dMi-2 and ISWI chromatin remodeling factors have distinct nucleosome binding and mobilization properties. EMBO J. *19*, 4332-4341.

Carmen,A.A., Milne,L., and Grunstein,M. (2002). Acetylation of the yeast histone H4 N terminus regulates its binding to heterochromatin protein SIR3. J. Biol. Chem. *277*, 4778-4781.

Carruthers,L.M. and Hansen,J.C. (2000). The core histone N termini function independently of linker histones during chromatin condensation. J. Biol. Chem. *275*, 37285-37290.

Carvin,C.D., Dhasarathy,A., Friesenhahn,L.B., Jessen,W.J., and Kladde,M.P. (2003). Targeted cytosine methylation for *in vivo* detection of protein-DNA interactions. Proc. Natl. Acad. Sci. USA *100*, 7743-7748

Cosgrove,M.S., Boeke,J.D., and Wolberger,C. (2004). Regulated nucleosome mobility and the histone code. Nat. Struct. Biol. *11*, 1037-1043.

Cotton,F.A., Hazen Jr.,E.E., and Legg,M.J. (1979). Staphylococcal nuclease: Proposed mechanism of action based on structure of enzyme-thymidine 3',5'-bisphosphate-calcium ion complex at 1.5-A resolution. Proc. Natl. Acad. Sci. *76*, 2551-2555.

Croce,L.D., Koop,R., Venditti,P., Westphal,H.M., Nightingale,K., Becker,P., and Beato,M. (1999). Two-steps synergism between progesterone receptor and the DNA binding domain of NF1 on MMTV minichromosomes. Mol. Cell *4*, 45-54.

Croston,G.E., Kerrigan,L.A., Lira,L.M., Marshak,D.R., and Kadonaga,J.T. (1991). Sequence-specific antirepression of histone H1-mediated inhibition of basal RNA polymerase II transcription. Science *251*, 643-649.

Cuthbert,G.L., Daujat,S., Snowden,A.W., Erdjument-Bromage,H., Hagiwara,T., Yamada,M., Schneider,R., Gregory,P.D., Tempst,P., Bannister,A.J., and Kouzarides,T. (2004). Histone deimination antagonizes arginine methylation. Cell *118*, 545-553.

Davey,C.A., Sargent,D.F., Luger,K., Maeder,A.W., and Richmond,T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J. Mol. Biol. *319*, 1097-1113.

DeRisi,J.L., Iyer,V.R., and Brown,P.O. (1997). Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale. Science *278*, 680-686.

Dumoulin,P., Oertel-Buchheit,P., Granger-Schnarr,M., and Schnarr,M. (1993). Orientation of the LexA DNA-binding motif on operator DNA as inferred from cysteine-mediated phenyl azide crosslinking. Proc. Natl. Acad. Sci. USA *90*, 2030-2034.

Emre,N.C., Ingvarsdottir,K., Wyce,A., Wood,A., Krogan,N.J., Henry,K.W., Li,K., Marmorstein,R., Greenblatt,J.F., Shilatifard,A., and Berger,S.L. (2005). Maintenance of low histone ubiquitylation by Ubp10 correlates with telomere-proximal Sir2 association and gene silencing. Cell *17*, 585-594.

Fan,Y., Nikitina,T., Morin-Kensicki,E.M., Zhao,J., Magnuson,T.R., Woodcock,C.L., and Skoultchi,A.I. (2003). H1 linker histone are essential for mouse development and affect nucleosome spacing *in vivo*. Mol. Cell. Biol. *23*, 4559-4572.

Feng,S.-Y., Ota,K., Yamada,Y., Sawabu,N., and Ito,T. (2004). A yeast one-hybrid system to detect methylation-dependent DNA-protein interactions. Biochem. Biophys. Res. Commun. *313*, 992-925.

Feng,X., Zhang,K., and Grunstein,M. (2005). Acetylation in histone H3 globular domain regulates gene expression in yeast. Cell *121*, 375-385.

Finch,J.T. and Klug,A. (1976). Solenoidal model for superstructure in chromatin. Proc. Natl. Acad. Sci. USA *73*, 1897-1901.

Flaus,A. and Owen-Hughes,T. (2001). Mechanisms for ATP-dependent chromatin remodeling. Curr. Opin. Genet. Dev. *11*, 148-154.

Fogh,R.H., Ottleben,G., Ruterjans,H., Schnarr,M., Boelens,R., and Kaptein,R. (1994). Solution structure of the LexA repressor DNA binding domain determined by ^{1}H NMR spectroscopy. EMBO J. *13*, 3936-3944.

Francis,N.J., Kingston,R.E., and Woodcock,C.L. (2004). Chromatin compaction by a Polycomb group protein complex. Science *306*, 1574-1577.

Georgel,P.T., Horowits-Scherer,R.A., Adkins,N., Woodcock,C.L., Wade,P.A., and Hansen,J.C. (2003). Chromatin compaction by Human MeCP2. J. Biol. Chem. *278*, 32181-32188.

Ghirlando,R., Litt,M.D., Prioleau,M.-N., Recillas-Targa,F., and Felsenfeld,G. (2004). Physical properties of a genomic condensed chromatin fragment. J. Mol. Biol. *336*, 597-605.

Gilbert,N., Boyle,S., Fiegler,H., Woodfire,K., Carter,N.P., and Bickmore,W.A. (2004). Chromatin architecture of the human genome: gene rich domains are enriched in open chromatin fibres. Cell *118*, 555-566.

Harvey,A.C. and Downs,J.A. (2004). What functions do linker histones provide? Mol. Micro. *53*, 771-775.

Hayashi,T., Hayashi,H., and Iwai,K. (1987). Tetrahymena histone H1. Isolation and amino acid sequence lacking the central hydrophobic domain conserved in other H1 histones. J. Biochem. *102*, 369-376.

Hayes,J.J. and Hansen,J.C. (2001). Nucleosomes and the chromatin fibre. Curr. Opin. Genet. Dev. *11*, 124-129.

Hegde,P., Abernathy,K., Gay,C., Dharap,S., Gaspard,R., Hughes,J.E., Snesrud,E., Lee,N., and Quackenbush,J. (2000). A concise guide to cDNA microarray analysis. Biotechniques *29*, 548-562.

Hellauer,K., Sirard,E., and Turcotte,B. (2001). Decreased expression of specific genes in yeast cells lacking histone H1. J. Biol. Chem. *276*, 13587-13592.

Hendzel,M.J., Lever,M.A., Crawford,E., and Th'ng,P.H. (2004). The C-terminal domain is the primary determinant of histone H1 binding in chromatin *in vivo*. J. Biol. Chem. *279*, 20028-20034.

Jenuwein,T. and Allis,C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Kadosh,D. and Struhl,K. (1997). Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. Cell *89*, 365-371.

Kadosh,D. and Struhl,K. (1998a). Histone deacetylase activity of Rdp3 is important for transcriptional repression in vivo. Genes. Dev. *12*, 797-805.

Kadosh,D. and Struhl,K. (1998b). Targeted recruitment of the Sin3-Rpd3 histone deacetylase complex generates a higly localized domain of repressed chromatin in vivo. Mol. Cell. Biol. *18*, 5121-5127.

Khochbin,S. and Wolffe,A.P. (1994). Developmentally regulated expression of linker-histone variants in vertebrates. Eur. J. Biochem *225*, 501-510.

Koop,R., Croce,L.D., and Beato,M. (2003). Histone H1 enhances synergistic activation of the MMTV promoter in chromatin. EMBO J. *22*, 588-599.

Krebs,J.E., Fry,C.J., Samuels,M.L., and Peterson,C.L. (2000). Global role for chromatin remodeling enzymes in mitotic gene expression. Cell *102*, 587-598.

Krebs,J.E., Kuo,M.-H., Allis,C.D., and Peterson,C.L. (1999). Cell cycle-regulated histone acetylation required for expression of the yeast HO gene. Genes. Dev. *13*, 1412-1421.

Kurdistani,S.K. and Grunstein,M. (2003). Histone acetylation and deacetylation in yeast. Nat. Rev. Mol. Cell Biol. *4*, 276-284.

Lachner,M., O'Carroll,D., Rea,S., Mechtler,K., and Jenuwein,T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-120.

Lohr,D. and Hereford,L. (1979). Yeast chromatin is uniformlydigested by DNaseI. Proc. Natl. Acad. Sci. USA *76*, 4285-4288

Lucchesi,J.C. (1998). Dosage compensation in flies and worms: the ups and downs of X-chromosome regulation. Curr. Opin. Genet. Dev *8*, 179-184.

Luger,K., Mäder,A.W., Richmond,R.K., Sargent,D.F., and Richmond,T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 angstrom resolution. Nature *389*, 251-260.

Maile,T., Kwoczynski,S., Katzenberger,R.J., Wassarman,D.A., and Sauer,F. (2004). TAF1 activates transcription by phosphorylation of serine 33 in histone H2B. Science *304*, 1010-1014.

Margueron,R., Trojer,P., and Reinberg,D. (2005). The key to development: interpreting the histone code? Curr. Opin. Genet. Dev. *15*, 163-176.

Nakamura,Y., Gojobori,T., and Ikemura,T. (1999). Codon usage tabulated from the international DNA sequence databases; its status 1999. Nucleic Acids Res. *27*, 292.

Nehme,D., Li,X.-Z., Elliot,R., and Poole,K. (2004). Assembly of the MexAB-OprM Multidrug Efflux System of P*seudomonas aeruginosa:* Identification and Characterization of Mutations in *mexA* Compromising MexA Multimerization and Interaction with MexB. J. Bacteriol. *186*, 2973-2983.

Panetta,G., Buttinelli,M., Flaus,A., Richmond,T.J., and Rhodes,D. (1998). Differential nucleosome positioning on Xenopus oocyte and somatic 5 S RNA genes determines both TFIIIA and H1 binding: a mechanism for selective H1 repression. J. Mol. Biol. *282*, 683-697.

Patterton, H.G., Landel, C.C., Landsman, D., Peterson, C.L., and Simpson, R.T. (1998). The biochemical and phenotypic characterization of Hho1p, the putative linker histone H1 of Sa*ccharomyces cerevisiae*. J. Biol. Chem. *273*, 7268-7276.

Patterton,H.G. and Graves,S. (2000). DNAssist, a C++ program for editing and analysis of nucleic acid and protein sequences on PC-compatible computers running Windows 95, 98, NT4.0 or 2000. Biotechniques *6*, 1192-1197.

Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C., and Ferrin,T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605-1612.

Pringle, J.R., Broach,J.R., and Jones,E.W., eds. (1991). The molecular and cellular biology of the yeast *Saccharomyces*., Boston: Cold Spring Harbor Laboratory Press.

Rea,S., Eisenhaber,F., O'Carroll,D., Strahl,B.D., Sun,Z.-W., Schmid,M., Opravil,S., Mechtler,K., Ponting,C.P., Allis,C.D., and Jenuwein,T. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. Nature *406*, 593-599.

Richards,E.J. and Elgin,S.C.R. (2002). Epigenetic code for heterochromatin formation and silencing: rounding up the usual suspects. Cell *108*, 489-500.

Santos-Rosa,H., Bannister,A.J., Dehe,P.M., Geli,V.V., and Kouzarides,T. (2004). Methylation of H3 lysine 4 at euchromatin promotes Sir3p association with heterochromatin. J. Biol. Chem. *279*, 47506-47512.

Santos-Rosa,H., Schneider,R., Bernstein,B.E., Karabetsou,N., Morillon,A., Weise,C., Schreiber,S.L., Mellor,J., and Kouzarides,T. (2003). Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. Mol. Cell *12*, 1325-1332.

Sera,T. and Wolffe,A.P. (1998). Role of Histone H1 as an Architectural Determinant of Chromatin Structure and as a Specific Repressor of Transcription on X*enopus* Oocytes 5S rRNA Genes. Mol. Cell. Biol. *18*, 3668-3680.

Shen,X. and Gorovsky,M.A. (1996). Linker Histone H1 Regulates Specific Gene Expression but Not Global Transcription In V*ivo*. Cell *86*, 475-483.

Shen,X., Yu,L., Weir,J.W., and Gorovsky,M.A. (1995). Linker Histones Are Not Essential and Affect Chromatin Condensation In *Vivo*. Cell *82*, 47-56.

Simpson,R.T. (1978). Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. Biochemistry *25*, 5524-5531.

Strahl,B.D. and Allis,C.D. (2000). The language of covalent histone modifications. Nature *403*, 41-45.

Struhl,K. (1998). Histone acetylation and transcriptional regulatory mechanisms. Genes Dev. *12*, 599-606.

Sun,Z.-W. and Allis,C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature *418*, 104-108.

Thoma,F., Koller,T., and Klug,A. (1979). Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. J. Cell Biol. *83*, 403-427.

Vignali,M., Hassan,A.H., Neely,K.E., and Workman,J.L. (2000). ATP-Dependent Chromatin-Remodeling Complexes. Mol. Cell. Biol. *20*, 1899-1910.

van Steensel.B. and Henikoff.S. (2000). Identification of *in vivo* DNA targets of chromatin protein using tethered Dam methyltransferase. Nat. Biotechnol. *18*, 424-428

Vogelauer.M., Wu.J., Suka.N., and Grunstein.M. (2000). Global histone acetylation and deacetylation in yeast. Nature *408*, 495-498.

Wang.X. and Simpson,R.T. (2001). Chromatin structure mapping in *Saccharomyces cerevisiae in vivo* with DNaseI. Nucleic Acids Res. *29*, 1943-1950.

Wang,Y., Wysocka,J., Sayegh,J., Lee,Y.-H., Perlin,J.R., Leonelli,L., Sonbuchner,L.S., McDonald,C.H., Cook,R.G., Dou,Y., Roeder,R.G., Clarke,S., Stallcup,M.R., Allis,C.D., and Coonrod,S.A. (2004). Human PAD4 regulates histone arginine methylation levels via demethylimination. Science *306*, 279-283.

Wei,Y., Yu,L., Bowen,J., Gorovsky,M.A., and Allis,C.D. (1999). Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. Cell *97*, 99-109.

Wodicka,L., Dong,H., Mittmann,M., Ho,M.H., and Lockhart,D.J. (1997). Genome-wide expression monitoring in Saccharomyces cerevisiae. Nat. Biotech. *15*, 1359-1367.

Wolffe,A.P. (1998). Chromatin structure and function. (London: Academic Press.).

Wolffe,A.P. (2001). Transcriptional regulation in the context of chromatin structure. Essays Biochem. *37*, 45-57.

Wu,M., Allis,C.D., Richman,R., Cook,R.G., and Gorovsky,M.A. (1986). An intervening sequence in an unusual histone H1 gene of Tetrahymena thermophila. Proc. Natl. Acad. Sci. USA *22*, 8674-8678.

Zhang,H., Pan,K.-H., and Cohen,S.N. (2003). Senescence-specific gene expression fingerprints reveal cell-type-dependent physical clustering of up-regulated chromosomal loci. Proc. Natl. Acad. Sci. USA *100*, 3251-3256.