

NEGATIVE REGULATORS OF GENE EXPRESSION IN YEAST: a1/α2 AND SIR

by

Allan Malcolm Miller

SUMMARY

This dissertation describes work investigating two different repressors of transcription in Saccharomyces cerevisiae. In the first system, the HML and HMR loci are repressed by the action of the four gene products, SIR1-4. Deletion analysis of HML and HMR revealed that the DNA sequences that are required in cis are located up to 1700 base pairs from the affected promoters, and will in fact repress a promoter 2600 bp away. Thus SIR-mediated repression appears to be unlike any other characterised system of repression, in that it is capable of action at a distance. The deletion analyses also implied that these DNA sequences are tightly linked to putative origins of DNA replication. I have looked at the role of DNA replication in SIR-mediated repression. If the silent loci are derepressed by growing strains carrying temperature-sensitive mutations in SIR3 or SIR4 at the restrictive temperature, then shifted to the permissive temperature, I have found that the ensuing repression is blocked by inhibitors of DNA replication, but not by an inhibitor of mitosis. This suggests that DNA replication may be necessary for the establishment of the repressed state, and possible mechanisms are considered. When a sir3^{ts} strain is shifted from the permissive temperature to the restrictive temperature, however, the silent loci can be switched on (derepressed) in the absence of DNA replication, implying that the action of the SIR3 gene product is not restricted to S phase.

The second system of repression that I have studied is a1/α2-mediated repression. The MATα2 gene product represses the transcription of one set of genes, but if both MATα2 and MATα1 gene products are present within the same cell, then an additional set of genes is repressed. I have shown that this "a1/α2-mediated repression" can be established in the absence of DNA replication, in contrast to SIR-mediated repression. I have also identified, using the technique of DNA sequence comparisons, the DNA sequence motifs that are the targets for α2- and a1/α2-mediated repression. The functions of the sequences thus identified were confirmed by inserting DNA fragments containing these sequences into the promoter of an unrelated yeast gene, CYC1. The insertions were able to bring the CYC1 promoter under the appropriate control (α2-mediated or a1/α2-mediated).

The DNA sequences responsible for a1/α2-mediated repression do not seem to be able to act at a distance, in contrast to the situation with SIR-mediated repression. This suggests that there may be two distinct forms of repression of transcription in yeast, one involving DNA replication and acting over considerable distances, and another, which acts only locally and does not involve DNA replication.

PHD 14474

NEGATIVE REGULATORS OF GENE EXPRESSION IN YEAST:

a1/α2 and SIR

by



Allan Malcolm Miller

A dissertation submitted to the University of
Cambridge for examination for the degree of
Doctor of Philosophy.

July 1986

Clare College

For Mum and Dad

DECLARATION

This dissertation is entirely the result of my own work and includes nothing which is the outcome of work done in collaboration. It has not been submitted, either in whole or in part, for a degree or qualification at any other University.

Allan Malcolm Miller

ACKNOWLEDGEMENTS

Just as writing a thesis is tough, so trying to list the many people to whom I owe so much is tougher.

Undoubtedly, my thanks go primarily to all the inhabitants of the cell biology division for the continual encouragement and incisive advice they have offered me; in particular I would like to thank Andy Brand, David Shore, Linda Breeden, Rolf Sternglanz, David Stillman, Gustav Ammerer, Hugh Pelham, Mariann Bienz and of course my thesis advisor Kim, whose vigilant supervision has been essential.

Dexterous technical assistance was provided by Mike Squire and Kim Nasmyth.

I would also like to thank all the kitchen ladies and secretaries of the second floor for their cheerful and efficient support.

This work was supported financially by a Medical Research Council Studentship, and a Research Fellowship at Clare College, and in addition by the generosity of my parents.

However, this thesis would never have been written but for the unethical pursuit of hard cash by the Nestle Company and Philip Morris Inc.

Lastly I would like to thank those people who have taken on the task of putting up with me at one time or another, in particular my landlord Ian. I guess Nick and Gos belong here too.

ABBREVIATIONS

Units

g	gram
mg	milligram
µg	microgram
l	litre
ml	millilitre
µl	microlitre
M	molar
mM	millimolar
bp	base pair
kb	kilobase

Chemicals

DNA	Deoxynucleic acid
RNA	Ribonucleic acid
mRNA	messenger RNA
A	Adenine
C	Cytosine
G	Guanine
T	Thymine
DMSO	Dimethyl sulphoxide
EDTA	Ethylenediamine tetra-acetic acid
HU	Hydroxyurea
MBC	Methyl-benzimidazol-2-yl carbamate
Tris	Tris (hydroxymethyl) aminomethane
SDS	Sodium dodecyl sulphate

Miscellaneous

- UAS Upstream Activating Sequence
- ARS Autonomously replicating sequence
- O.D. Optical density
- Δ deletion
- Lk Linking Number
- Tw Twist
- Wr Writhe
- Δ Lk Change in Linking Number
- Δ Tw Change in Twist
- Δ Wr Change in Writhe

CONTENTS

Title Page		
Dedication		
Declaration		
Acknowledgements		
Abbreviations		
Contents		
		Page
CHAPTER 1	Introduction	1
1.1	The Control of Gene Expression	1
1.2	Control of Transcription, Yeast as a Model System	4
1.3	Control of Cell type in Yeast	5
1.3.1	The <u>MAT</u> Genes are Master Regulators	6
1.3.2	Mating type Switching	8
1.3.3	<u>SIR</u> Repression	11
CHAPTER 2	Materials and Methods	13
2.1	Bacterial Growth Media	13
2.2	<u>E.Coli</u> Transformation	13
2.3	Preparation of DNA	13
2.4	Gel Electrophoresis	14
2.5	Enzymes and Chemicals	14
2.6	Yeast Media and Transformation	15
2.7	Yeast Plasmid Preparation and Supercoiling Gels	15
2.8	Southern Transfer	16
2.9	DNA Sequencing	16

2.10	Preparation of DNA probes for	
	S1 Nuclease Protection	16
2.11	RNA Preparation and S1 analysis	17
2.12	Construction of isogenic strains by	
	DNA Transformation	18
2.13	Preparation of G ₀ Daughters and	
	Release from the Alpha Factor Block	18
2.14	Preparation of Alpha Factor	19
2.15	Sequence Comparisons	19
2.16	Sequencing <u>STE2</u>	20
2.17.1	Cloning of <u>a1/α2</u> elements into	
	the <u>CYC1</u> promoter	20
2.17.2	Inserting a Fragment of <u>STE2</u> into	
	the <u>CYC1</u> promoter	21
2.18	Beta-galactosidase Assays	21
2.19	Cloning of synthetic oligonucleotides	21
2.20	Cloning <u>MATα</u> alleles from M30, 5-9, and 8-2	22
CHAPTER 3	Assays for <u>SIR</u> Repression	24
3.1	Introduction	24
3.2	Linking Number as an Assay for Sir status	25
3.2.1	Background	25
3.2.2	Linking number of plasmids in <u>sir^{ts}</u> strains	27
3.3	RNA levels as an assay for Sir status	28
3.4	Mapping the intron-exon boundaries	31
3.5	Half-life of the mature message	32
3.6	Investigation of the 400 nucleotide fragment	33
3.7	Discussion on Splicing	34
3.8	Conclusions	36

CHAPTER 4	Role of DNA Replication in <u>SIR</u> Repression	38
4.1	Introduction	38
4.2	Repression of <u>HML</u> and <u>HMR</u> requires cell cycle progress	39
4.3	Repression is Blocked by Inhibitors of DNA Synthesis, but not by Inhibitors of Mitosis	42
4.4	<i>Cell Cycle Progress</i> is not required for derepression	44
4.5	DNA Replication is not required to repress <u>MATα1</u>	45
4.6	Discussion	46
CHAPTER 5	Identification and Comparison of Two Cell-type Specific Operator Sequences	51
5.1	Introduction	
5.1.1	Background	51
5.1.2	<u>α1/α2</u> control of the <u>HO</u> gene	53
5.2	The Search for a Repeated Sequence	55
5.3	The Repeated Sequence is Sufficient to Confer <u>α1/α2</u> Control	57
5.4	The Search of Other Genes	58
5.5	Effect of an Element, when Placed outside of the Promoter	60
5.6	<u>α2</u> -mediated Repression Recognizes a Different Sequence	62
5.7	Discussion	65
5.7.1	Summary of Results	65
5.7.2	Possible mechanisms of repression	67

5.7.3	The <u>a1/α2</u> and <u>α2</u> elements have related, but distinct sequences	68
5.7.4	Possible Mechanism of <u>a1</u> gene product action	69
CHAPTER 6	Isolation of Mutations in <u>MATα2</u>	73
6.1	Introduction	73
6.2	Isolation of Mutations	74
6.3	Analysis of the Mutants	76
6.3.1	Analysis of the Mutant 5-8	77
6.3.2	Analysis of the Mutants 3-6, 5-9, and 8-2	78
6.3.3	Conclusion	81
6.4	Cloning and Sequencing the Mutations	82
6.5	The Cloned DNA is Sufficient to Confer the Phenotype	83
6.6	Discussion	86
CHAPTER 7	Discussion	89
REFERENCES		96

FIGURES AND TABLES

CHAPTER 1

- Figure 1.1 Life cycle of a typical laboratory yeast strain
- 1.2 Mating type control
- 1.3 Regions of Homology between the three Mating Type
Loci on Chromosome III
- 1.4 Mating Type Switching Rules
- Table 1.1 Genes involved in Mating Type

CHAPTER 2

- Figure 2.1 Partial Restriction Map of the Plasmid pLG-312
- Table 2.1 Yeast Strains Used
- Table 2.2 Bacterial Strains Used

CHAPTER 3

- Figure 3.1 Linking Number
- 3.2 Effects of sir^{ts} alleles on the linking numbers of
various plasmids
- 3.3 RNA Measurement by the Protection of Radiolabelled
DNA from S1 nuclease
- 3.4 S1 Mapping of MATa1
- 3.5 Extent of DNA probes used, and Interpretation of
the Mapping Data
- 3.6a Mapping the 245 bp Exon
- 3.6b Effect of the rna2^{ts} mutation
- 3.7 Mapping the RNA species in the rna2^{ts} strain
- 3.8 Proposed splicing pattern

CHAPTER 4

- Figure 4.1 Levels of a1 RNA after a shift to the permissive temperature
- 4.2 Effect of alpha factor on SIR repression in exponentially growing cells
- 4.3 a1 Transcription in a sir3^{ts} strain after release from the alpha factor block
- 4.4 Effect of the cell cycle blocks on H0 and Histone mRNA levels
- 4.5 A typical S1 protection experiment
- 4.6 a1 Transcription in a sir4^{ts} strain after release from the alpha factor block
- 4.7a The derepression of HMRa without DNA replication
- 4.7b The repression of MAT α 1 without DNA replication

CHAPTER 5

- Figure 5.1 a1/ α 2 control of the H0 gene
- 5.2 The a1/ α 2 element is sufficient to confer a/ α control
- 5.3 Effect of inserting a1/ α 2 elements in different places
- 5.4 The α 2 element, and its effect on the CYC1 promoter
- 5.5 Effect of a synthetic a1/ α 2 element
- 5.6 Possible model for the action of a1 gene product
- Table 5.1 Aligned Sequences of the a1/ α 2 elements

CHAPTER 6

- Figure 6.1 Sequence comparison of MAT α alleles

6.2	Effect of <u>MATα2</u> [*] on <u>HO</u> and <u>STE2</u> transcription
Table 6.1	Mating behaviour of the Mutants
CHAPTER 7	
Figure 7.1	Distribution of <u>a1/α2</u> and <u>α2</u> elements
7.2	Different insertions of <u>a1/α2</u> elements into the <u>CYC1</u> promoter
Table 7.1	Examples of the <u>α2</u> element

INTRODUCTION

1.1 The Control of Gene Expression

A complete blueprint for an organism is carried on its genetic material, the deoxyribonucleic acid. It is convenient (principally for researchers, but also for the organism,) to divide the information carried on the DNA into units of action, called 'genes'. Perhaps the simplest example of a gene is a contiguous region of DNA which carries the information for a single protein. In addition to protein-encoding genes, there are genes that code for RNA molecules required in protein synthesis and there are regions of DNA which do not code for a product but which are the substrates upon which other gene products act.

The proteins in a cell are its principal catalysts: nearly all the reactions which occur in a cell are catalysed by proteins. Therefore the rate at which any process occurs in a cell depends on the amount and the activity of the relevant proteins in the cell. Cells in various tissues do different things because they contain different mixtures of these protein catalysts. With a few notable exceptions, related cells contain the same complement of genes, but different proteins are produced because they are expressing different subsets of their total repertoire of genes. In this way most of the processes involved in the growth and reproduction of any living organism are controlled via the regulation of expression of the organism's various genes.

A considerable amount of work is in progress to study the ways in which genes are expressed at different levels in different cells. At each step in the process by which a protein product is made from the information in the gene, there is the possibility of control. Firstly an ribonucleic acid (RNA) copy of the coding region of the gene is

made. RNA polymerase binds to the promoter and initiates synthesis, polymerase processes through the gene, transcribing an RNA copy of the gene by adding nucleotides to the 3' end of the nascent polymer as directed by the DNA template. This continues until a transcriptional termination signal is reached. The generation of this primary transcript is often controlled by the rate of initiation of RNA synthesis; sometimes changing the point of termination is used to control the fate of the primary transcript or the gene product produced. (For review of transcriptional control in prokaryotes, see Jacob and Monod, 1961 and Miller and Reznikoff, 1978). The rate of elongation of the nascent transcript is not believed to be used to control gene expression.

In eukaryotes the primary transcript undergoes various modifications before it is exported from the nucleus: a 5' "cap" of 7-methyl-guanosine (Shatkin, 1976) is attached and a poly-adenosine tail is attached at a site near the 3' end of the primary transcript. Intervening sequences within the coding region may be excised (Berget et al, 1977). The rate of production of this processed message and its export from the nucleus may be a control point; changes in the intervening sequences which are spliced out can lead to different products being produced from a single primary transcript (Chow et al, 1977; Kitchingman et al, 1977; Berk and Sharp, 1978). Once the processed RNA is exported to the cytoplasm it is ready to be translated into protein. The mRNA binds to the small subunit of a ribosome and an initiation complex is formed. As translation proceeds the coding information of the mRNA is read (in a 5' to 3' direction) and amino acids are added to the nascent protein chain as directed by the RNA. This continues until a translational termination signal is reached. All three processes, initiation , elongation , and

termination may be used to control the rate of production of the protein product. Finally the protein may undergo various modifications such as glycosylation, ADP-ribosylation, phosphorylation, the attachment of prosthetic groups or the binding of effector molecules. These are often reversible processes, and may be used to modulate the activity of the protein product without changing the amount of protein in the cell. For example the reversible phosphorylation of metabolic enzymes is frequently used to modulate their activity very rapidly in response to changes in metabolite concentrations or hormonal signals. This method of control has the advantages of rapid response and low cost (in terms of energy used) that are attractive for this sort of regulation of metabolism. Longer term changes in the activity of proteins are achieved by modulating the amount of protein present. Different forms of control are used by the cell to solve different problems.

The changes in the patterns of gene expression that are involved in differences in the long-term activity of cells are often controlled at the level of transcriptional initiation. Genes that are only required under particular circumstances, such as growth on a particular carbon source for a microorganism, or genes required only in a specialised cell type (globins in reticulocytes) are controlled at the level of RNA production, with the initiation of transcription being a particularly important step for regulation. Thus under circumstances where the gene product is not required there is no initiation of transcription , and so the subsequent steps in the expression pathway do not occur, and therefore cannot be regulated.

1.2 Control of Transcription, Yeast as a Model System

The budding yeast Saccharomyces cerevisiae is a useful model system in which to study the control of transcription, thanks to the powerful classical genetics of the organism, and the easy application of recombinant DNA techniques to the organism. Stably propagated plasmids, both circular and linear, exist as vectors, and any altered version of a gene constructed in vitro can be used to replace the chromosomal version of that gene. As a result, there exist a number of genetically characterised systems of control of gene expression, such as GAL (Laughon and Gesteland, 1982; Johnston and Hopper, 1982; Bram and Kornberg, 1985), general amino acid control (Hinnebusch and Fink, 1983a; Hope and Struhl, 1985), PPR1 (Losson and Lacroute, 1983), control of CYC1 (Guarente et al, 1984), and the controls involved in mating type, which are discussed below. These systems are not essential for growth on rich media, allowing the isolation or construction of null alleles in the genes involved in their control.

When we refer to control at the level of transcription, we usually mean at the level of initiation of RNA synthesis. This is normally the rate-limiting step for the production of mature mRNA. The elements which are involved in controlling the rate of initiation of transcription fall into two classes: the DNA elements at or near the transcription start site which affect initiation (the promoter) and the trans-acting factors which interact, directly or indirectly, with these elements and with RNA polymerase.

A yeast promoter can be divided into the following elements: the TATA box and its associated initiation site (IS)(Hahn et al, 1985, Nagawa and Fink, 1985). The TATA box is required for efficient initiation, while removal of the IS causes initiation to occur at other sites nearby, without any significant reduction in rate. These

two elements are not sufficient for efficient initiation - they require assistance from additional activating sequences. These sequences are found upstream (5') to the TATA box and are termed Upstream Activating Sequences (UAS's). (Guarente et al., 1984; Sarokin and Carlson, 1984; Struhl, 1985). So far these elements have not been shown to work when placed downstream of the IS (Guarente and Hoar, 1984). There is however evidence to suggest that downstream sequences may promote transcription in certain circumstances. This would be analogous to the action of a mammalian 'enhancer' sequence, which can activate transcription even when placed downstream of the gene.

1.3 Control of Cell Type in Yeast

In the yeast Saccharomyces cerevisiae there are three distinct cell types, a, α , and a/ α (Lindegren and Lindegren, 1943). A cell of mating type α will mate with a cell of mating type a to produce an a/ α cell which does not mate but which is capable of undergoing meiosis and sporulation. α and a cells are usually haploid, while a/ α cells are diploid. The mating type of a cell is determined by a single locus, MAT, which maps to chromosome three. Thus an α cell carries the MAT α allele, and the MATa allele is found in cells of mating type a. When these cells mate they produce a diploid which is heterozygous at MAT: MAT α /MATa.

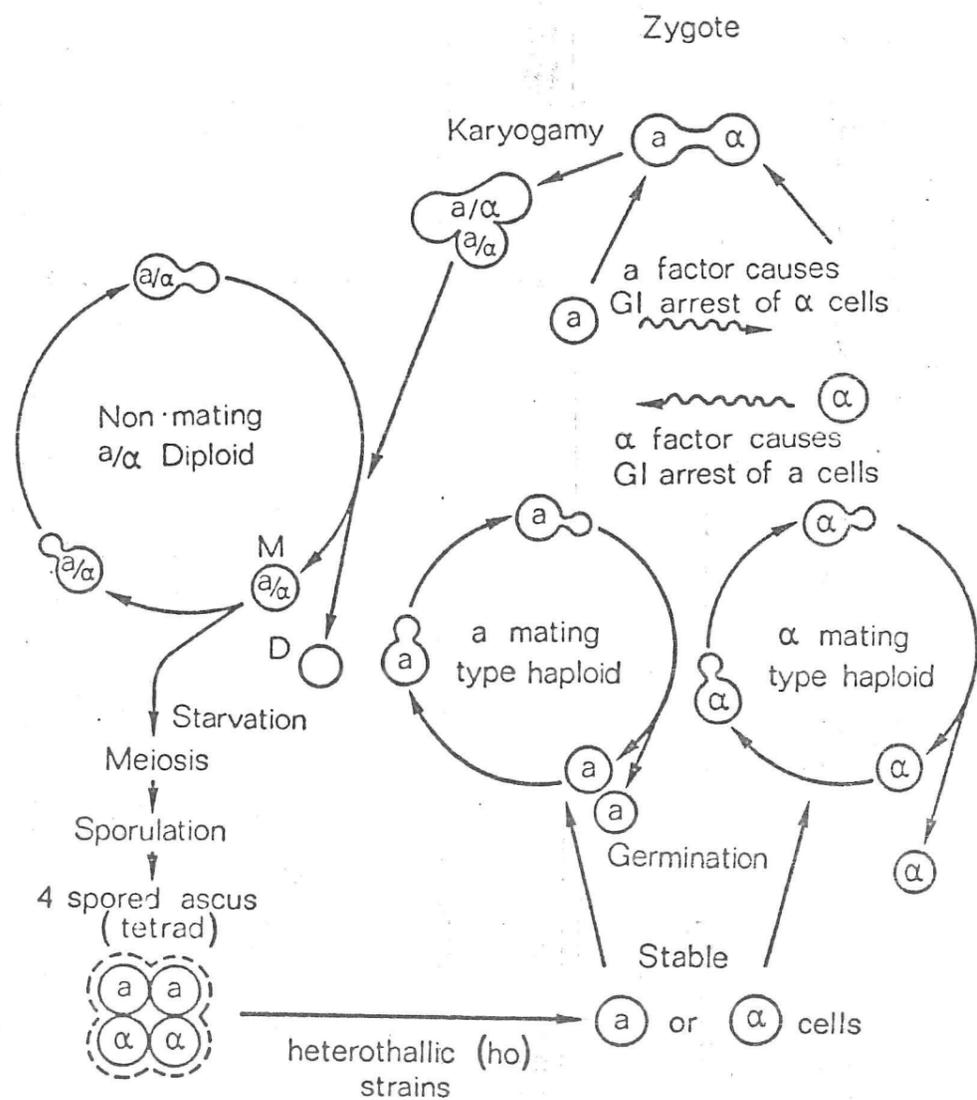
We can see this in the life cycle of a typical laboratory strain. (See Figure 1.1). This strain is heterothallic, that is mating type is a heritable trait which is stable within clones of cells. The three cell types, a haploids, α haploids, and a/ α diploids can all reproduce vegetatively by a budding process in which a mother cell produces a small bud that grows and eventually separates, thus forming another cell, called the daughter cell.

FIGURE 1.1

The life cycle of a typical laboratory yeast strain.

This yeast strain can reproduce vegetatively by budding, both in the haploid (a or α) and diploid (a/ α) phases of its life cycle. When haploid cells of opposite mating type encounter each other, mating can occur to produce an a/ α diploid, which cannot mate.

Under starvation conditions, however, this diploid will undergo meiosis and sporulation to produce an ascus containing four haploid spores, which are the products of a single meiosis. Since MAT is a stable Mendelian allele, two of these spores will be of mating type a, and two α .



Both haploid cell types produce mating pheromones (Levi, 1956, Dunzte *et al*, 1970; Wilkinson and Pringle, 1974) which affect cells of the opposite mating type, causing them to arrest cell division (Bucking-Throm *et al*, 1973; Chan, 1977) and prepare themselves for cell fusion. At this stage directional cell surface growth can be observed (Levi, 1956; Dunzte *et al*, 1973; Herman, 1971). The two pheromones are short peptides of 11 to 13 amino acids (Betz and Dunzte, 1979; Stotzler and Duntze, 1976) produced by the cleavage of larger precursor peptides. This aspect of mating is reviewed by (Thorner, 1981).

Cell fusion and nuclear fusion ensue, producing a diploid cell which has the genotype MAT_a/MAT_α. This diploid does not mate or respond to mating pheromones, but it is capable of undergoing meiosis and sporulation to produce an ascus containing the four haploid progeny of a single meiosis. Since MAT is a Mendelian locus, two spores will be of mating type a, and two α.

1.3.1 The MAT genes are Master Regulators

MacKay and Manney (1974a,b) isolated and analysed mutants defective in mating. Some mutations at MAT_α were found, but the majority of sterile mutations were unlinked to MAT. Some of these mutations only affected mating by MAT_α strains, while others only affected mating by MAT_a strains. The remainder affected mating in both a and α strains (Hartwell, 1980)(See Table 1.1). No mutations in MAT_a were isolated in this study.

These results implied that a number of genes were involved in mating and that these genes were not physically linked to MAT, but their expression was controlled by the MAT locus. Although MacKay and Manney (1974a,b) did not observe any role for MAT_a in mating, Kassir

TABLE 1.1

Genes involved in Mating Type.

Various genes involved in mating type in *Saccharomyces cerevisiae* are listed, along with what is known about their expression and function. Alternative names are given in brackets. d.n.a.= data not available.

Gene	expressed in cell types	mutations cause sterility in	Function
MATa1	a and a/α		acts with MATα2 to determine a/α cell type
MATα1	α	α	positive regulator of α-specific genes
MATα2	α and a/α	α	negative regulator of a-specific genes, acts with MATa1 to determine a/α cell type
MFa1	a		a cell mating pheromone
MFa2	a		a cell mating pheromone
MFα1	α		α cell mating pheromone
MFα2	α		α cell mating pheromone; the mfa1 ⁻ mfa2 ⁻ double is α sterile
NUL3	d.n.a.	a and α	
SST1 (BAR1)	a		inactivates α-factor; mutation makes a cells supersensitive to α-factor.
SST2	d.n.a.		mutation sensitizes a cells to α-factor, and α cells to a-factor
STE2	a	a	α-factor receptor
STE3	α	α	a-factor receptor
STE4	d.n.a.	a and α	
STE5	a and α	a and α	
STE6	a	a	involved in a-factor production
STE7	all	a and α	mutations all show varying defects in the transcription of other cell-type-specific genes.
STE11	all	a and α	
STE12	all	a and α	
STE13	all	α	aminopeptidase involved in α-factor processing
STE14	d.n.a.	a	endopeptidase involved in α-factor processing
KEX2	all	α	
RME	a and α		repressor of sporulation
HO	a and α		endonuclease that initiates mating type switching
SIR1			} all required for the repression of HML and HMR loci
SIR2 (MAR1)			
SIR3 (MAR2, STE8, CMT)			
SIR4 (STE9)			

and Simchen (1976) found that a mutation in MAT_a caused a defect in diploid functions: MAT_a⁻/MAT_α diploids maintain α mating ability and are not able to sporulate.

The sterile mutations that mapped to MAT_α fell into two complementation groups, termed MAT_{α1} and MAT_{α2}. The phenotypes of mat_{α1}⁻ and mat_{α2}⁻ mutations (Hicks and Herskowitz, 1976a; Tkacz and MacKay, 1979; Kassir and Simchen, 1976; Klar *et al.*, 1979b) led Strathern *et al.* (1981) to propose the α₁-α₂ model for the control of cell type. Further molecular studies (reviewed by Nasmyth 1982a) have confirmed and elaborated on this model, which is shown in Figure 1.2. MAT_a and MAT_α each direct the synthesis of two unique transcripts: a₁ and a₂, or α₁ and α₂ (Nasmyth *et al.*, 1981b; Tatchell *et al.*, 1981). In an a cell, there appears to be no role for either the a₁ or the a₂ transcript in determining cell type (See Figure 1.2), and therefore a loss of MAT_a function does not affect mating ability, as mentioned above. In an α cell, both MAT_{α1} and MAT_{α2} gene products are required for the α mating phenotype. MAT_{α1} is required for the transcription of the "α-specific genes" (such as MF_{α1} MF_{α2} STE3: Strathern *et al.*, 1981; Sprague *et al.*, 1983), whereas MAT_{α2} is a repressor of the a-specific genes (MF_{a1} MF_{a2} STE2 STE6 BAR1: see chapters 5 and 6, and Wilson and Herskowitz, 1984; Strathern *et al.*, 1981).

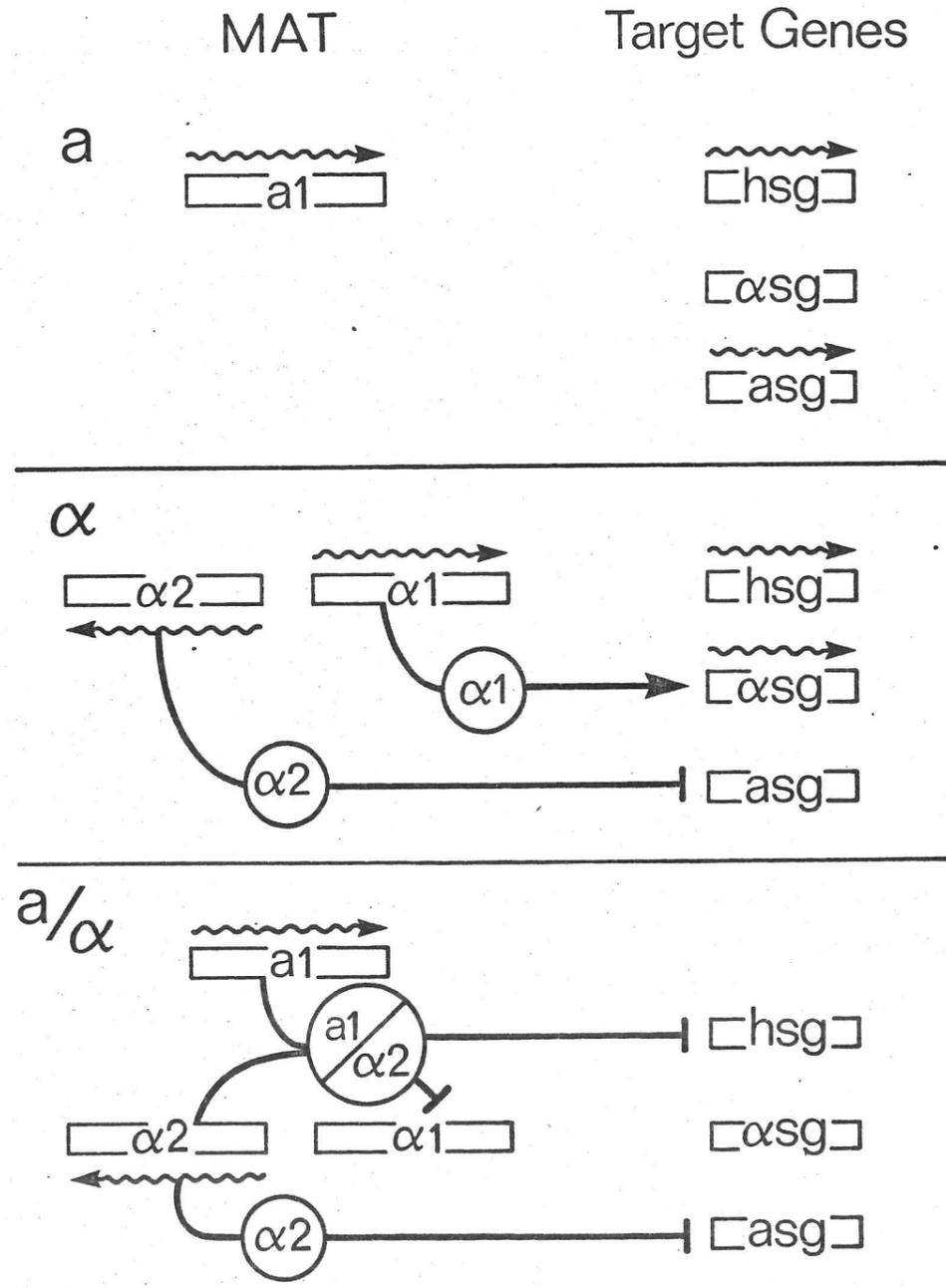
Thus if MAT_{α1} function is lost, sterility ensues since α-specific genes are not expressed. Alternatively, loss of MAT_{α2} function leads to the expression of both a-specific and α-specific genes in the same cell. It is not clear why this causes sterility, but presumably the a- and α-specific gene products interfere with each other's function in some way. If however both α₁ and α₂ functions are lost, the cell will now mate as an a cell. This double mutant (called an a-like faker) and the a₁⁻ mutant described above both mate as a's since they lack

FIGURE 1.2

Mating type control

In a cell of mating type a, the a-specific genes (asg) and the haploid specific genes (hsg) are expressed constitutively. In a cell of mating type α , the asg's are repressed by the action of the $\alpha 2$ gene product while the α -specific genes (α sg) are activated by the $\alpha 1$ gene product (Strathern et al, 1981). When these two cell types mate to produce an a/ α diploid, the $a 1$ and $\alpha 2$ gene products together repress the expression of the haploid-specific genes, including MAT $\alpha 1$.

Mating Type Control



all MAT functions, but produce $\underline{a}^-/\underline{\alpha}$ diploids which cannot sporulate (see below).

When an \underline{a}^+ cell and an $\underline{\alpha}$ cell have mated, a novel situation arises in that both MAT_a and MAT_α gene products are present within the same cell. The MAT_{a1} and MAT_{α2} gene products act in combination to repress an additional set of genes, which are not required in diploids. These "haploid specific genes" include MAT_{α1}, HO and STE5 (Nasmyth *et al.*, 1981a; Jensen *et al.*, 1983; J.Thorner, personal communication). Since MAT_{α1} is repressed, the α -specific genes are not expressed. The a -specific genes are still repressed, presumably by the presence of MAT_{α2}. The cell has therefore a non-mating phenotype.

There remains the question of why only $\underline{a}/\underline{\alpha}$ cells are capable of sporulation. Mutations have been isolated that remove this restriction and allow $\underline{a}/\underline{a}$ and $\underline{\alpha}/\underline{\alpha}$ diploids to sporulate (Hopper and Hall, 1975; Kassir and Simchen, 1976). These mutations are recessive and lie within the RME gene, suggesting that RME encodes a repressor of sporulation functions, the removal of which is sufficient to allow sporulation to proceed in $\underline{a}/\underline{a}$ or $\underline{\alpha}/\underline{\alpha}$ cells. Recent experiments by Mitchell and Herskowitz (1986) show that RME transcription is repressed in $\underline{a}/\underline{\alpha}$ diploids. Thus MAT_{a1} and MAT_{α2} in combination may allow sporulation to occur by repressing the transcription of a repressor of sporulation (RME).

In this thesis I describe work identifying DNA sequences that are the targets of $\underline{a1}/\underline{\alpha2}$ -mediated repression and also those sequences recognised by $\underline{\alpha2}$ -mediated repression (See Chapter 5).

1.3.2 Mating type Switching

Most strains of Saccharomyces cerevisiae found in the wild have an added complication to their life cycle. MAT alleles are not stable.

Haploid wild type strains are able to switch their mating types at a frequency of about 80% per generation, by the directed mutation of one MAT allele to the other (Hawthorne, 1963a). This switching is catalysed by the product of the H0 gene (originally called D for Diploidization) (Hicks and Herskowitz, 1976b, Takano and Oshima, 1970a). In a strain carrying the wild type H0 gene, a single haploid cell of either mating type rapidly gives rise to a mixture of a, α , and a/α cells, hence the name H0 for homothallic.

How is this mating type interconversion achieved. Genetic studies revealed that two other loci on chromosome three were required for mating type switching, and that these loci behaved genetically as if they were donors of information for either the a mating type or the α mating type (Takano and Oshima, 1967, 1970a,b; Harashima et al, 1974). Oshima and Takano (1971) proposed that a "controlling element" (McClintock, 1956) was transferred to the MAT locus, and that this element could then cause the expression of either MATa or MATα functions. Then Hicks et al (1977) proposed that the other two loci (now called HML and HMR) contained the transcription units found at MAT but that these loci were silent since they lacked a functional promoter. Transfer of these unexpressed transcription units to MAT led to their expression. This model arose from the observation that mutations at MAT could be healed by switching events of the form $\alpha^- \rightarrow a \rightarrow \alpha^+$ (Hicks and Herskowitz 1977), and later Strathern et al, 1979).

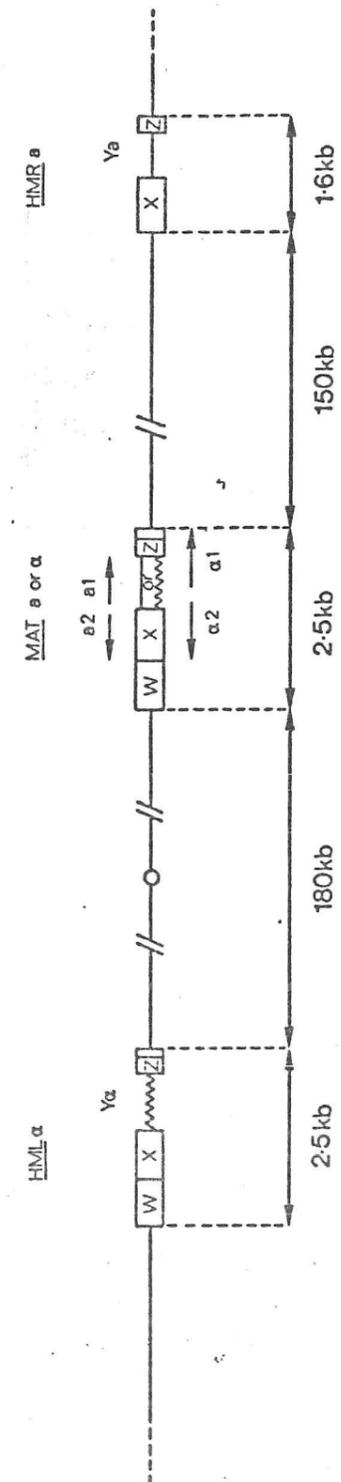
Physical analysis of HML, HMR and MAT (Hicks et al, 1979; Nasmyth and Tatchell, 1980, Strathern et al, 1980) has shown that the silent loci contain the transcription units of the mating type genes, but also contain the divergent promoter which directs expression of these genes (Figure 1.3). At the silent loci, however, sequences

FIGURE 1.3

Regions of Homology between the three Mating Type Loci on Chromosome III.

The chromosome is denoted by a straight line, with the centromere marked as an open circle. The boxes marked W, X, and Z represent the regions of identical sequence shared between the loci, while the arrows show the extent of the transcribed regions. Y_α is a 747 bp α -specific sequence, while Y_a is a 642 bp a-specific sequence.

Although the promoter and the transcription units found at MAT are also present at HMR and HML, they are transcriptionally repressed due to sequences outside the regions of homology. Therefore if Y_α DNA is present at MAT, then α_1 and α_2 will be expressed, but if Y_a is present at MAT, then a_2 and a_2 will be expressed.



outside the transposed region repress the transcription of those genes (Klar *et al*, 1980; Nasmyth *et al*, 1980).

The transposition of DNA sequences from a silent locus to MAT is initiated by a double-stranded cut at the MAT locus (Strathern *et al*, 1982). The cut is made by a sequence-specific endonuclease which is the product of the HO gene (Kostriken *et al*, 1983). The switching process is under a variety of controls. Switching occurs frequently in haploids of either mating type, but extremely rarely in a/α diploids (Hicks and Herskowitz, 1977). This control is essential for the yeast's life cycle, since if switching occurred in diploids, then a/a and α/α diploids would rapidly arise. These diploids, homozygous at MAT, would mate as if they were haploids, producing a triploid yeast strain that cannot sporulate properly. In order to maintain control of ploidy, switching must not occur in diploids.

Different lines of evidence imply that the switching event occurs in the G_1 period of the cell cycle, after cells have been committed to a mitotic cell cycle, but before DNA replication (Hicks *et al*, 1977; Klar *et al*, 1982). As a consequence, when a switching event occurs, both progeny produced by the switching cell have changed mating type (Hicks and Herskowitz, 1976b). Also, since switching does not occur until after the commitment to the mitotic cell cycle, cells cannot switch when they are arrested by mating pheromone in preparation for mating. This latter control is clearly adaptive.

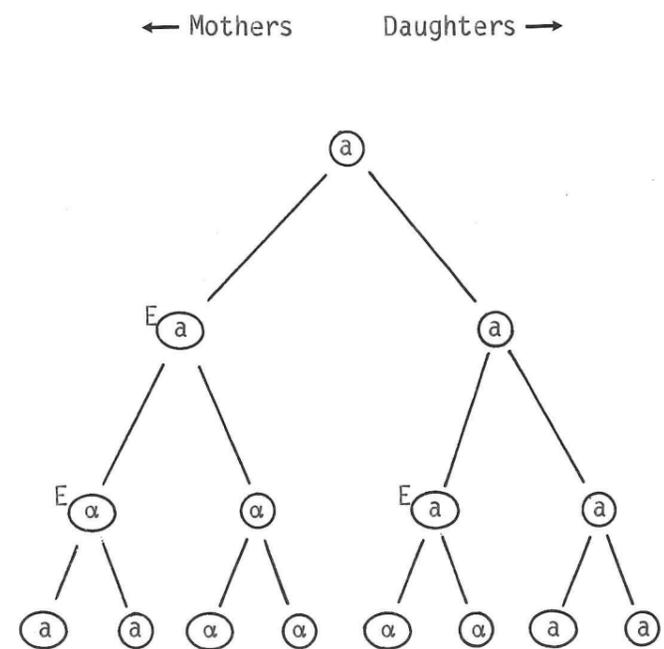
Finally, daughter cells do not switch in their first cell division cycle (See Figure 1.4).

Most of these controls of switching can be explained by the temporal control of HO gene transcription: the HO gene is only expressed within a narrow window of G_1 in Mother cells (i.e. cells that have experienced at least one cell division since they were

FIGURE 1.4

Mating Type Switching Rules: the Pattern of Switching in a Homothallic (H0) Strain.

Only cells that have already produced at least one bud, marked 'E' for experienced, are able to give rise to cells of opposite mating type. When switching does occur, both progeny change mating type.



produced as buds)(Nasmyth, 1983). If the H0 gene is expressed constitutively, then daughters do switch mating types, but the switching still occurs predominantly in pairs, implying that the G_2 switching of one MAT chromatid is still a rare event (Jensen and Herskowitz, 1984) and suggesting that other controls on the switching event may also be operating.

H0 is not transcribed at all in a/α cells (Jensen et al, 1983) due to MAT α 1/MAT α 2 mediated repression (see section 1.3.1 and chapter 5).

1.3.3 SIR repression

As mentioned in the previous section, the silent mating type loci contain the entire transcription units found at MAT, and the promoter that directs their transcription. They are not expressed, however, due to some effect mediated by DNA sequences outside the region of homology. The sequences which are required in cis for repression have been mapped by constructing deletions within plasmids carrying HML or HMR DNA, and are found to lie up to 1700 base pairs from the affected promoters (Abraham et al, 1984; Feldman et al, 1984).

Also required for the repression of the silent loci are four trans-acting gene products SIR1,2,3,4 (Table 1.1)(Hopper and Hall, 1975; Hicks, 1975; Rine, 1979; Rine et al, 1979; Haber and George, 1979; Klar et al; 1979; Hartwell, 1980). Recessive mutations in any of these genes causes both HML and HMR to be expressed. These mutations were isolated and characterized as repressors of the silent loci; Klar et al (1981a) have demonstrated that the repression is exerted at the level of transcription.

In addition to its effect on transcription, the action of the SIR gene products on the silent loci ensures that these loci are only used

as donors of sequence information in the switching process and are not the recipients. In sir⁻ strains the silent loci become targets for the HO endonuclease and therefore for the transposition events (Klar et al, 1981b). It is possible that this effect on transposition is a consequence of the effect on transcription. Nasmyth (1982b) however found that the SIR-dependent change in chromatin structure at the HO cut site was not merely a consequence of transcription through that region of DNA.

The four SIR genes have been cloned and sequenced (Shore et al, 1984, Ivy et al, 1986) and they all appear to encode proteins. When in vitro constructed disruptions of SIR2,3, or 4 are used to replace the chromosomal copy of the gene, full expression of the silent loci occurs. The disruption of SIR1 only leads to partial expression, however (Ivy et al, 1986).

SIR-mediated repression appears to be rather unusual, in that the sequences which are necessary for repression lie some distance from the affected promoters, and we must postulate that some kind of 'action at a distance' is involved. This thesis describes work which investigates a possible connection between SIR-mediated repression and DNA replication.

CHAPTER TWO

MATERIALS AND METHODS

2.1 Bacterial Growth Media

Bacteria were grown in 2xTY medium (16g tryptone, 10g Yeast Extract, 5g NaCl per litre, pH 7.4). Bacterial strains were maintained, and single colonies isolated, on TYE plates (15g agar, 8g NaCl, 10g Bactopectone, 5g Yeast Extract per litre).

Ampicillin selection was performed at final concentrations of 40 μ g/ml for liquid media, and 40 - 100 μ g/ml for solid growth media.

2.2 E.Coli transformation

Transformation of E.Coli was performed according to the procedure of Hanahan (1983). Transfection with M13 DNA was performed as described by Messing (1983).

2.3 Preparation of DNA

Plasmid DNA was prepared from E.Coli cultures as described by Birnboim and Daly (1979), and banded in an isopycnic caesium chloride/ethidium bromide gradient.

When large numbers of different DNA clones were being analysed in order to identify the desired recombinant, "miniprep" DNA was prepared using a modification of the method of D.Holmes and M.Quigley (1981). Individual colonies were inoculated into 3ml of 2xTY and grown for 5 - 6 hours at 37^oC with vigorous shaking. Cells were harvested by centrifugation, resuspended in 1ml of water and recentrifuged. These pellets were resuspended in 380 μ l of a mixture comprising: 350 μ l of STET (8% Sucrose (w/w), 5% Triton X-100, 50mM Tris HCl, 50mM EDTA, pH

8.0) and 30 μ l of 5mg/ml lysozyme. The samples were vortexed briefly then placed in boiling water for a minute and cooled on ice. Centrifugation for about 30 minutes in an eppendorf centrifuge produced a glutinous pellet which was removed with a toothpick and discarded. Nucleic acids were precipitated by the addition of 0.2ml 5M ammonium acetate and 1ml isopropanol, and centrifugation for 3 mins at 4 $^{\circ}$ C. The pellets were washed with 80% Ethanol, dried under vacuum, and dissolved in 30 - 50 μ l of (10mM Tris HCl, 1mM EDTA, pH7.4).

Typically, 2.5 μ l of this DNA preparation would be digested with appropriate restriction enzymes in a final volume of 15 μ l.

2.4 Gel Electrophoresis

Plasmid DNA and restriction fragments were analysed on 0.8 to 2.0% agarose submarine gels and electrophoresed in 1 x TBE buffer (90mM Tris base, 90mM Boric acid, 2.5mM EDTA, pH 8.3. Made up as 10 x TBE : 108g, 55g, and 9.3g per litre respectively) and ethidium bromide (1 μ g/ml). Nucleic acid was visualised under ultraviolet and photographed with a polaroid camera.

2.5 Enzymes and chemicals

Restriction enzymes and T4 DNA ligase were bought from New England Biolabs or East Anglia Biotechnology, who were also a source of E. Coli polymerase I Klenow fragment. Klenow and polymerase I holoenzyme were also supplied by Boehringer Mannheim, and S1 nuclease came from Bethesda Research Laboratories. Glusulase was purchased from Du Pont, and Zymolyase from Seikagaku Kogyo Co. Ltd, Tokyo. T4 polynucleotide kinase came from Pharmacia, and Exonuclease VII was the kind gift of J.Chase.

Radiochemicals were purchased from Amersham, while other

chemicals came either from BDH or Sigma. Nitrocellulose sheets were purchased from Sartorius.

2.6 Yeast Media and Transformation

Yeast genetics were performed as described by Mortimer and Hawthorne (1969).

DNA-mediated yeast transformation was performed as described by Beggs (1978); the simplification of this protocol described by MacKay (1983) was sometimes used.

2.7 Yeast Plasmid Preparation and Supercoiling gels

Cells were harvested by centrifugation, resuspended in water and transferred to a 1.5ml eppendorf tube. The cells were re-pelleted and resuspended in 0.2ml SCE/Zymolyase/ β ME mix (1M Sorbitol, 0.1M Sodium Citrate, 0.06M EDTA, made 0.6mg/ml Zymolyase and 0.1M β -mercaptoethanol just before use). After incubation at 37°C for about 40 minutes, 0.2ml of (2% SDS, 100mM Tris HCl, 10mMEDTA; pH 9.0) was added, the suspension mixed briefly and heated to 65°C for 5 minutes. 0.2ml of Potassium acetate was added and the samples were held on ice for 20 minutes. After a 5 minute centrifugation in an eppendorf centrifuge, the supernatant was taken and nucleic acids precipitated by the addition of 0.2ml 5M Ammonium acetate and 1ml isopropanol. The pellet, collected by a gentle spin, was redissolved in 90 μ l water and reprecipitated by the addition of 10 μ l 5M Ammonium acetate and 200 μ l isopropanol. The fibrous pellet so formed was washed with 80% ethanol, dried, and dissolved in 50 μ l TE. Insoluble contaminants may be removed by centrifugation.

To measure the linking number of plasmids in such preparations, the DNA prep was first digested with RNase (boiled) and then

electrophoresed through a 0.8% agarose 1 x TPE gel (50mM Tris phosphate, 1mM EDTA; pH7.2) with Chloroquine added to a final concentration of 5 μ g/ml (Shure et al, 1977). Electrophoresis was performed in the dark at 4 $^{\circ}$ C. For a 300mm gel, conditions were 93V for 42 hours, with approximately 4l of buffer circulating.

After electrophoresis the gel was stained with ethidium bromide prior to photography. The portion of the gel below the chromosomal DNA was then prepared for southern transfer (see below)

2.8 Southern Transfer

Gels were prepared for Southern transfer (Southern, 1975). Hybridisation to radiolabelled DNA and subsequent washing of blots was performed as described in (Nasmyth 1982b).

2.9 Sequencing

Chain termination sequencing reactions and the electrophoresis of their products were performed as described in (Sanger et al, 1977; Bankier and Barrell, 1983). The M13 vectors are described in (Messing and Vieira, 1982).

2.10 Preparation of DNA probes for S1 Nuclease protection

Various fragments of the MATa1 gene were subcloned into M13 vectors as follows. Plasmids containing HMRa DNA that had been subjected to Xho-linker mutagenesis were digested with XhoI and another enzyme and were subcloned into the M13 sequencing vectors mp9 and mp10. Three mutations were used, numbers 3, 144 and 238, all Xho derivatives of the plasmid 82.6 (Abraham et al, 1982, 1984). Numbers 3 and 238 were digested with XhoI and BglII and ligated into SalI/BamHI cut M13mp9. (Probes C and A respectively in Figure 3.5). Number 144

was digested with XhoI and XbaI and ligated into SalI/XbaI cut M13mp10. The identity of the recombinants was confirmed by chain-termination sequencing. Over the region of DNA covered by these probes, the HMRa sequence is identical to the MATa sequence, the former was used as the source of DNA because a wider range of Xho linker positions was available.

The SIR3 probe is described in (Shore *et al.*, 1984). The H0 probe is described in (Nasmyth, 1983) and the Histone H2B probe in (K.A.Nasmyth, 1985a). The $\alpha 1$ DNA probe was made by subcloning the Xho-linker mutation $\alpha X109$, which is a Xho linker mutation 157 nucleotides upstream from the 5' terminus of the MAT $\alpha 1$ message (Tatchell *et al.*, 1981). This DNA was cut with XhoI and HincII to produce a 655 nucleotide fragment that was ligated into SalI/SmaI cut M13mp9. The identity of the recombinant plasmid was confirmed by sequencing.

The STE2 probe was made by digesting plasmid pZV37 with HindIII and EcoRV to produce a 680 bp fragment which was ligated into HindIII/SmaI cut M13mp9; this produces a probe that extends 140 bp upstream of the STE2 transcription start site.

2.11 RNA preparation and S1 analysis

Total yeast RNA was prepared and levels of RNA were determined by S1 nuclease protection of radiolabelled DNA, as described in (Nasmyth, 1983).

The single-stranded radiolabelled DNA was prepared by *in vitro* DNA synthesis. With single-stranded phage DNA as a template, Klenow polymerase was used to extend a 17 nucleotide primer (Duckworth *et al.*, 1981) across the insert in the presence of 200 curies/millimole ³²P deoxy-adenosine triphosphate. The resulting double-stranded DNA was

cleaved with a restriction enzyme cutting distal to the insert, and the radioactive strand isolated by electrophoresis through a 7M urea 4% polyacrylamide gel.

2.12 Construction of isogenic strains by DNA transformation

Strain M26 (Table 2.1) was made from strain M20 by transformation with a plasmid carrying the SIR3 gene in YRP7 (Struhl et al., 1979). Southern analysis revealed that the plasmid had integrated at the TRP1 locus. Strains M28 and M29 were made by transplacing (Rothstein, 1983) a sir^{ts} strain to sir⁰, and then transplacing both the parent and the transformant to mat::LEU2 to give M28 and M29 respectively. The constructions used were an insertion of the TRP1 gene into the open reading frame of SIR3 and an insertion of the LEU2 gene into the Xho-linker deletion of MAT α , α X8 (Tatchell et al., 1981).

RS3 and M48 were made from RS1 by transplacement at MAT

2.13 Preparation of G₀ daughters and release from the alpha factor block.

Cells were grown into G₀ on YEPD plates and daughter cells were purified as described in (Nasmyth, 1983). The daughters were inoculated into YEPD to an O.D.₆₆₀ of 0.25, then after one hour at 23°C alpha factor was added to 3 units/ml, and the cells were grown for a further 5 hours, then harvested by centrifugation and inoculated into YEPD containing 3 units/ml alpha-factor, or 200mM HU, or 50µg/ml MBC (1/200th volume of 10 mg/ml MBC in DMSO), or no cell cycle block at all. Cell samples for RNA preparation were chilled by the addition of about one third volume of ice.

2.14 Preparation of alpha-factor

α factor was prepared according to the procedure of Bucking-Throm et al ., 1973.

2.15 Sequence Comparisons

The 5' flanking sequence of the H0 gene (1950 bp) was compared with itself, and with its complementary strand, using DIAGON (Staden, 1982). This revealed a number of examples of the putative α_1/α_2 element. The H0 sequence was DIAGONed against the MAT α intergenic region, revealing the same sequence at MAT α_1 : these sequences were aligned to produce a frequency matrix (using GETFRQ) and the H0 sequence was reanalysed using the ANALYSEQ program suite (Staden, 1984). This revealed a total of ten examples of this motif within the H0 gene, and two examples within 705 bp of 5' flanking sequence at STE5. The frequency matrix used to search the EMBL library was derived from these examples, except that the MAT α_1 sequence was included three times, to give it greater weight. At the time, this was the only example which was known to be sufficient for repression. I have re-evaluated all the occurrences I have found, testing their score against a frequency matrix that lacks this arbitrary weighting. Thus all the "scores" mentioned in this dissertation refer to a frequency matrix derived from the set of examples shown in Table 5.1.

The sequences of the 5' flanking regions of the BAR1(690 bp) and MFa1(628 bp) genes were compared with each other, and with the 400bp HindIII fragment of STE2 using DIAGON (Staden, 1982).

2.16 Sequencing STE2

pZV37 was digested with HindIII and ligated with HindIII cut M13mp10. The sequence was then derived from each end by the dideoxy technique (Bankier and Barrell, 1983). The $\alpha 2$ element starts 23 bp from one of the HindIII sites. The entire STE2 gene has been sequenced recently by Nakayama *et al*, (1985).

2.17.1 Cloning of the $\alpha 1/\alpha 2$ elements into the CYC1 promoter.

The H0 plasmid was digested with BamHI and BglII, filled in and blunt-end ligated into the XhoI site of pLG- Δ 292S. Xho-linker mutant H204 was digested with XhoI and NruI and ligated into the XhoI site of pLG- Δ 292S. Recombinants were identified by colony hybridisation, and their identity was confirmed by subcloning and sequencing: the SmaI/BamHI fragment of pLG- Δ 292S carrying the $\alpha 1/\alpha 2$ element insert was cloned into HincII/BamHI cut M13mp10. M13 carrying the relevant insert were identified by plaque hybridisation and sequenced by the dideoxy technique. DNA's used pH0 DNA source was a 4.3kilobase(kb) Sau3A partial fragment, and the Xho linker mutant No. 204 derived from it (Nasmyth, 1985a). pLG- Δ 292S was the kind gift of L.Guarente (Guarente *et al*, 1982): this plasmid has been renamed pLG-312 (Guarente and Mason, 1983) and a partial restriction map of the plasmid is shown in Figure 2.1. The STE2 plasmid was pZV37 (carrying a PstI-EcoRI fragment in pUC.), The vector for sequencing was M13mp10 (Messing and Vieira, 1982).

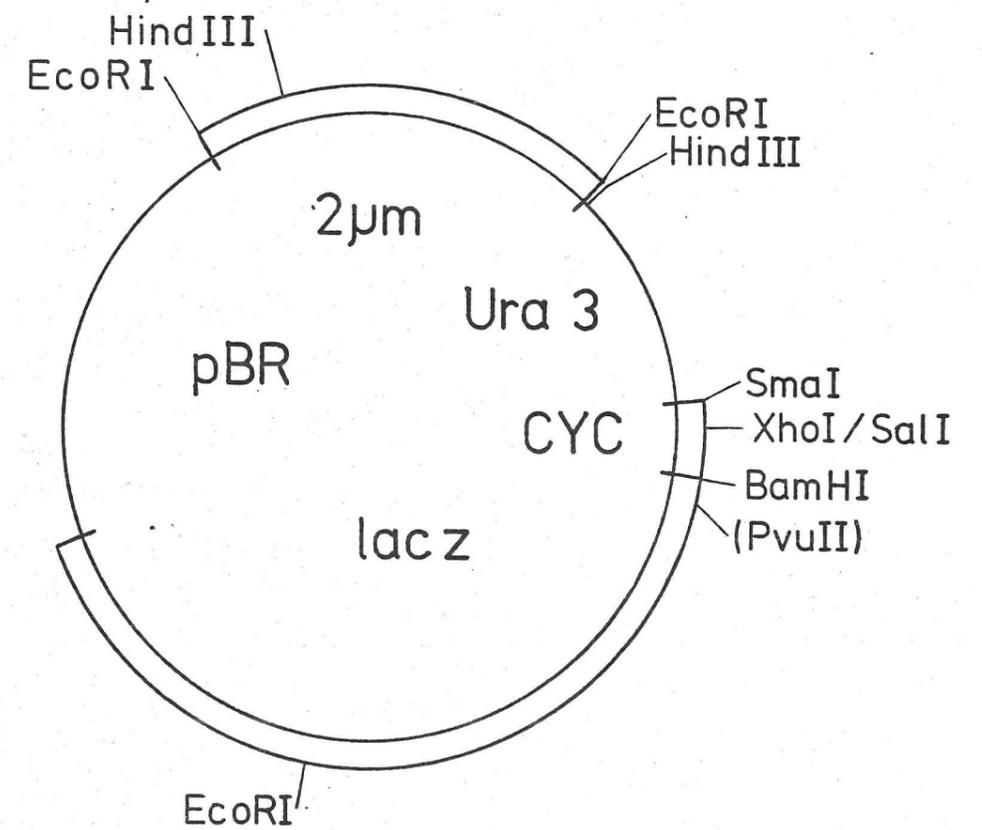
Similarly the constructions carrying an $\alpha 1/\alpha 2$ element inserted at the SmaI site or at the BamHI site of pLG-312 were constructed from the BamHI-BglII fragment of the H0 gene. Plasmids containing inserts at the SmaI site were analysed by subcloning the relevant XhoI-HindIII

FIGURE 2.1

Partial Restriction Map of the Plasmid pLG-312

The map shows the construction of pLG-312, and the recognition sites of restriction enzymes utilised in sub-cloning operations. There are a number of PvuII sites in the plasmid. For clarity I have only shown the one which was used for subcloning.

The map is to scale; the overall size of the plasmid is 10 kb.



fragment into SalI/HindIII cut M13mp9 and sequencing, whereas plasmids containing inserts at the BamHI site were analysed by subcloning the relevant XhoI-PvuII fragment into SalI/SmaI cut M13mp9 and sequencing. In this way the sequence and orientation of the inserts could be ascertained.

2.17.2 Inserting a fragment of STE2 into the CYC1 promoter.

The plasmid carrying the STE2 gene (pZV37) was digested with HindIII and AvaII and blunt-end ligated into XhoI cut pLG-312. Recombinants were identified by colony hybridisation, and their identity was confirmed by subcloning and sequencing. The SmaI/BamHI fragment of pLG-312 carrying the $\alpha 2$ element insert was cloned into HincII/BamHI cut M13mp10. These recombinants were sequenced.

2.18 Beta-galactosidase assays

β -galactosidase assays were performed using the technique of Miller (1972). Cells were collected by centrifugation, resuspended in Z buffer and permeabilized with CHCl_3/SDS . After the reaction was stopped with base, cell debris was removed by centrifugation prior to measuring the O.D._{420} . Units were measured as $1000 \times \text{OD}_{420} / (\text{OD}_{660} \times \text{time (in mins)})$.

2.19 Cloning of synthetic oligonucleotides

A synthetic oligonucleotide having the sequence 5'-TCGATTCATGTTATTATTTACATCAT-3' was synthesised using the phosphotriester method (See for example Gait, 1984) on a Biosearch SAM ONE oligonucleotide synthesis machine. This sequence is identical to an example of the $\alpha 1/\alpha 2$ element found at the HO gene (HO -411),

with an additional sequence of five nucleotides 5'-TCGAT...-3' to facilitate ligation into a SalI site. Also synthesised was a mixture of three oligonucleotides which are essentially complementary to the first oligonucleotide, but differ at a single nucleotide. 5'-TCGAATGATXTAAATAATAACATGAA-3', where X represents A or T or C. The oligonucleotides were purified on a 20% polyacrylamide 1xTBE gel, annealed to each other and treated with polynucleotide kinase to produce 5' phosphate groups. The heteroduplex DNA molecules thus formed were ligated into SalI cut pLG-312 (Guarente and Mason, 1983). After ligation, the DNA ligase was inactivated by heating, and the recircularised vector molecules were eliminated by redigestion with SalI. Recombinant DNA molecules were identified by restriction digestion, and the sequence of the insert they carried was determined by digesting with XhoI, and cloning and sequencing the small restriction fragment produced. This procedure does not reveal the orientation of the oligonucleotide inserts within pLG-312. The orientation was determined by taking advantage of the HinFI site at one end of the oligonucleotide inserts. Insert-containing plasmids were digested with BamHI, then end-labelled with Klenow polymerase as described in Maniatis (1982), and digested with HinFI. The sizes of the labelled fragments produced were compared with the size of the BamHI-SalI fragment produced by the parent molecule, pLG-312. In this way it was possible to determine both the sequence and the orientation of the oligonucleotide inserts.

2.20 Cloning MAT α alleles from M30, 5-9, and 8-2

Yeast DNA was prepared as described in section 2.7, digested with HindIII, and electrophoresed through a 0.8% agarose 1xTBE gel for 14.5 hours at 2.33 V cm⁻¹. DNA fragments of approximately 4.2 kb were

extracted from the gel by cutting a slot in the gel and inserting a piece of GF/C paper (Whatman) into the slot. DNA was electrophoresed into the paper for 30 minutes at 3.33 V cm^{-1} . The paper was then removed and incubated in (1.5M NaCl, 50mM Tris HCl pH=8.0, $10 \mu\text{g/ml}$ tRNA) for 15 minutes at 65°C , to elute the bound DNA. This solution was then phenol extracted, chloroform extracted and passed over a 0.6ml G-50 coarse column equilibrated with TE before DNA was precipitated with ethanol. These DNA fragments were ligated into HindIII cut M13mp10 and plaques carrying MAT α DNA were identified by plaque hybridisation. M13 template DNA was sequenced using primers complementary to regions of MAT α DNA: ATTATCAACTTACACAG and CAGCTTAGAAGTGGGCA.

The sir3^{ts} and sir4^{ts} alleles used are ste8-59a and ste9-62c respectively (Hartwell, 1980). cdc28-4 is from (Reed, 1980). sir3⁰::TRP1 denotes the insertion of the TRP1 gene into the open reading frame of the SIR3 gene. This SIR3 gene disruption was the kind gift of D.Shore and the MAT α 2 mat α 1::LEU2 construction was the kind gift of L.Breeden. mat::LEU2 denotes an insertion of the LEU2 gene into the XhoI-linker mutation MAT α X8, which is $\alpha 1^-$, $\alpha 2^-$ (Tatchell et al, 1981).

Strains W303-1A and W303-1B are isogenic (except for MAT), and were the kind gift of R.Rothstein. Strains beginning RS were the gift of R.Sternglanz, and Y65, which carries a deletion of HMRE that leads to expression of HMR α 1 and a non-mating phenotype, was the kind gift of A.Brand.

The ho:lacZ construction was made by L.Breeden (Breeden and Neigh-Smith, 1985).

The rna2^{ts} strains were the kind gift of J.Warner, (Fried and Warner, 1982)

TABLE 2.1
Genotypes of the yeast strains used

STRAIN	GENOTYPE
M9	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>sir3^{ts}</u> <u>cdc28-4^{ts}</u> <u>ade2</u> <u>leu2</u> <u>met2</u> <u>tyr1</u> <u>ura1</u>
M10	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>sir3^{ts}</u> <u>CDC⁺</u> <u>ade2</u> <u>leu2</u> <u>met2</u> <u>ura1</u>
M20	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>sir3^{ts}</u> <u>cdc28-4^{ts}</u> <u>ade2</u> <u>met2</u> <u>trp1</u> <u>tyr1</u>
M26	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>sir3^{ts}</u> <u>cdc28-4^{ts}</u> <u>ade2</u> <u>met2</u> <u>TRP1</u> <u>SIR3</u> <u>tyr1</u>
M28	<u>HMLα</u> <u>mat::LEU2</u> <u>HMRα</u> <u>sir3^{ts}</u> <u>trp1</u> <u>ura1</u> <u>ade2</u> <u>his4</u>
M29	<u>HMLα</u> <u>mat::LEU2</u> <u>HMRα</u> <u>sir3⁰::TRP1</u> <u>trp1</u> <u>ura1</u> <u>ade2</u> <u>his4</u>
M37	<u>HMLα</u> <u>mat::LEU2</u> <u>HMRα</u> <u>sir4^{ts}</u> <u>ade2</u> <u>can1-100</u>
M48	<u>HMLα</u> <u>MATα2</u> <u>matα1</u> <u>LEU2</u> <u>hmr::TRP1</u> <u>ade2-1</u> <u>can1-100</u> <u>his3-11,15</u> <u>leu2-3</u> <u>trp1-1</u> <u>ura3</u>
Y65	K700 transplaced to <u>HMRΔ 77-268</u>
RS1	<u>HMLα</u> <u>MATα</u> <u>HMR::TRP1</u> <u>trp1</u> <u>leu2</u> <u>ura3</u> <u>his3</u> <u>ade2-1</u> <u>can1-100</u>
RS3	<u>HMLα</u> <u>mat::LEU2</u> <u>HMR::TRP1</u> <u>trp1</u> <u>leu2</u> <u>ura3</u> <u>his3</u> <u>ade2-1</u> <u>can1-100</u>
K123	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>sir3^{ts}</u> <u>ade2</u> <u>leu2</u> <u>trp1</u>
K124	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>sir4^{ts}</u> <u>ade2</u> <u>leu2</u> <u>trp1</u>
K163	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>SIR⁺</u> <u>ade1</u> <u>gal1</u> <u>his3</u> <u>leu1</u> <u>met14</u> <u>trp1</u> <u>ura3</u>
K699	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>ade2-1</u> <u>can1-100</u> <u>his3-11,15</u> <u>leu2-3</u> <u>trp1-1</u> <u>ura3</u> aka W303-1A
K700	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>ade2-1</u> <u>can1-100</u> <u>his3-11,15</u> <u>leu2-3</u> <u>trp1-1</u> <u>ura3</u> aka W303-1B
K736	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>trp1</u> <u>leu2-1</u> <u>ade2-1</u> <u>can1-100</u> <u>his3</u> <u>his4(?)</u> <u>Met⁻</u>
K765	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>ho::SUP4-o</u> <u>trp1-1</u> <u>ade2-1</u> <u>can1-100</u> <u>his3</u> <u>leu2-3,112</u> <u>ura3</u>
K822	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>ade2</u> <u>can1-100</u> <u>his4</u> <u>leu2</u> <u>trp1-1</u> <u>ura3</u>
K1114	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>ho:lacZ</u> <u>can1-100</u> <u>his4</u> <u>leu2</u> <u>trp1</u> <u>ura3</u>
ts368	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>rna2^{ts}</u> <u>ade1</u> <u>ade2</u> <u>his7</u> <u>lys2</u> <u>tyr1</u> <u>ura1</u>
ts368 α 25	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>rna2^{ts}</u> <u>ade1</u>
AB10	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>SIR⁺</u> <u>can1</u> <u>gal2</u> <u>his6</u> <u>leu1</u> <u>met1</u> <u>trp5-1</u>
AB20	<u>HMLα</u> <u>MATα</u> <u>HMRα</u> <u>SIR⁺</u> <u>can1</u> <u>gal2</u> <u>his6</u> <u>leu1</u> <u>met1</u> <u>trp5-1</u>

TABLE 2.2

Bacterial strains used

E. Coli strain	Genotype	Source or Reference
JM101	$\Delta(\text{lac}, \text{pro}), \text{supE}, \text{B1}^-$, F'(traD36, proAB lacI ² , lacZ Δ M15)	Messing, 1979
C600	thr ⁻ , leuB6, tonA21, supE44, galK ⁻	K. McKenney
TG1	$\Delta(\text{lac}, \text{pro}), \text{supE}, \text{B1}^-$, hsdd5 F'(traD36, proAB, lacI ^q , lacZ Δ M15)	T.J. Gibson
DH1	recA1, hsd, R17(r _k ⁻ m _k ⁺), endA1 supE44, thi-1, gyrA96, rela1	Hanahan, 1983

CHAPTER THREE

ASSAYS FOR SIR REPRESSION

3.1 Introduction

The first system of gene control which I have looked at is the SIR system. The sequences that are required in cis for repression have been mapped by deletion analysis of the repressed loci and have been found to lie some distance from the affected promoters. In experiments performed on plasmids, a site to the right, "E", is found to be essential for repression, while a site to the left, "I", is important. deletion of "I" leads to partial derepression (Abraham et al., 1984; Feldman et al., 1984). If however the in vitro constructed deletions are used to replace the chromosomal copy of the HMR locus, then only the "E" sequence appears to be required (Brand et al., 1985). Brand has shown that the "E" site can be thought of as a negative 'enhancer' for a number of reasons; firstly it can act at a distance - repressing transcription from promoters up to 2.6 kb away. Secondly, the relative orientations of the E element and the affected promoter are not significant, and finally the E element can repress transcription from promoters unrelated to the mating type system.

Mapping of the E element has shown that it is intimately associated with an ARS element, as is the I element (Abraham et al., 1984). ARS elements confer the ability to replicate autonomously and are therefore believed to be specific origins of DNA replication (Stinchcomb et al., 1979; Beech et al., 1980; Chan and Tye, 1980; Newlon and Burke, 1980; Celniker and Campbell, 1982). This suggests that DNA replication may be involved in the mechanism of SIR action.

To test this idea, one can ask whether the SIR-mediated repression of the silent loci can be established (or lost) in the absence of DNA replication. This can be done using temperature-

sensitive (ts) alleles of SIR genes. The paradigm is to shift the temperature of a yeast culture and to follow the subsequent change in status of the silent loci. One can then ask whether inhibitors of cell cycle progress affect the change in HML and HMR.

There are various possible assays for the Sir status of the silent loci. Perhaps the most obvious is the rate of transcription, since the primary consequence of SIR action is the repression of transcription. If one monitors a silent copy transcript with a reasonably fast turnover rate, then the steady state level of that transcript should reflect the status of the silent loci. Alternatively, one could use the apparent change in chromatin structure that is associated with SIR-mediated repression. Nasmyth (1982b) has shown a change in pattern of DNase I sensitivity that is dependent on SIR control, and Abraham *et al* (1982) have shown a shift in the linking number of plasmids carrying HMR between SIR⁺ and SIR⁻ strains. Either of these measures of chromatin structure could be used to monitor the change in Sir status.

In this chapter I describe preliminary investigations into the use of plasmid linking number and into the measurement of transcript levels as assays for Sir status.

3.2 Linking number as an assay for Sir status

3.2.1 Background

Abraham *et al* (1982) observed that the linking number of a plasmid containing HMR differed, depending on whether the plasmid was isolated from a SIR⁺ or a SIR⁻ strain, this shift was not observed in a control plasmid that lacked the E site. In order to understand the implications of this result we must consider what "linking number" is

and how it is measured. Linking number is a topological property of a pair of closed loops, describing the number of times the path of one loop passes through the other loop. It can only be changed by operations which involve cutting and religating one of the loops (Figure 3.1). If we think of the Watson and Crick strands of an unnicked DNA circle, then we can see that one strand passes through the circle formed by the other strand approximately once every 10.5 base pairs as long as the plasmid is relaxed and is under no torsional stress. If however one strand has been nicked, and a section unwound and religated, then this plasmid is now under torsional stress because its linking number has been reduced. The plasmid can do two things to accommodate this change.

1) It can change the winding angle of the DNA so that one strand rotates about the other once every 11 base pairs, say, rather than 10.5. This is referred to as ΔTw , for change in "twist".

2) It can twist the path of the double helix back upon itself to alleviate the torsional stress. This is known as ΔWr , for change in "writhe".

$$\Delta Lk = \Delta Tw + \Delta Wr$$

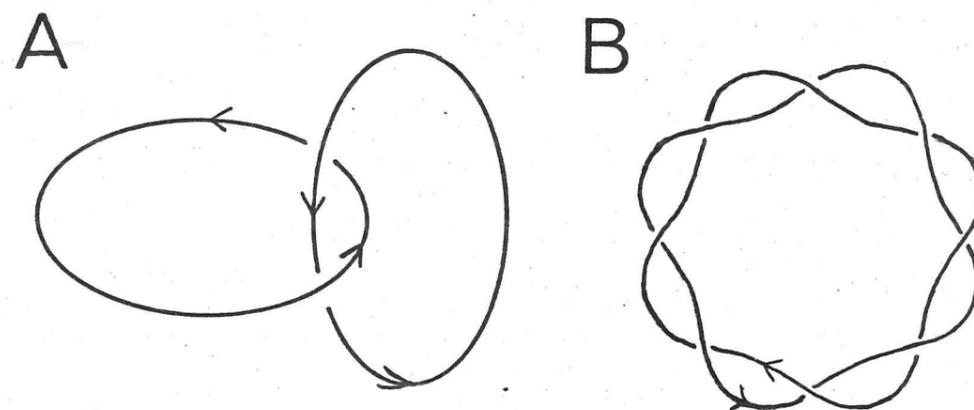
Note that ΔLk is a mathematical property of the plasmid which cannot be altered without breaking a strand, whereas Tw and Wr can vary so long as their sum remains a constant.

Circular DNA molecules isolated from cells tend to be negatively supercoiled, i.e. their linking number is lower than would be expected for relaxed DNA. In eukaryotes torsionally stressed DNA is rapidly relaxed by the action of "nicking-closing enzyme", which nicks one DNA strand, allows it to rotate about the other strand and then reseals the nick. These two facts appear to be contradictory, but in

FIGURE 3.1

Linking Number

Linking Number is a topological property of two closed loops, which describes the number of times the path of one loop passes through the other loop. In (A), each loop passes through the other once. By giving each of the paths a direction, indicated by the arrow, we can distinguish right-handed twisting from left-handed. In DNA topology, the arrangement shown in (A) is defined as $Lk=+1$, and its mirror image would be -1 . As a result, (B) has $Lk=+4$, and right-handed DNA has positive linking number, left-handed DNA negative. (The reverse convention is adopted in Mathematics.)



eukaryotes the DNA is wrapped around histone proteins, which leads to a negative writhe (Klug and Lutter, 1981): thus when the proteins are removed, the DNA becomes torsionally stressed. Higher order chromatin structures which affect the way the DNA's path twists about itself will affect the writhe, and local melting of the double helix (as might be caused by RNA polymerase) will affect twist. Thus we can learn something about the conformation of a circular DNA molecule in vivo by isolating the plasmid in an unnicked form and measuring its linking number.

Lk can be calculated by measuring the electrophoretic mobility of the plasmid through an agarose gel. a totally relaxed "open circle" migrates slowly through such a gel, while a highly writhed plasmid migrates much faster. For the assay of Sir status we are only interested in changes in linking number and therefore in changes in writhe for the naked DNA.

In order to aid the resolution of the different topoisomers of a given plasmid, the gel is run in the presence of an intercalating agent, which serves to reduce the winding angle of the DNA. this produces a negative ΔTw and therefore reduces the (negative) Wr that the plasmid has. Smaller negative Wr 's are resolved more easily.

3.2.2 Linking number of plasmids in sir^{ts} strains

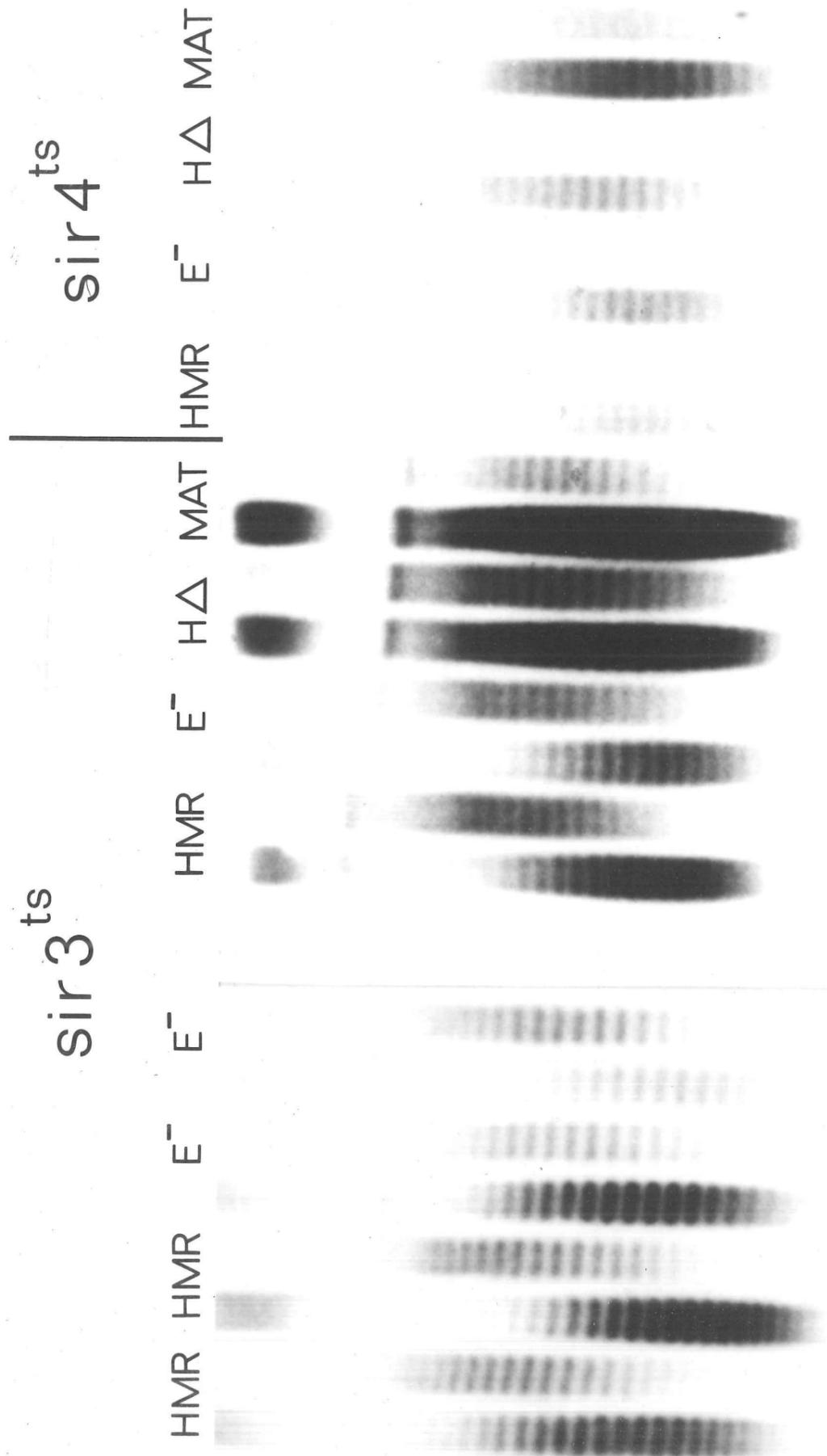
As mentioned above, Abraham et al found that $Lk_{sir^+} < Lk_{sir^-}$ when the plasmid JA82.6 was tested in a SIR⁺ versus a sir2⁻ strain. In contrast, a plasmid lacking the E element had identical linking numbers when isolated from the two strains. I therefore repeated this experiment, but using a sir3^{ts} strain grown at 23°C or 37°C to give the Sir⁺ or Sir⁻ states. Figure 3.2a shows that pJA82.6 has more negative writhe when extracted from Sir⁺ cells [a difference of 7 or

FIGURE 3.2

Effects of sir^{ts} alleles on the linking numbers of various plasmids

Plasmids were prepared from exponentially growing cells and electrophoresed as described in Materials and Methods. Gel lanes are arranged in pairs, and each pair shows a particular plasmid isolated from cells grown at 23°C (on the left) or 37 °C (on the right). Topoisomers with greater negative writhe will migrate further down the gel.

The yeast strains used were K123 (sir3^{ts}) and K124(sir4^{ts}). The plasmids used are all based on the plasmid pJA82.6, but with various alterations. (HMR) denotes pJA82.6, while (E⁻) denotes the deletion no. 77-268. In (HΔ), the left-hand side of HMRa has been replaced with MATa DNA by in vitro reconstruction, while (MAT) has the entire HMRa fragment replaced with MATa DNA.



8]. An identical effect is seen, however, with the plasmid no.77-268, which carries a deletion of the entire E region. Although Abraham's data imply that SIR2^{+/-} does not affect the linking number of plasmid no.77-268, it may however be affected by changes in the activity of the SIR3 gene product. To investigate this further, I looked at plasmids with less silent locus DNA: 1) the "Hawthorne deletion"(Hawthorne, 1963b), which contains the right hand side of the HMR locus fused to the left hand side of MAT (contains the I element but not the E element). 2) the MAT locus. Figure 3.2b shows that these two plasmids both show a shift in linking number when grown at 23°C or 37°C in the sir3^{ts} strain M28, the change is somewhat smaller - 4 or 5 - as compared with 7 or 8 for the E deletion plasmid or for the complete HMR plasmid JA82.6. When the four plasmids were tested in a sir4^{ts} strain, a similar effect was seen (Figure 3.2b).

The size of the shift seen in these different plasmids does not correlate with the sensitivity of the constructions to SIR-mediated repression. Some confusing factor must exist, which is leading to an artefactual shift. The fact that the MAT plasmid shows a shift of 4 to 5 in linking number suggests that the shift I see in pJA82.6 is largely not related to Sir status, but rather is a consequence of the temperature change.

From the data presented here, we cannot conclude that the linking number shifts are unconnected with Sir action, since we have not tested the effect of temperature on these plasmids in strains without sir^{ts} alleles. It is clear, however, from the failure to correlate the changes in linking number with SIR-dependence that the linking number shift observed in these strains is not an assay for SIR action which can be interpreted unambiguously.

3.3 RNA levels as an assay for Sir status

Steady state RNA levels were also considered as an assay for the Sir status of the silent loci. The expression of the mating type gene a1 was investigated, using the technique of protecting radiolabelled DNA probes from digestion by S1 nuclease. Such experiments were first done using DNA probes that had been end labelled and separated from their complementary strand (Berk and Sharp, 1978). The DNA probe is then hybridised to RNA and the hybrids are digested with S1 endonuclease. S1 will digest single stranded DNA but not double stranded DNA nor RNA-DNA hybrids (Figure 3.3). After digestion, the products are electrophoresed through a gel and the labelled fragments are visualised by autoradiography. In a modification of this technique by D.Bentley, the radio-labelled probe DNA is prepared by in vitro DNA synthesis, using an M13 template and Klenow polymerase (The "prime cut" method, see Section 2.11). This method produces a DNA probe that is labelled throughout its length, rather than at only one end.

If the hybridisation is performed with excess probe DNA then the strength of the signal produced is proportional to the steady state level of the RNA complementary to the probe, and therefore the signal can be used to quantitate RNA levels.

The MATa1 transcript had been mapped previously by Nasmyth (Nasmyth et al, 1981a), using end-labelled probes. It was decided to use continuously labelled probes because they should provide a stronger radioactive signal and allow a simpler technique for probe preparation. Templates for the DNA probes were prepared by subcloning fragments of MATa1 DNA into M13 sequencing vectors. Due to the paucity of restriction enzyme sites in the MATa1 gene, Xho-linker insertions (Abraham et al 1982) were used as the source of DNA for subcloning. The

regions of DNA sequence covered by the probes is shown in Figure 3.5.

Single stranded DNA, radiolabelled throughout its length, was prepared by the 'prime cut' method, and used for the hybridisation and S1 digestion procedures. From the previous mapping data, Probe A was expected to produce a single protected fragment of approximately 490 nucleotides. Figure 3.4a shows that it did not. A series of radiolabelled fragments were produced; substoichiometric bands at 490 and 310, a strong band at 245 nucleotides and a heterogeneous band at 118. This dissection of the DNA probe into smaller fragments is consistent with the MATa1 transcript being processed. Such a result could, however, be caused by a number of possible artefacts which are considered below.

1) Alterations in the probe DNA sequence during Xho-linker mutagenesis and subsequent subcloning would cause digestion within the region covered by the transcript. This possibility was discounted by sequencing directly the single-stranded 'phage DNA used to make the probe. Its DNA sequence was shown to be identical to the known genomic sequence.

2) The fragments could be the result of an artefact of S1 digestion. To test this, the conditions and extent of S1 digestion were varied (Figure 3.4b, lanes 1-7). A slight movement of the bands, consistent with higher concentrations of nuclease 'chewing back' the protected fragments (Hentschel et al, 1980), but the relative intensity of the bands did not change significantly with digestion conditions.

It is also possible that the MATa1 gene contains sequences that are susceptible to S1 digestion even when in duplex form. However, when double-stranded MATa1 DNA was end-labelled and digested with S1, a potential S1 sensitive site 5' to the transcribed region was revealed, but no specific cleavage within the transcribed region was

regions of DNA sequence covered by the probes is shown in Figure 3.5.

Single stranded DNA, radiolabelled throughout its length, was prepared by the 'prime cut' method, and used for the hybridisation and S1 digestion procedures. From the previous mapping data, Probe A was expected to produce a single protected fragment of approximately 490 nucleotides. Figure 3.4a shows that it did not. A series of radiolabelled fragments were produced; substoichiometric bands at 490 and 310, a strong band at 245 nucleotides and a heterogeneous band at 118. This dissection of the DNA probe into smaller fragments is consistent with the MATa1 transcript being processed. Such a result could, however, be caused by a number of possible artefacts which are considered below.

1) Alterations in the probe DNA sequence during Xho-linker mutagenesis and subsequent subcloning would cause digestion within the region covered by the transcript. This possibility was discounted by sequencing directly the single-stranded 'phage DNA used to make the probe. Its DNA sequence was shown to be identical to the known genomic sequence.

2) The fragments could be the result of an artefact of S1 digestion. To test this, the conditions and extent of S1 digestion were varied (Figure 3.4b, lanes 1-7). A slight movement of the bands, consistent with higher concentrations of nuclease 'chewing back' the protected fragments (Hentschel *et al*, 1980), but the relative intensity of the bands did not change significantly with digestion conditions.

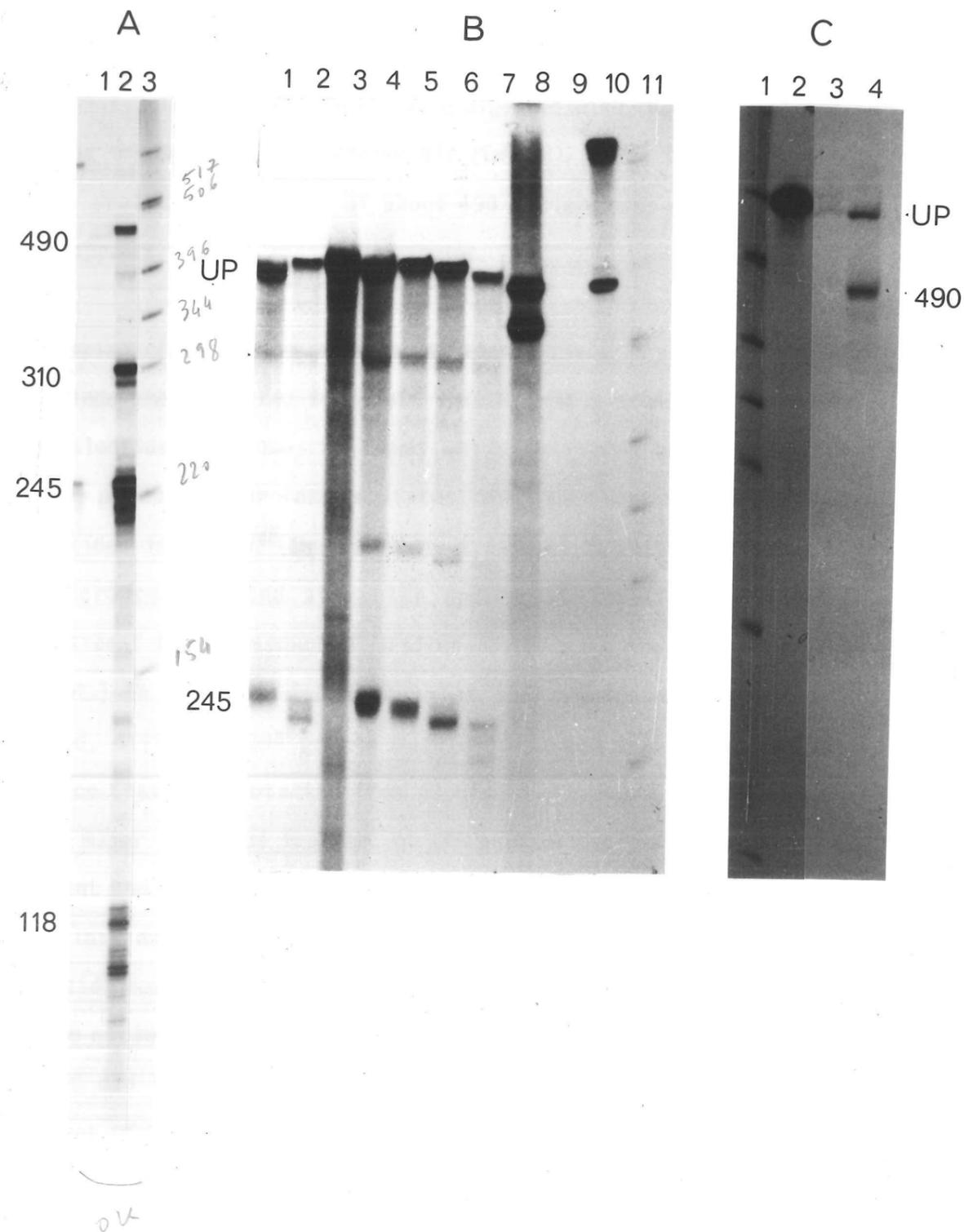
It is also possible that the MATa1 gene contains sequences that are susceptible to S1 digestion even when in duplex form. However, when double-stranded MATa1 DNA was end-labelled and digested with S1, a potential S1 sensitive site 5' to the transcribed region was revealed, but no specific cleavage within the transcribed region was

FIGURE 3.4

(A) S1 Mapping with probe A (Figure 3.5): DNA protected by 12 μ g of total RNA 1) from an α cell, 2) from an a cell, 3) pBR322 x HinfI (Sizes are 517, 506, 396, 344, 298, 221, 220, and 154 nucleotides.)

(B) Titration with S1: 50 μ g total RNA was digested under the following conditions, 1) 20u/ml, 20 $^{\circ}$ C; 2) 320u/ml, 20 $^{\circ}$ C; 3) 5u/ml, 37 $^{\circ}$ C; 4) 10u/ml, 37 $^{\circ}$ C; 5) 20u/ml, 37 $^{\circ}$ C; 6) 80u/ml, 37 $^{\circ}$ C; 7) 320u/ml, 37 $^{\circ}$ C. "UP" denotes undigested probe DNA. Digestion of duplex DNA: the duplex DNA from which probe A was made was end-labelled at the BglIII site and subjected to S1 digestion at 20u/ml, 37 $^{\circ}$ C (See Materials and Methods): Lane 8) digested DNA, 9) no DNA, 10) undigested end-labelled DNA, 11) pBR322 x HinfI. (The undigested DNA band is shorter than that in the RNA protection experiments, since it lacks the M13 primer. The higher molecular weight band in lane 10 is end-labelled vector DNA.)

(C) Protection of Probe A from Exo VII digestion: 25 μ g of total RNA was used; Lane 1) pBR322 x HinfI, 2) undigested probe, 3) α cell RNA, 4) a cell RNA.



observed (Figure 3.4b, lanes 8-10).

We conclude from these data that the 310, 245, and 118 nucleotide protected fragments correspond to segments of RNA. The protection could result from a single transcript that is spliced, or from a number of short, unspliced transcripts. To eliminate the latter possibility, the RNA/DNA hybrids were digested with the single-strand-specific exonuclease, exonuclease VII (ExoVII). Since this produces a single protected fragment of about 490 nucleotides (Figure 3.4c), there appears to be a single transcript.

3.4 Mapping of the intron-exon boundaries

Since the protected fragments result from a transcript covering 490 nucleotides of probe, the bands which correspond to exons, rather than to precursors or intermediates, must be those of 245 and 118 nucleotides in length. (The 118 nucleotide band appears faint since it is heterogeneous, and since the amount of label in a band is proportional to the number of uridine residues in the protecting RNA segment.) In order to determine what intervening sequences were excised, different DNA probes were used to map the regions of DNA sequence that are protected from S1 digestion. Probe B (Figure 3.5) gave a major protected fragment of 200 nucleotides instead of the 245 fragment and did not affect the 118 nucleotide fragment (See Figure 3.6a). This demonstrates that the 245 nucleotide fragment begins near nucleotide number 1684 and should therefore end at 1924-1934, and that the 118 nucleotide exon lies to the left of the 245 nucleotide exon. Probe C did not give any protection at 118 and gave a protected fragment of 190 nucleotides, thus confirming that the 3' end of the 245 nucleotide exon is at 1935, and the 118 nucleotide exon is to the left. Hence there is an exon extending from approximately nucleotide

FIGURE 3.5

Extent of DNA probes used, and interpretation of the S1 mapping data.

The arrow at the top of the figure denotes the transcript mapped previously by Nasmyth *et al.*, (1981). The dotted lines represent the extent of the probes used to map the transcript. The wide boxes represent the major protected fragments of probe DNA, and are aligned with the model for splicing proposed at the bottom of the figure. The narrow boxes represent the sub-stoichiometric fragments which may be due to RNA intermediates. The numbers on the scale correspond to those of Astell *et al.*, (1981) and in Figure 3.8.

Interpretation of the mapping data

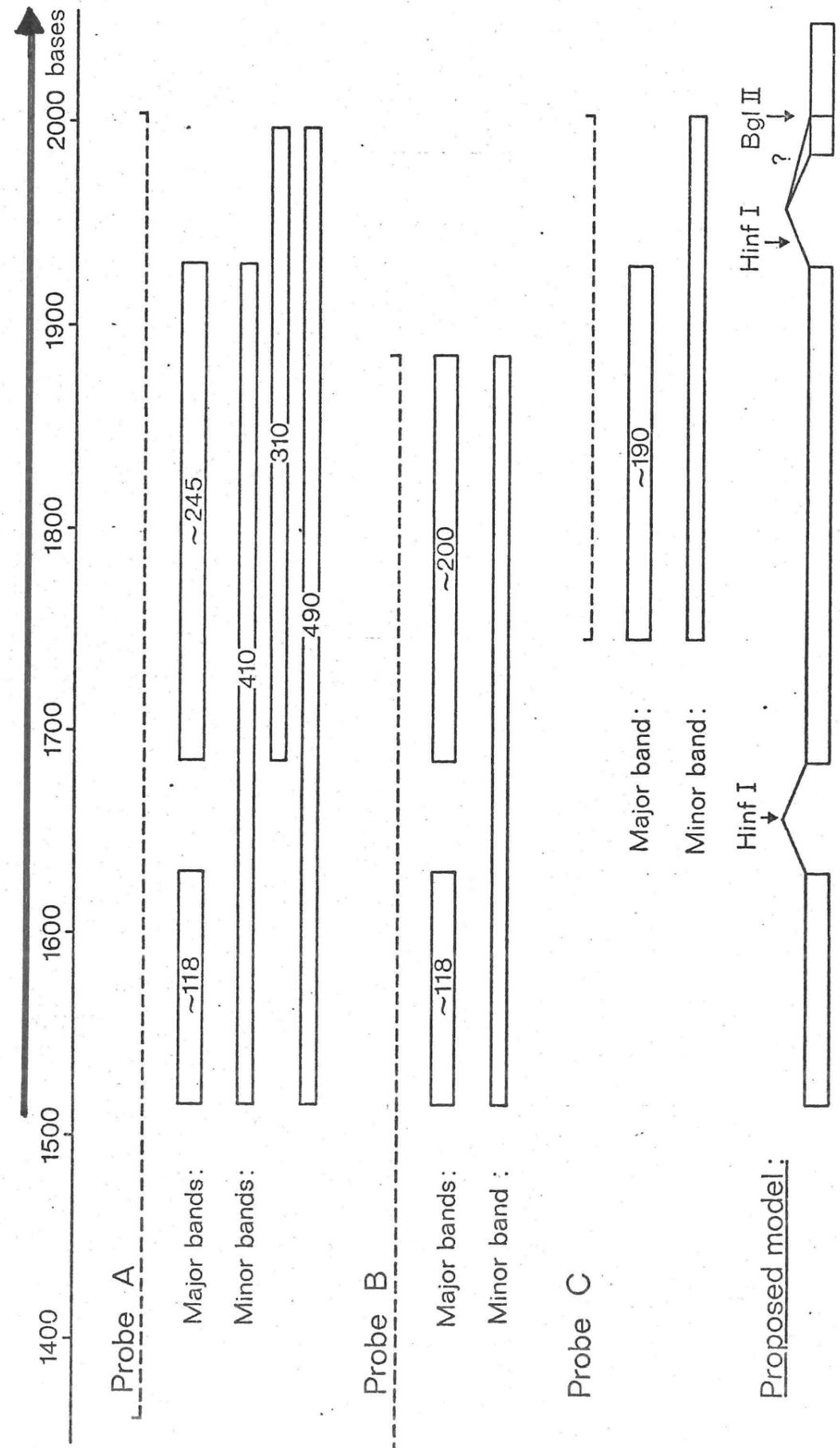


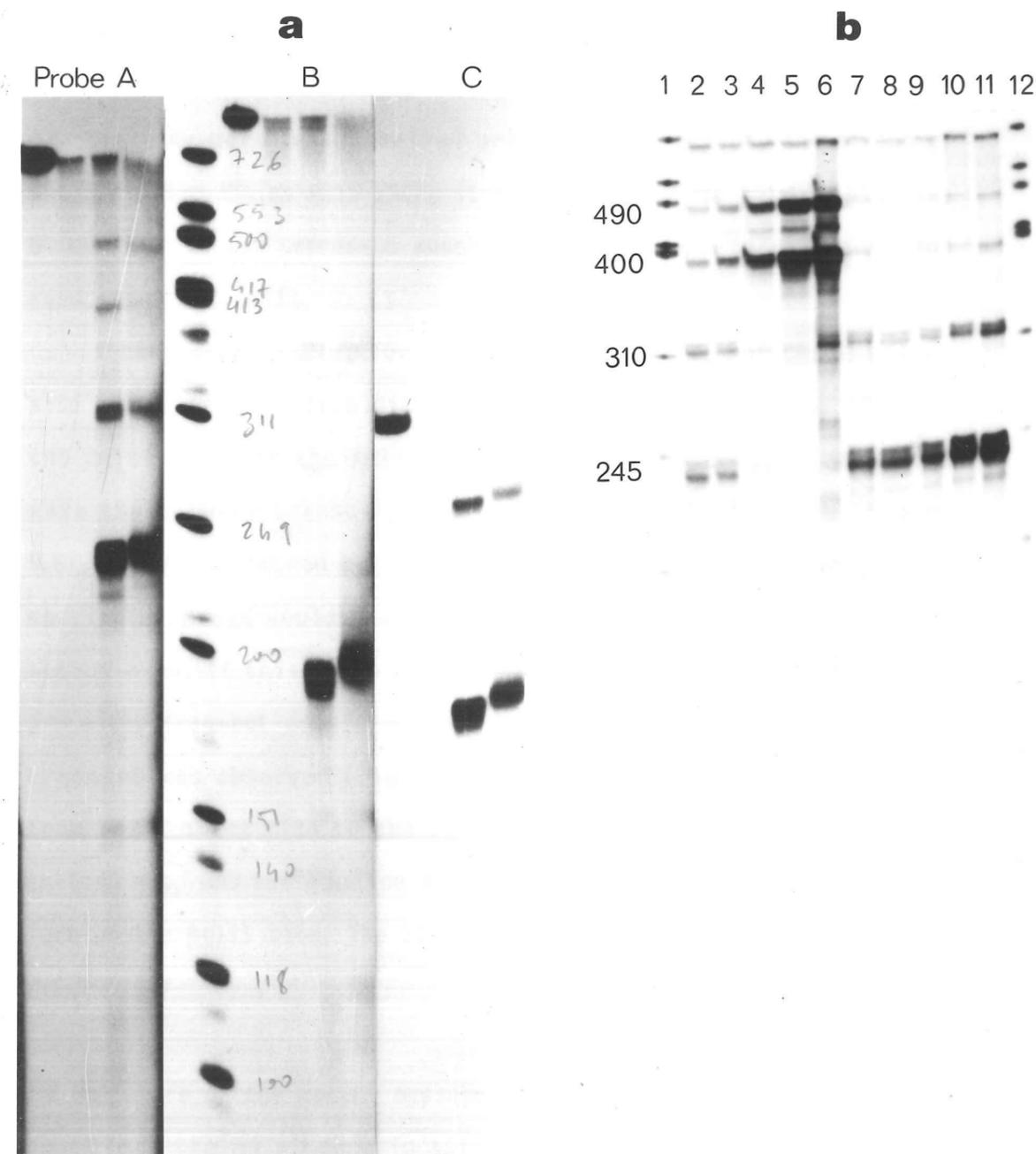
FIGURE 3.6

(A) Mapping the 245 bp Exon.

20 μ g of total RNA was hybridised to one of three probes and subjected to S1 digestion. Lanes 1-4) probe A, Lanes 6-9) probe B, Lanes 10-13) probe C. Within each group the order of tracks is: first undigested probe DNA, second α cell RNA, third a cell RNA and fourth a cell RNA, but digested at 20 $^{\circ}$ C rather than 37 $^{\circ}$ C. Lane 5 is ϕ X174 x HinfI. (Sizes are 726, 713, 553, 500, 427, 417, 413, 311, 249, 200, 151, 140, 118, 100 nucleotides.) The 118 nucleotide band is only faintly visible at this exposure; over-exposure of the autoradiogram demonstrates its presence.

(B) Effect of the rna2^{ts} mutation.

Lane 1) ϕ X174 x HinfI, Lanes 2-6, DNA protection by 25 μ g of total RNA taken from the rna2 mutant strain ts 368. Lane 2) at the permissive temperature (23 $^{\circ}$ C), 3) after 0 minutes at 34 $^{\circ}$ C, 4) after 10 min at 34 $^{\circ}$ C, 5) after 20 min at 34 $^{\circ}$ C, 6) after 60 min at 34 $^{\circ}$ C. Lanes 7-11, 25 μ g of total RNA taken from a spontaneous revertant of ts 368. Lane 7) at 23 $^{\circ}$ C, 8) after 0 min. 9) after 10 min, 10) after 20 min, 11) after 60 min at 34 $^{\circ}$ C. Lane 12) ϕ X x HinfI.



1684 to nucleotide 1935, with introns on either side.

The mapping of the 118 nucleotide exon was less direct. The data above demonstrate that it lies to the left of the 245 nucleotide exon. ExoVII digestion using Probe A gave a single RNA-dependent band of about 490 nucleotides, thus mapping the 5' terminus of the mRNA to nucleotide number 1513. This is consistent with the previous mapping of this gene using end-labelled probes, which defined the 5' end of the precursor RNA as 1515-1517; it would appear that this is also the 5' terminus of the processed message. The first exon therefore extends from about 1515-1633.

I have been unable, for technical reasons, to map the 3' splice site of the second intron directly, but it cannot lie to the right of the BglIII site for the following reason. The homology of the probe with the primary transcript ends at this restriction site (Figure 3.5), so if the intron extended beyond the BglIII site, then the RNA species produced would cease to protect the probe at the end of the second exon. If this were so, then ExoVII digestion would produce a protected fragment of 420 nucleotides. In fact only a 490 nucleotide fragment was observed (Figure 3.4c), corresponding to the distance from the 5' start site to the BglIII site. This site is therefore not spliced out, and the 3'splice site of the second intron must be to the left of the BglIII site. The S1 mapping data are summarized in Figure 3.5.

3.5 Half-life of the mature message

Since the MATa1 gene is spliced, we are able to put a maximum value on the half-life of the mature message under certain circumstances. Yeast strains carrying the rna2^{ts} mutation fail to carry out splicing reactions at the restrictive temperature (Rosbash

et al, 1981; Fried and Warner, 1982; Teem et al, 1983). The precursor accumulates, and levels of the processed message decay as this RNA species turns over. Thus if we shift an rna2^{ts} strain to the restrictive temperature and watch the processed message disappear we are able to estimate the half-life of the message under these conditions. Figure 3.6b shows the result when this is done with MATa1. After only 10 minutes at the restrictive temperature the intensity of the 245 nucleotide fragment has fallen more than 10 fold: the half-life of the mature message must be less than 3 minutes under these conditions. The half-life of the message does not appear to be altered dramatically due to the effects of heat shock, since the isogenic RNA2⁺ strain shows no significant change in the relative amounts of the various species. I cannot, however, exclude the possibility that the rna2^{ts} mutation itself affects the half-life of the processed message.

The protected fragments which accumulate at the restrictive temperature have been mapped using probes B and C. This confirms that the 490 nucleotide fragment corresponds to unspliced precursor, while the 400 nucleotide fragment seems to extend from about nucleotide 1600 - 2003, since it produces a new protected fragment of 280 nucleotides with probe B (Figure 3.7). The appearance of this band is strain-dependent and is correlated with the appearance of a fragment at 85 nucleotides, implying that both fragments result from an S1 sensitive site at 1600.

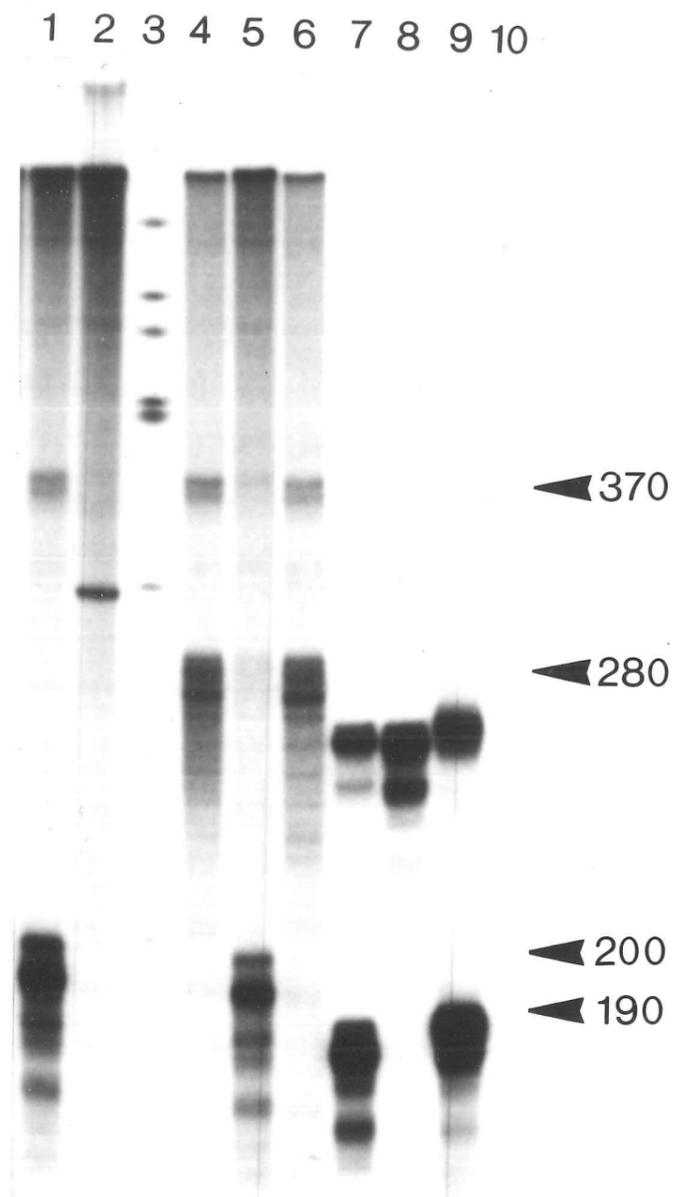
3.6 Investigation of the 400 nucleotide fragment

It is not clear what is causing the protected fragment of 400 nucleotides. Since it accumulates at the restrictive temperature it could be a consequence of the rna2^{ts} mutation. To test this two

FIGURE 3.7

Mapping the RNA species in the *rna2^{ts}* strain.

Some of the RNAs used in Figure 3.6b were hybridised with probes B and C and S1 nuclease protection measured. Lanes 1-6, probe B, lanes 7-10, probe C (Figure 3.5). Lanes 5 and 7 use RNA from a spontaneous revertant of ts368, taken after 60 min at 34°C. Lanes 6 and 8 use RNA from ts368 taken after 60 min at 34°C, while lane 4 uses the 20 min timepoint. 25µg of total RNA were used per lane. Lanes 1 and 9 show *a1* RNA from an unrelated strain (M29), for comparison, and lanes 2 and 10 use RNA that lacks *a1* message (M28 grown at 23°C.) For these four lanes, 40µg of total RNA were used per lane.



crosses were performed. The strain used in Figure 3.6b was crossed to strain AB20, and an α rna2^{ts} strain was crossed to AB10 (which is the strain from which the MATa locus was cloned and sequenced (Astell et al, 1981): this allele of MATa is known to have a sequence identical to the probe DNA sequence.)

In the first cross, 6 MATa segregants were tested - all showed the band at 400 despite the fact that 4 were RNA2⁺.

In the second cross, 8 MATa segregants were tested - none showed the band at 400 although 6 of them were rna2^{ts}. The 400 nucleotide fragment is therefore not a necessary consequence of the rna2^{ts} mutation. The first cross furthermore suggests that the band phenotype may be linked to the MATa locus. These results, and a survey of the presence of the 400 band in a number of different yeast strains, are consistent with the hypothesis that there is a polymorphism at MATa, which could then cause partial S1 digestion at the point of mismatch between the probe and the RNA.

3.7 Discussion of the Data on the Splicing of the a1 gene

The S1 protection data described above maps the first two exons to an accuracy of 5 - 10 nucleotides. One way of mapping the splice sites exactly would be to isolate and sequence cDNA clones derived from a1 RNA; such clones should reveal the sequence of the mRNA after the excision of the intervening sequences. This approach was attempted without success. As an alternative approach, the MATa1 sequence was searched for homologies with the sequences found at other yeast introns. These are: G \blacktriangle GTATGT at the 5' splice site and a signal sequence 5'-TACTAACA-3' that directs splicing at the next AG following, thus defining the 3' splice site (Langford and Gallwitz, 1983; Pikielny et al, 1983). In the MATa1 sequence (Figure 3.8), 5'

splice sites are found at 1631 and 1931 and the TACTAACA signal is found preceding AGs at 1684 and 1982, in agreement with the mapping data. Although the 3' splice site of the second intron has not been mapped directly, we have deduced that it must lie to the left of the BglIII site. There is only one AG between the TACTAACA signal and the BglIII site - the one at 1982. I conclude, therefore, that the MATa1 gene contains an intron from nucleotides 1631 to 1684 and another from 1931 to 1982. The previous mapping of the MATa1 gene failed to detect this processing since the DNA probes were end-labelled at HinfI restriction sites that fall within introns. The use of continuously labelled probe DNA avoids this danger (See Figure 3.3).

The splicing pattern proposed is consistent with the previous data on this gene. The first intron preserves the reading frame and removes the UGA codon that was previously believed to be read through by a natural suppressor tRNA. The nonsense mutations that have been sequenced fall within the second exon (Klar, 1980; J.Abraham and K.Nasmyth, personal communication), and the processing pattern is consistent with the deletion analysis of this gene. One Xho-linker mutation gives a surprising phenotype: the mutation $\alpha X20$ (Tatchell et al, 1981) falls entirely within the first intron (Figure 3.8), but is MATa1⁻. This deletion removes most of the intron including part of the TACTAACA signal, and is probably defective in processing. The splicing pattern also explains the originally curious result (K.Nasmyth, personal communication) that deleting from the BglIII site rightwards abolishes MATa1 function - in the processed message the open reading frame extends beyond this site. It appears that the processing model proposed is sufficient to explain the previous data and no read-through of stop codons need be postulated.

In the S1 mapping experiments, sub-stoichiometric bands of higher

molecular weight were observed in addition to the major bands seen at about 245 and 118 nucleotides. The 490 nucleotide fragment accumulates in rna2^{ts} strains at the restrictive temperature, and is the result of protection by the unspliced precursor RNA. The band at 310 nucleotides does not disappear with increasing S1 concentration, so it is probably not an intermediate in S1 digestion (i.e. a partially digested hybrid), but rather represents protection by another RNA species. A protected fragment of 310 nucleotides would be produced by an RNA species in which only the first intron had been excised (see Figure 3.5). This RNA species may be a splicing intermediate.

Shepherd et al (1984) have shown that the MATa1 and the MATa2 gene products show significant homology with the 'Homeo domain' found in Drosophila and Xenopus. Furthermore, if the MATa1 RNA with only the first intron excised is translated then it too produces a protein homologous to the 'Homeo domain'. Surprisingly, this homology extends into the region encoded by the second intron, suggesting that the coding capacity of the second intron is under selective pressure. This observation, combined with the fact that the second intron does not appear to be excised efficiently, suggests that the RNA species with only the first intron excised may be translated to produce a functional protein. The MATa1 gene may therefore, by differential splicing, encode two proteins of related functions.

3.8 Conclusions

In this chapter I have looked at two different assays, by which one might be able to monitor the Sir status of the silent loci following a temperature shift in a sir^{ts} strain. The first technique - measuring the linking number of a plasmid carrying the HMR locus - appears to be beset by an artefactual linking number shift which

results from the temperature change, or from the use of the sir3^{ts} and sir4^{ts} alleles. Whatever the cause, the linking number shift seen with different plasmids bears little relation to their sensitivity to SIR-mediated repression. Since it is not clear what the linking number shift is measuring, linking number cannot be used satisfactorily as an assay of Sir status in the sir3^{ts} or sir4^{ts} strains tested.

The second technique - measuring the steady state level of a1 RNA - appears suitable. The discovery that the a1 gene is spliced has allowed the measurement ^{of} its half-life, which is approximately 3 minutes or less, under the conditions tested. Thus, over a timescale of tens of minutes, changes in the rate of transcription of a1 should be reflected in changes in the steady state level of the transcript. Furthermore, since the unprocessed message and what may be a splicing intermediate can also be detected, dramatic changes in the turnover rate of the a1 RNA should not go unnoticed.

CHAPTER FOUR

ROLE OF DNA REPLICATION IN SIR REPRESSION

4.1 Introduction

In this chapter I will describe experiments investigating the role of DNA replication in the SIR-mediated repression of the silent loci.

Deletion analysis of the silent copies has demonstrated that HML and HMR are under the negative control of their flanking DNA sequences (Abraham *et al.*, 1982). A small segment of DNA about a kilobase to the left of each silent copy promoter (called HMLE for HML and HMRE for HMR) is essential for their repression. Also required for the repression of both silent copies are four trans-acting gene products SIR 1,2,3 and 4 (Haber and George, 1979, Klar *et al.*, 1979, Rine *et al.*, 1979, Rine, 1979).

Several observations suggest that one can consider HMRE as a negative enhancer: first, HMRE can repress a promoter up to 2.5 kilobases away. Second, HMRE can repress other promoters when substituted for the a promoter at HMR. Third, the relative orientations of both HMRE and the regulated promoter are unimportant (Brand *et al.*, 1985).

The observation that the HMRE sequence is very tightly linked to an ARS sequence (Abraham *et al.*, 1984) suggests that DNA replication may be important in SIR-mediated repression (ARS sequences confer the ability to replicate autonomously and are probably specific origins of DNA replication (Stinchcomb *et al.* 1979; Beech *et al.*, 1980; Chan and Tye, 1980; Newlon and Burke, 1980; Celniker and Campbell, 1982)). In this chapter I describe experiments

which look at the possible role of DNA replication in SIR repression. More specifically, I have investigated the expression of HML and HMR in strains carrying *ts* alleles of SIR3 or SIR4. I have asked whether DNA replication is required for repression or for derepression of HML and HMR in these strains. The results suggest that DNA replication is required for the onset of repression (after a shift to the permissive temperature), whereas derepression (after a shift to the restrictive temperature) can occur in the absence of cell cycle progress.

4.2 Repression of HML and HMR requires cell cycle progress

The first question I wish to address is whether DNA replication is required for SIR repression. I have therefore tested whether the onset of SIR repression can occur in growing cells in the absence of any cell cycle progress. To do this, I have used strain M28 (HMLa MATa HMRa sir3^{*ts*} (See table 2.1)). This strain should produce a1 transcripts from both of its silent copies at the restrictive temperature but none at all at the permissive temperature. Since the MATa1 primary transcript is processed and the mature mRNA produced has a short half-life ($t_{1/2}$ 3 min. (See chapter 3)), the level of the mature RNA or of its precursors can be used as a measure of the instantaneous rate of transcription. Cells of M28 grown into stationary phase (G_0) at the restrictive temperature transcribe HMLa1 and HMRa1 at a full level. I have analysed the rate of a1 mRNA transcription after these cells are inoculated into fresh medium at the permissive temperature and tested whether it is affected by blocking cell cycle progress with alpha factor. (Alpha-factor is a yeast pheromone which blocks a cells in early G_1 in preparation for mating but does not directly

affect macromolecular synthesis. (For review, see Thorner, 1981))

Figure 4.1a shows the results of such an experiment, using S1 mapping to measure a1 mRNA. In the absence of alpha-factor, repression (as measured by a reduction in the level of a1) starts between 2.5 and 4 hours and is virtually complete by 10 hours. In the presence of alpha-factor, however, no repression occurs even after 10 hours. The effect of alpha factor cannot be due to some gross inhibition of cell growth, since a significant degree of repression occurs within 4 hours (i.e. at a time when alpha-factor has no appreciable effect on overall cell growth (as measured by optical absorbance at 660 nanometres or RNA synthesis (data not shown)).

The inhibition of a1 mRNA disappearance by alpha factor suggests that cell cycle progress is required for the repression of previously derepressed copies of HML and HMR. This result could, alternatively, arise if SIR3 synthesis or the activity of the a1 promoter were directly affected by alpha factor or by cell cycle progress. Figure 4.1a shows that the level of the SIR3 transcript is unaltered throughout the experiment, both in the presence and absence of alpha factor. The effect of these treatments on the activity of the a1 promoter can be followed in strain M29, which is isogenic with M28 but sir3⁰ (Table 2.1). When G₀ cells from M29 are inoculated into fresh medium (plus or minus alpha-factor), no significant change in the level of a1 is seen (Figure 4.1b) - demonstrating that the a1 promoters at the silent copies are not affected by these treatments.

The changes observed in the level of a1 transcript cannot be due to a change in its half-life, since we can follow the processing of this message. The MATa1 gene is spliced and the

FIGURE 4.1a Levels of a1 RNA in strain M28 after shift to the permissive temperature.

Strain M28 (Table 2.1) was grown into G₀ and purified daughter cells were inoculated into fresh medium at 23°C in the presence or absence of 10 units/ml alpha-factor.

Lanes 1-8 cells grown into G₀ at the permissive temperature (23°C).

Lanes 10-20 Cells grown into G₀ at the restrictive temperature (34°C).

The time elapsed (in hours) since inoculation is shown.

25µg total RNA were used per reaction.

This experiment yields the same result if purified mothers or total G₀ cells are used. Daughters are used here to allow more SIR3 synthesis at the permissive temperature in the alpha factor blocked cells.

FIGURE 4.1b Transcription from the a1 promoter is not affected by alpha factor or by cell cycle progress

Strain M29 (Table 2.1) is isogenic with M28, except that it is sir3⁰, and therefore expresses HML and HMR constitutively.

The effect of alpha factor on the level of a1 RNA.

cells of strain M29 were grown into G₀ at 34°C and inoculated into fresh YEPD at 23°C. After 7 hours growth, RNA samples were taken.

Lane 1 Cells grown without alpha factor

Lane 2 Cells grown in the presence of 10 units/ml alpha factor.

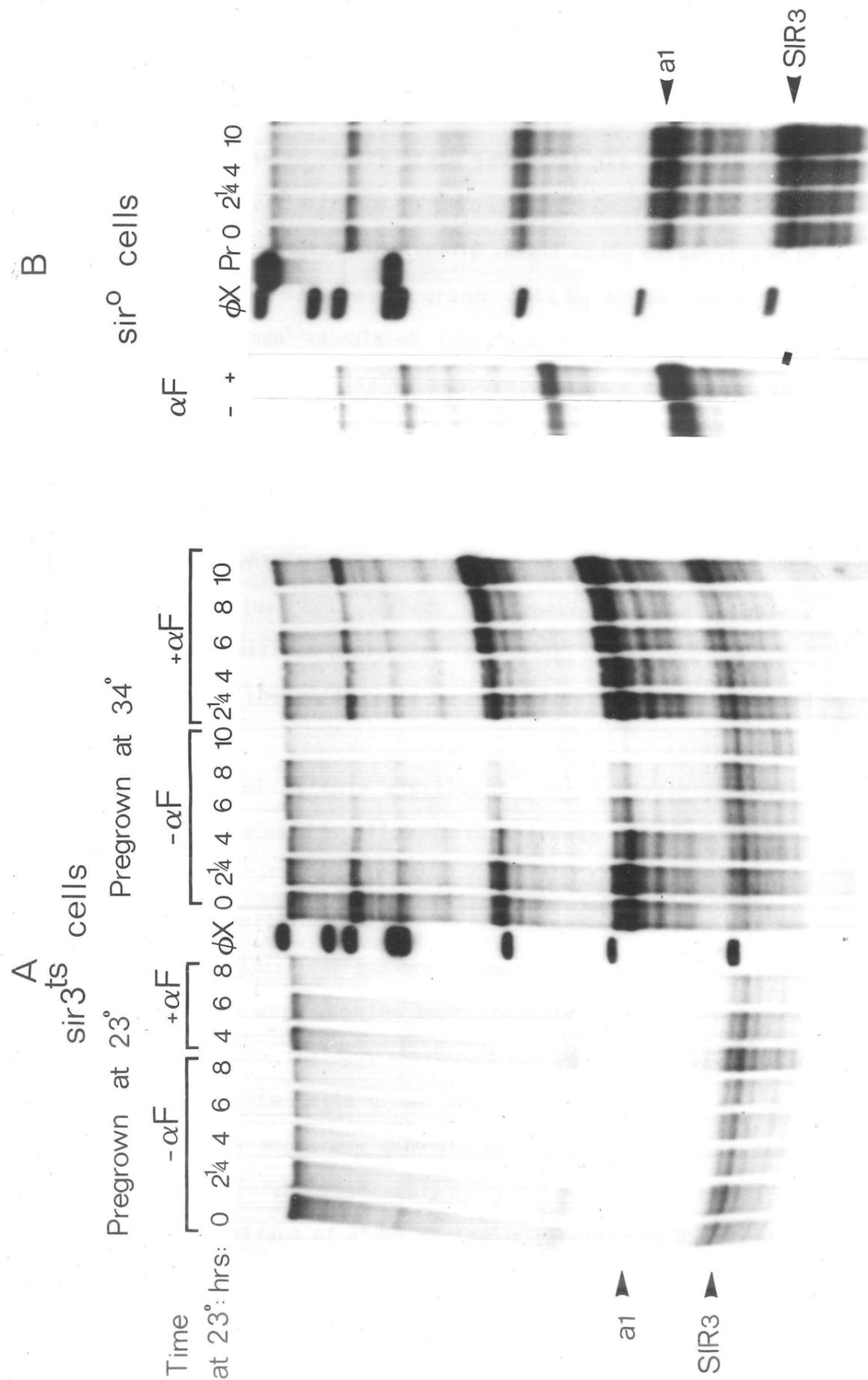
Lane 3 phi X HinfI.

Lane 4 Undigested probe. (Pr)

40µg of total RNA were used per reaction.

The effect of cell cycle progress on the level of a1 RNA.

cells of strain M29 were grown into G₀ at 34°C and purified daughters were inoculated into fresh YEPD at 23°C. RNA samples were taken at intervals: 5) after 0 hours, 6) after 2.25 hours, 7) after 4 hours; 8) after 10 hours. 25µg total RNA were used per reaction.



larger fragments observed correspond to unspliced precursor and what is probably a splicing intermediate (Chapter 3); the coordinate fall in the steady state levels of all these species must therefore reflect a change in the rate of transcription.

Finally, alpha factor does not in itself cause derepression of the silent copies. M28 cells grown into G_0 at the permissive temperature and then inoculated into fresh medium (with or without alpha factor) at the permissive temperature show no silent copy transcription (Figure 4.1a).

Thus, alpha factor does not directly affect a1 or SIR3 transcription, and neither of these transcripts shows any cell cycle dependence, yet alpha factor blocks the onset of repression of both HMLa and HMRa. Since this effect is observable prior to any difference in overall macromolecular synthesis, we conclude that progress through the cell cycle is necessary for SIR-mediated repression.

The cells used in this experiment were G_0 daughter cells; daughter cells were used to allow maximum growth at the permissive temperature in the cell-cycle arrested cells, and these daughters were purified from a culture arrested in stationary phase, since the separation of daughters from mothers is easier with these cells. Thus, although the cells were showing approximately exponential growth during the experiment, the result observed may be peculiar to daughter cells, or peculiar to cells grown out of stationary phase. (For example, stationary phase may generate some special chromatin state which cannot be altered immediately by cells growing out of G_0 .) To test whether the effect of alpha factor is peculiar to G_0 daughter cells, the experiment was repeated using an exponentially growing culture containing both mother and daughter cells. Figure 4.2 shows

FIGURE 4.2 Effect of alpha factor on SIR repression in exponentially growing cells

Strains M28(*sir3^{ts}*) and M29(*sir3⁻*) were grown in exponential culture at 34°C and were transferred to 23°C either in the presence or absence of alpha factor at 10 units/ml.

Lanes 1 and 2 M28 growing at 23°C

Lane 3 M28, pregrown at 23°C, after 6 hours in alpha factor.

Lanes 4 and 5 M28 growing at 34°C

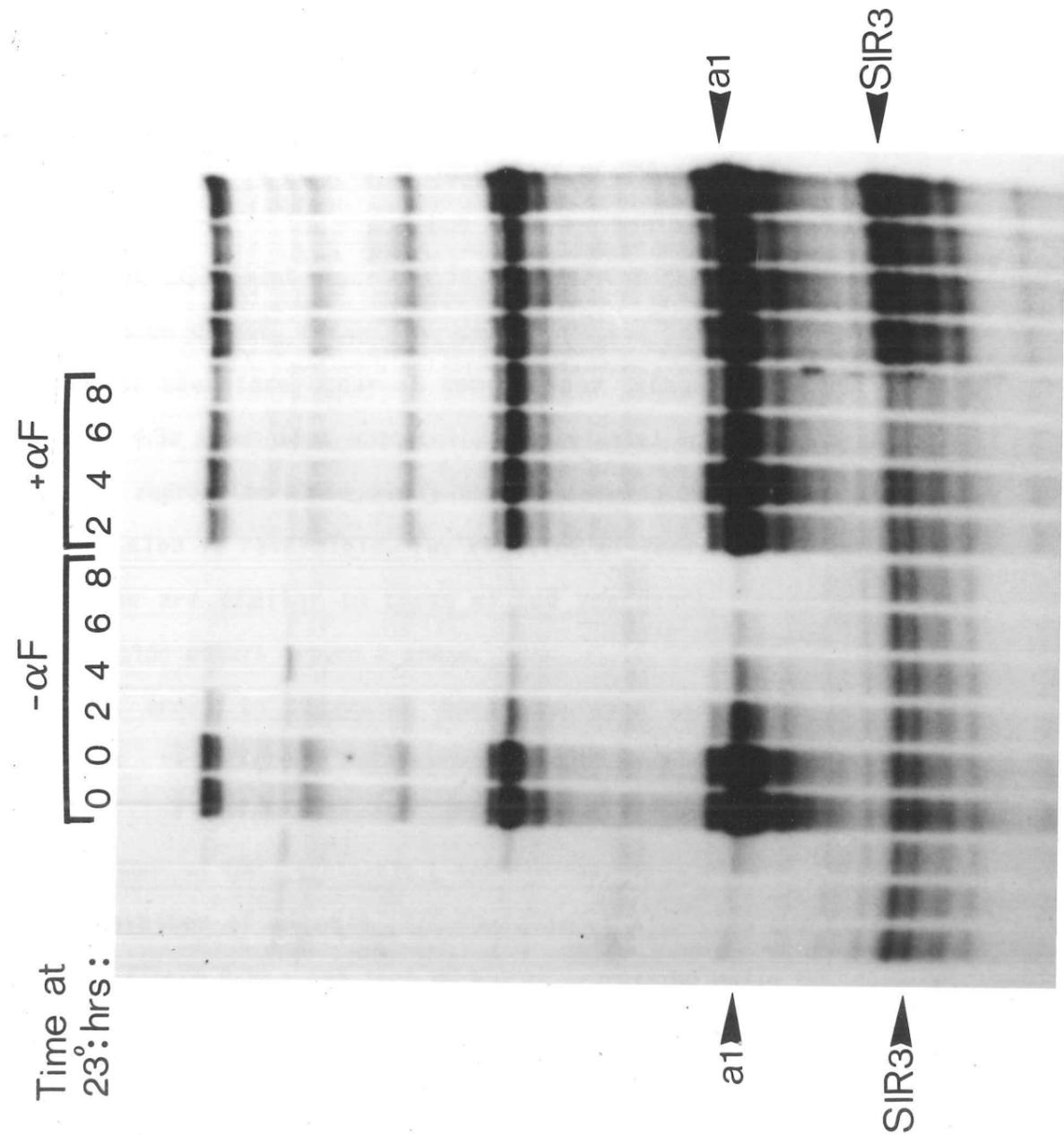
Lanes 6-9 M28 2,4,6,8 hours after a shift to 23°C

Lanes 10-13 M28 2,4,6,8 hours after a shift to 23°C in the presence of alpha factor

Lanes 14 and 15 M29 grown at 34°C

Lane 16 M29 after 6 hours at 23°C

Lane 17 M29 after 6 hours at 23°C, in the presence of alpha factor.



that repression of the silent loci occurs rapidly upon a shift to the permissive temperature, but is blocked by the presence of alpha factor. Figure 4.2 also shows control experiments similar to those performed in Figure 4.1. The effect of alpha factor on the repression of the silent loci in strain M28 is not, therefore, peculiar to G_0 daughter cells.

4.3 Repression is blocked by inhibitors of DNA synthesis,

but not by inhibitors of mitosis.

The experiment described in the previous section shows that SIR repression cannot occur in cells blocked in G_1 by alpha-factor. It must therefore occur at some later stage in the cell cycle. Figure 4.3a shows that when cells are released from the alpha-factor block, repression subsequently occurs, showing that the inhibition of repression is reversible. The kinetics of repression during this release are similar to those of bud emergence, suggesting that repression occurs around S phase.

In order to determine more precisely upon what cell cycle events repression is dependent, I have tested whether the repression which occurs upon release is blocked either by an inhibitor of DNA synthesis (hydroxyurea, HU (Hartwell, 1976)) or by an inhibitor of mitosis (MBC (Quinlan *et al.*, 1980 and Kilmartin, 1981). Figure 4.3b shows that HU blocks repression while MBC does not, suggesting that repression is dependent on some cell cycle event after the HU arrest point, but before the MBC arrest point. This result could, however, merely reflect the failure of MBC to block cell cycle progress. In order to demonstrate that the blocks were effective, the cells were examined microscopically. Cells blocked at mitosis or at DNA replication arrest as cells with one and

FIGURE 4.3a a1 Transcription in a sir3^{ts} strain after release from the alpha factor block.

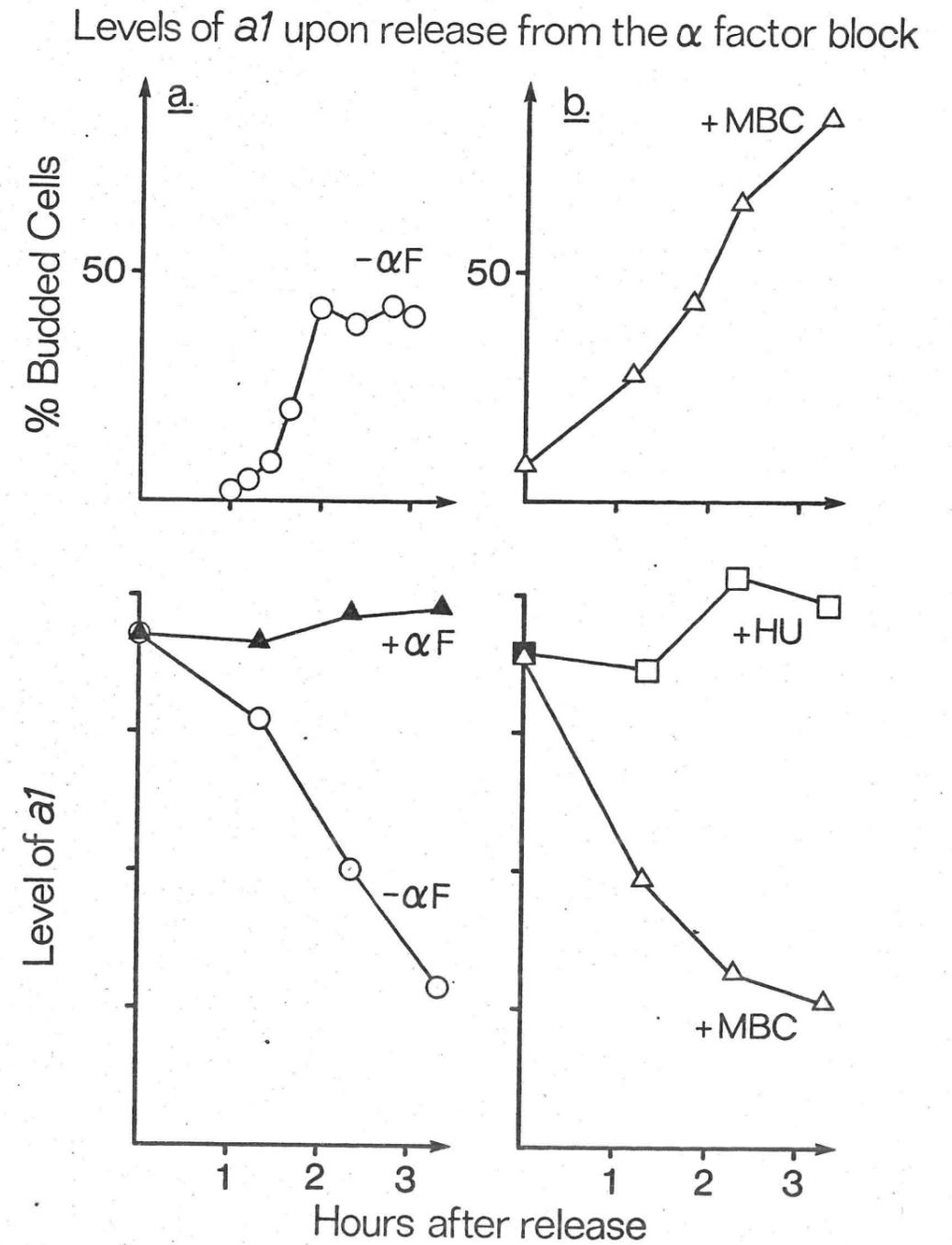
Purified G₀ daughters of strain M28 were shifted to 23°C and grown out in YEPD for 6 hours in the presence of alpha factor, then inoculated into fresh YEPD in the presence or absence of alpha factor. The levels of a1 RNA are plotted against time after release.

The proportion of budded cells in the release condition is also shown.

FIGURE 4.3b The effect of Hydroxyurea or MBC in a sir3^{ts} strain

The cells were released from alpha factor as described in Figure 4.3a, and inoculated into YEPD containing HU or MBC. The levels of a1 RNA are plotted against time after release.

The proportion of budded cells in the MBC condition is also shown. Cells grown into G₀ at the permissive temperature and subjected to identical treatments did not show any significant silent copy transcription.



only one bud. A time lapse microscopic analysis showed that less than 3% of the cells produced a second bud in either the HU or the MBC block. The data points shown in Figure 4.3 are typical results from a series of experiments, all of which gave similar results. In one experiment, the levels of HO and Histone H2B mRNA's were also measured, in order to confirm the effectiveness of the cell cycle inhibition. Since daughter cells were used, there will be no transcription of the HO gene in the first cell cycle, but if cells enter G₁ of the second cell cycle HO transcription will begin (Nasmyth, 1983; 1985b). H2B is transcribed in late G₁ in all cells (Hereford *et al.*, 1974). Thus cells in alpha factor should produce neither transcript, while those released into no cell cycle block at all should show considerable transcription of both HO and H2B. Cells released into HU or MBC should show H2B transcription, but negligible HO transcription. This is exactly what was observed (Figure 4.4), demonstrating that both these blocks are effective. Figure 4.5 shows the levels of SIR3 and a1 mRNA's in the RNA samples used to measure HO and H2B levels.

The procedure of blocking cells with alpha factor and then releasing them under different conditions minimises the period in which the cells are exposed to HU or to MBC, thus reducing the effect of growth inhibition by these reagents. It also ensures that the cells do not become habituated to the poison and thereby escape from its effect. The procedure therefore ensures that the cell cycle block is tight and minimises its intrusiveness.

In summary, I conclude that repression is occurring at a stage after the HU block but before the MBC block, suggesting that DNA replication itself is required.

When this experiment is repeated using a sir4^{ts} allele (strain

FIGURE 4.4 The effect of the cell cycle blocks on HO and Histone RNA levels.

The levels of HO and Histone RNA in some of the RNA samples from the experiment described in Figure 4.3. Each lane is labelled according to the cell cycle block used, and the time elapsed (in minutes) since the release from alpha factor.

Lane 1 Cells immediately after release.

Lane 2 Cells released into alpha-factor.

Lanes 3-5 Cells released into Hydroxyurea.

Lane 6 Cells in exponential growth in YEPD.

Lanes 7-10 Cells released into MBC

Lanes 11-14 Cells released into no cell cycle block at all.

The cells released into HU show reduced steady state levels of H2B transcript compared with the cells released into MBC or into no cell cycle block at all. This is probably caused by the increased turnover rate of Histone mRNA's that has been observed in cells treated with hydroxyurea (M.A.Osley and L.Hereford, personal communication.)

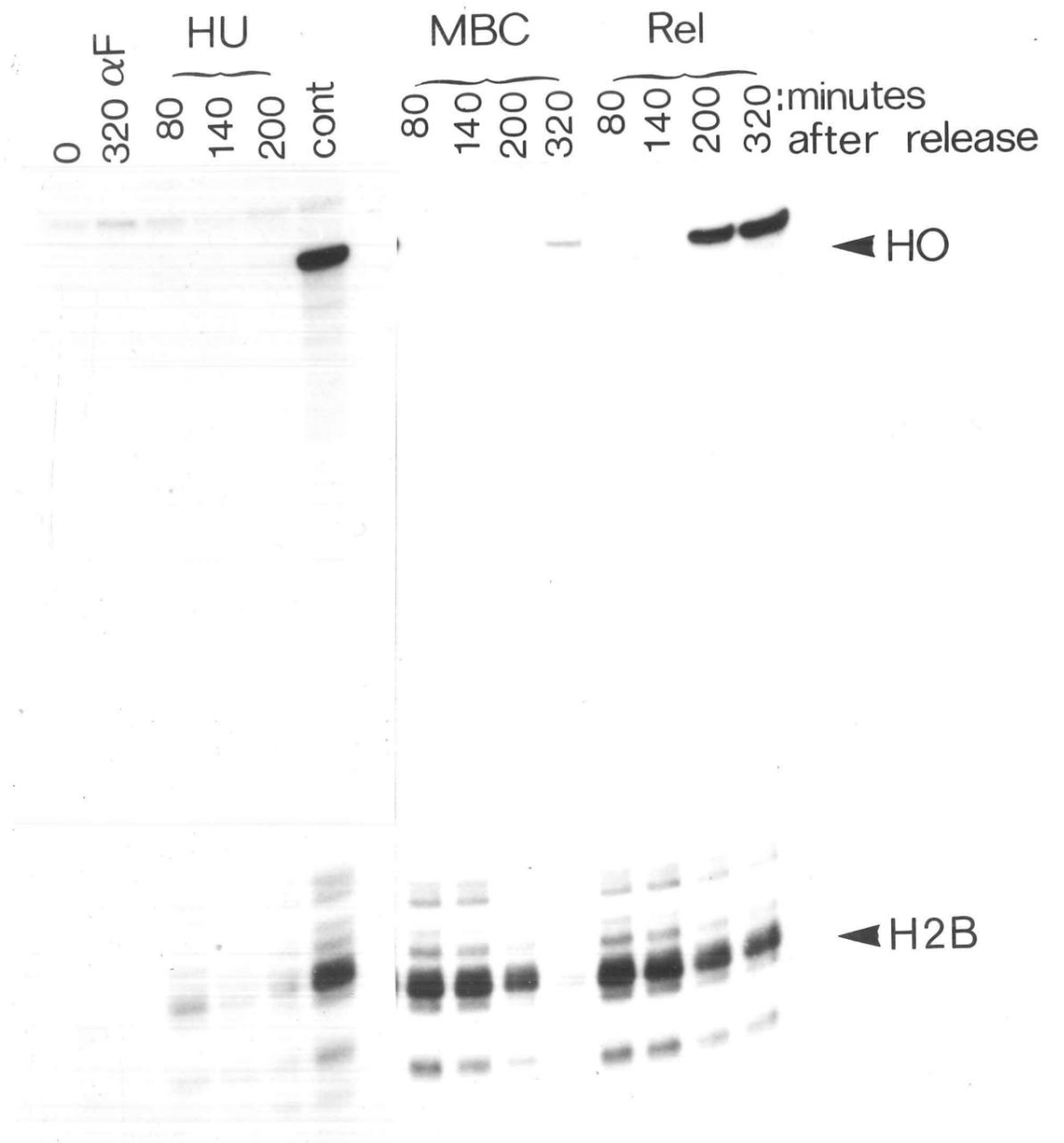


FIGURE 4.5 A typical S1 protection experiment

Daughters were prepared from cells grown into G_0 at either 23°C (lanes 1-9) or at 34°C (lanes 10-30.) They were inoculated into fresh YEPD and RNA samples taken (lanes 1 and 11), after one hour alpha factor was added to a final concentration of 3 units/ml and RNA samples were taken (lanes 2 and 12). After a further 4.5 hours the cells were harvested and released into alpha factor (αF), hydroxyurea (HU), MBC, or no cell cycle block at all (Rel), at this stage RNA samples were taken (lanes 3 and 13). Further RNA samples were taken at intervals, and the times (in minutes after release) are shown above the corresponding lanes. These treatments do not significantly affect a1 expression in the sir⁻ strain M29. As a positive control for the repression of the silent loci, daughters pregrown at 34°C were inoculated into YEPD and grown in the absence of any cell cycle block. RNA samples were taken after 5.5 hours (lane 10), i.e. when the cells in alpha factor were harvested, and after 11.6 hours (lane 30), i.e. shortly after the "320 minutes after release" timepoint.

25 μg of total RNA were used per lane.

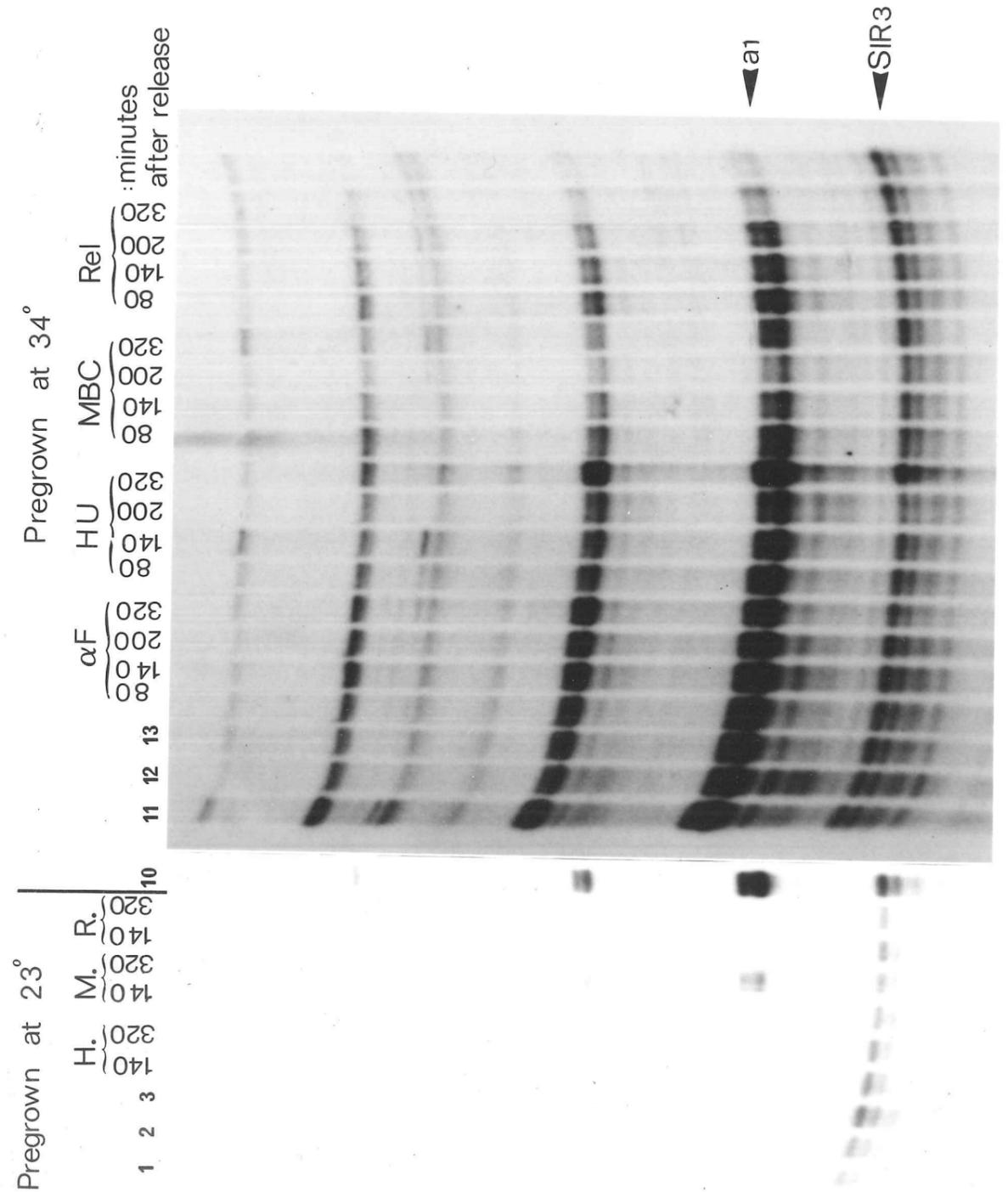


FIGURE 4.6a a1 Transcription in a sir4^{ts} strain after release from the alpha factor block.

Purified G₀ daughters of strain M37 were shifted to 23°C and grown out in YEPD for 6 hours in the presence of alpha factor, then inoculated into fresh YEPD in the presence or absence of alpha factor. The levels of a1 RNA are plotted against time after release.

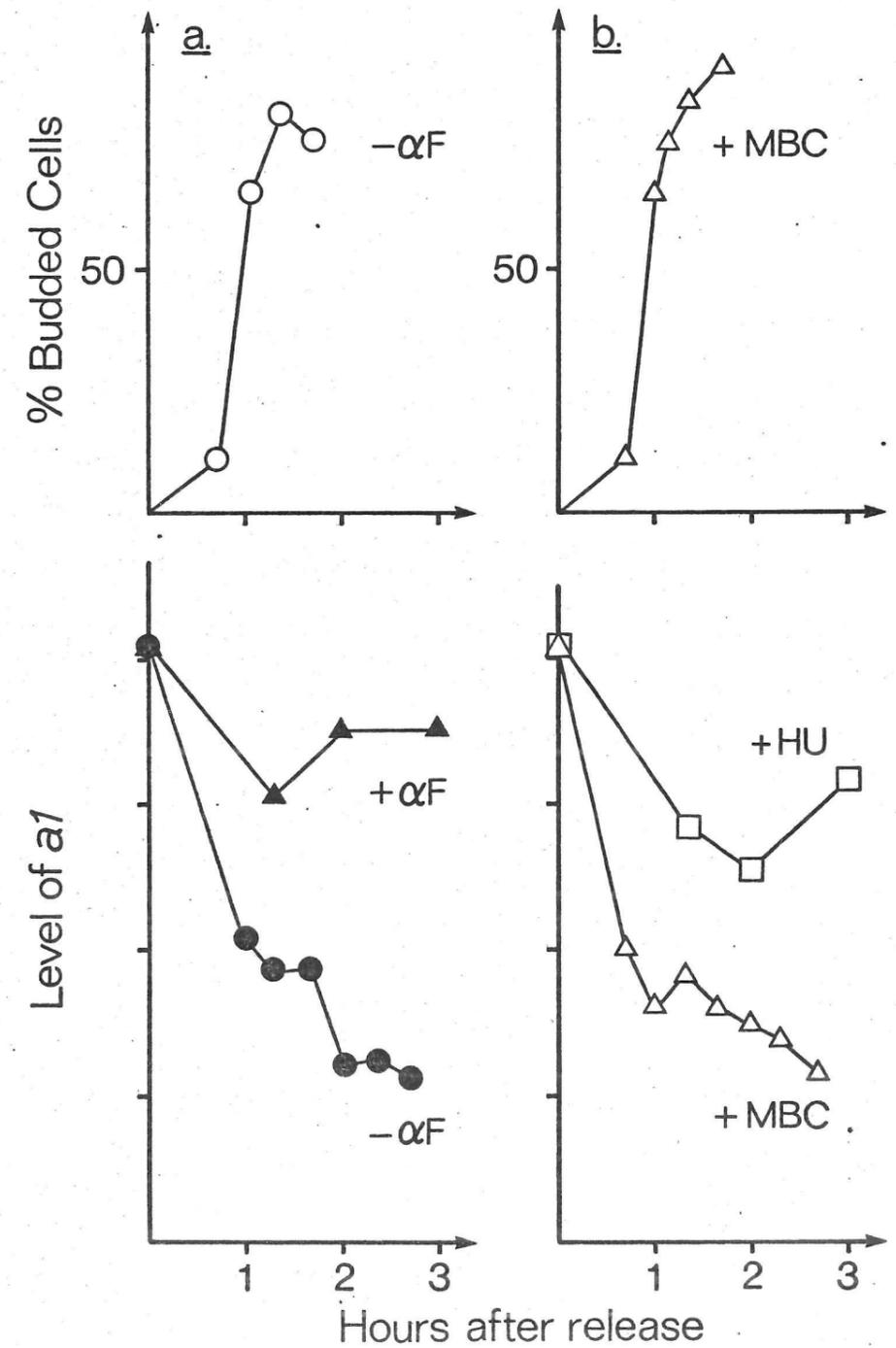
The proportion of budded cells in the release condition is also shown.

FIGURE 4.6b The effect of Hydroxyurea or MBC in a sir4^{ts} strain.

The cells were released from alpha factor as described in Figure 4.5a, and inoculated into YEPD containing HU or MBC. The levels of a1 RNA are plotted against time after release.

The proportion of budded cells in the MBC condition is also shown.

Levels of a1 upon release from the α factor block sir4^{ts}



M37, Table 2.1), the same result is observed (Figure 4.6), implying that the dependence on cell cycle progress for repression is not a peculiarity of the sir3^{ts} allele, but rather reflects a property of both the SIR3 and SIR4 gene products.

4.4 Cell Cycle Progress is not required for derepression.

In the previous section I described experiments which implied that SIR-mediated repression could only be re-established during S phase. One can also ask whether SIR⁺ activity is required throughout the cell cycle to maintain repression, or whether SIR⁺ activity is only required during S phase. sir3^{ts} cells were grown into G₀ at the permissive temperature and inoculated into fresh medium at the restrictive temperature. In this case cell cycle progress can be blocked either by alpha-factor or by a ts allele of CDC28 (CDC28 is required for progress through G₁. (Hereford and Hartwell, 1974))

Figure 4.7a shows that transcription from HMRa is induced within one to two hours of inoculation at the restrictive temperature, irrespective of whether the cells were blocked in G₁ (M20 sir3^{ts} cdc28^{ts}) or not (M10 sir3^{ts} CDC⁺). In fact the cdc28^{ts} strain M20 appears to derepress HMRa rather faster than the CDC⁺ strain M10. This difference in the time course of derepression seems to be due to background variation between nonisogenic strains, since M9 (sir3^{ts} cdc28^{ts}) gives a time course identical to M10. Furthermore, when strain M28 is subjected to the same temperature shift in the presence or absence of alpha factor, both conditions lead to full derepression within 30 minutes (data not shown). I conclude that cell cycle progress is not necessary for derepression and that SIR3 activity is required in early G₁ to maintain repression. SIR3's function, therefore, is not restricted to S phase.

FIGURE 4.7a The derepression of HMRa without DNA replication.

Levels of a1 RNA in a MAT α sir^{ts} strain after a shift to the restrictive temperature. Strains M20, M10 and M26 were grown into G₀ at the permissive temperature (23°C) on YEPD plates and inoculated into fresh YEPD medium at the restrictive temperature (37°C).

Lanes 1-5 M20(sir3^{ts} cdc28^{ts}): after 0, 0.5, 1, 2, and 3 hours at the restrictive temperature, lane 6 M20 grown out at the permissive temperature.

Lanes 7-11 M10(sir3^{ts} CDC⁺): after 0, 0.5, 1, 2, and 3 hours at the restrictive temperature, lane 12 M10 grown out at the permissive temperature. Lanes 13-17 M26(SIR⁺ cdc28^{ts}): after 0, 0.5, 1, 2, and 3 hours at the restrictive temperature, lane 18 M26 grown out at the permissive temperature.

25 μ g total RNA were used per reaction.

The difference in time course of derepression between M20 and M10 is due to background variation between nonisogenic strains: M9 (Table 1) gives a time course identical to M10, and when M28 is subjected to the same temperature shift in the presence or absence of alpha-factor, both conditions give full derepression after only 30 minutes (data not shown).

The effectiveness of the cdc block was confirmed by time lapse microscopic analysis; in the cdc28^{ts} strains, less than 6% of the cells produced buds.

FIGURE 4.7b The repression of MAT α 1 without DNA replication.

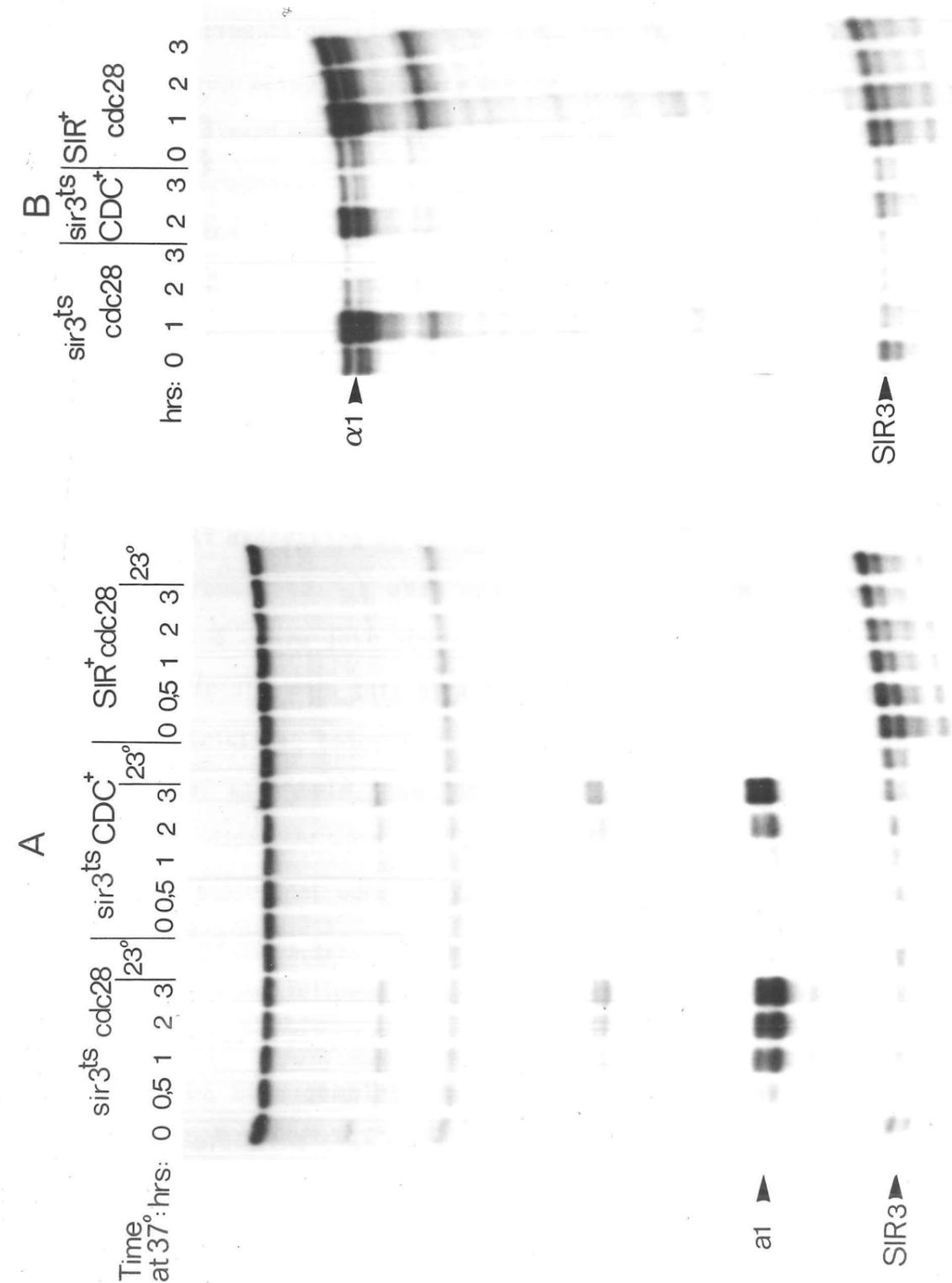
Levels of a1 RNA in a MAT α sir^{ts} strain after a shift to the restrictive temperature. Some of the RNA samples from the experiment described in Figure 4.6a were tested for levels of a1 RNA by S1 protection.

Lanes 1-4, M20 after 0, 1, 2, 3 hours at the restrictive temperature.

Lanes 5 and 6, M10 after 2 and 3 hours at the restrictive temperature.

Lanes 7-10, M26 after 0, 1, 2, 3, hours at the restrictive temperature.

25 μ g total RNA were used per reaction.



4.5 DNA replication is not required to repress MAT α 1

The experiments described above imply that the silent mating type loci can be repressed by SIR only during passage through S phase, or some closely linked event, whereas they may become derepressed without cell cycle progress. It is possible that an S phase requirement for repression is a general phenomenon and not peculiar to the mode of action of SIR regulation. Do there, therefore, exist genes which can be repressed in the absence of DNA replication?

This is simply tested. The transcription of MAT α 1 is repressed in diploids by the concerted action of the MAT α 2 and MAT α 1 gene products (Nasmyth *et al.*, 1981). α 1 transcription is also repressed in a sir⁻ cell expressing α 2 and α 1. A MAT α sir^{ts} strain (e.g. M20) therefore transcribes α 1 when grown into G₀ at the permissive temperature. I have just shown that α 1 transcription from HMR α commences rapidly when this strain is inoculated into fresh medium at the restrictive temperature (even if cells are blocked in G₁). Repression of α 1 should then occur. To determine whether such repression is dependent upon cell cycle progress, RNA samples from the previous experiment were analysed with an α 1 DNA probe (Figure 4.7b). Both M10(CDC⁺) and M20 (cdc28^{ts}) show a transient increase in α 1 transcription followed by a rapid reduction upon inoculation at the restrictive temperature. In fact repression appears to occur faster and be more complete in the cdc^{ts} strain. This probably reflects the different time courses with which α 1 transcript appears in these two strains (Figure 4.7a).

I conclude that α 1 transcription can be repressed by α 1/ α 2 control in cdc28-arrested cells. A requirement for cell cycle progress for repression is not, therefore, a general phenomenon.

4.6 Discussion

The experiments described in this chapter imply that if HML or HMR is derepressed by growing a sir3^{ts} or a sir4^{ts} strain at its restrictive temperature, then the ensuing repression during growth at the permissive temperature can only occur upon passage through S phase (Figures 4.1 - 4.6). This conclusion is based on the observation that repression can occur in the presence of MBC, which prevents mitosis (but not DNA replication), but not in the presence of alpha factor or hydroxyurea, which prevent entry into S phase. Unfortunately, the use of a temperature-sensitive sir⁻ allele to derepress HML and HMR has precluded the use of temperature-sensitive DNA replication mutants to pursue this question further.

The requirement for S phase is significant, however, because the cis-acting control sequences (HMLE and HMRE), via which SIR represses HML and HMR, have ARS activity (Abraham *et al.*, 1984; Feldman *et al.*, 1984). (ARS sequences confer the ability to replicate autonomously and are believed to be specific origins of DNA replication (Stinchcomb *et al.*, 1979; Beech *et al.*, 1980; Chan and Tye, 1980; Newlon and Burke, 1980; Celniker and Campbell, 1982).) More recent and more detailed deletion analyses by Brand (1986) have shown that HMRE activity is intimately associated with ARS activity. HMRE can be divided into three elements, called A, E, and B; the deletion of any two of these elements leads to expression of HMR. Both elements A and B are ARS sequences (Brand, 1986).

It seems likely that DNA replication is the critical event for SIR repression, although we cannot at present exclude the involvement of some other event in S phase. In the light of the results described in this chapter and the deletion analyses of HMLE

and HMRE, it has been proposed that DNA replication initiated at E is essential for repression.

The mechanism of SIR control can be formally divided into three steps. First, a protein involved in SIR repression recognizes and, presumably, binds at E. Secondly this signal is translocated to the target promoters, and thirdly the transcription of these promoters is repressed. In previously characterised systems of negative control (e.g. lac or bacteriophage lambda, (Gilbert and Muller-Hill, 1970; Miller and Reznikoff, 1978; Johnson et al. 1981)) these three steps, recognition, translocation, and repression, are coincident; that is, the binding of the repressor to the operator seems to be sufficient for repression, and there is no need to translocate any signal. SIR control appears different since the site of repression may be many hundreds of base pairs away from the recognition site (E), just as the recognition site for N anti-termination is distinct from the actual site at which anti-termination occurs (Salstrom and Szybalski, 1978; de Crombrughe et al., 1979).

DNA replication may be required for the recognition of E, or for the act of repression, or for the communication between the two. For example the functional SIR3 or SIR4 protein produced during growth at the permissive temperature may be unable to bind to E unless it is opened up at replication. This assembly would also be impossible if the E site were occupied by a non-functional SIR complex that could only be disassembled at replication. Alternatively, DNA replication may only be required for the act of repression: for instance, the HMLa1 and HMRa1 promoters may bind a transcription factor which cannot be removed from the DNA except at replication. Thus, once the silent copies have been derepressed,

this factor is bound to the promoter and the SIR⁺ or sir⁻ status of the cell is irrelevant until DNA replication occurs. Similarly, repression may be achieved by the construction of a repressed chromatin structure which can only be produced as the nascent DNA is packaged in S phase. Since derepression can occur without progress through the cell cycle (See Figure 4.6a), however, this repressed state must be dependent on the continued action of the SIR3 gene product. The latter point would seem to exclude replication- dependent methylation as the mode of SIR repression. Another possibility is that the translocation of the signal for repression is replication-dependent, as would be the case if translocation were achieved by the movement of a replication fork initiated at E to the affected promoters.

Whichever of these different hypotheses is correct, these results suggest that some modulator of transcription is unable either to exchange onto, or off, the DNA in the absence of DNA replication.

Gene expression is linked to DNA replication in both procaryotic and eukaryotic viruses. The early-late switch in transcription accompanies, and appears to be dependent upon, viral DNA replication. There are often important biological reasons for this pre-replication/ post-replication distinction in viral transcription. In such cases, the apparent "requirement" for DNA replication may be indirect, and have little or nothing to do with the actual mechanism of transcriptional control. For example, it has been proposed that DNA replication is required for late gene expression in SV40 (Contreras et al., 1982). More recent work, however, has demonstrated that the late promoter can be activated by T antigen in the absence of DNA replication (Keller and Alwine, 1984). In a bacteriophage T4 infection, late transcription requires

some form of template activation, but ongoing DNA replication appears to be just one way of achieving this (Geiduschek et al., 1980 and Elliott and Geiduschek, 1984). In Adenovirus, on the other hand, it has been demonstrated that a replicated DNA template is transcribed differently from an unreplicated DNA template within the same cell (Thomas and Mathews, 1980), suggesting that, in this virus, DNA replication may be directly involved in the mechanism of the early-late switch.

In viral systems, the coordination of gene expression and DNA replication is part of a developmental pathway. Because this coordination is often adaptive, various viruses have evolved a connection between DNA replication and gene expression which may or may not be a direct one. In contrast SIR repression is normally continuous throughout the cell cycle of both haploid and diploid yeast cells. Therefore the observed requirement for DNA replication more likely reflects the actual mechanics of SIR repression.

We have also demonstrated here that another system of repression, that of MAT α 1 by a1/ α 2, does not, in contrast to SIR, involve DNA replication (Figure 4.6b). The cis-acting DNA sequence necessary for this repression resides within a few base pairs of the MAT promoter (Siliciano and Tatchell, 1984; Chapter 5) whereas those for SIR lie a kilobase or more away from the affected promoter. It is possible that these represent two distinct modes of repression in eukaryotes, one involving DNA replication and capable of acting at a distance, and the other not involving DNA replication and only acting to full effect when interposed between the upstream activating sequence and the transcription start site.

A replication-dependent mechanism of transcriptional control could be important in the maintenance of developmental switches in

some form of template activation, but ongoing DNA replication appears to be just one way of achieving this (Geiduschek et al., 1980 and Elliott and Geiduschek, 1984). In Adenovirus, on the other hand, it has been demonstrated that a replicated DNA template is transcribed differently from an unreplicated DNA template within the same cell (Thomas and Mathews, 1980), suggesting that, in this virus, DNA replication may be directly involved in the mechanism of the early-late switch.

In viral systems, the coordination of gene expression and DNA replication is part of a developmental pathway. Because this coordination is often adaptive, various viruses have evolved a connection between DNA replication and gene expression which may or may not be a direct one. In contrast SIR repression is normally continuous throughout the cell cycle of both haploid and diploid yeast cells. Therefore the observed requirement for DNA replication more likely reflects the actual mechanics of SIR repression.

We have also demonstrated here that another system of repression, that of MAT α 1 by a1/ α 2, does not, in contrast to SIR, involve DNA replication (Figure 4.6b). The cis-acting DNA sequence necessary for this repression resides within a few base pairs of the MAT promoter (Siliciano and Tatchell, 1984; Chapter 5) whereas those for SIR lie a kilobase or more away from the affected promoter. It is possible that these represent two distinct modes of repression in eukaryotes, one involving DNA replication and capable of acting at a distance, and the other not involving DNA replication and only acting to full effect when interposed between the upstream activating sequence and the transcription start site.

A replication-dependent mechanism of transcriptional control could be important in the maintenance of developmental switches in

Brown (1984)

gene expression. Whereas the signals and circumstances inducing cell differentiation are often transient, the subsequent patterns of gene expression in the terminally differentiated cell are often permanent. One mechanism by which this can be achieved is the existence of positive feedback loops, as in bacteriophage lambda (Johnson et al, 1981). An alternative mechanism could involve the final division of a differentiating cell leading to a state of gene expression which cannot be altered in the absence of further DNA replication. For instance, the results described in this chapter suggest that yeast cells will continue to transcribe the HMRa1 gene despite the presence of SIR repressors as long as the cells are prevented from undergoing DNA replication.

CHAPTER FIVE

IDENTIFICATION AND COMPARISON OF TWO CELL-TYPE SPECIFIC
OPERATOR SEQUENCES

5.1 Introduction

5.1.1 Background

The previous chapters considered SIR-mediated repression which, unlike other characterised systems of repression, is capable of acting at a distance. I concluded that the establishment of the repressed state required passage through S phase. In chapter 4 I also described an experiment showing the establishment of a1/α2-mediated repression in the absence of DNA replication. This chapter describes experiments which study two systems of negative control involved in determining cell type - the a1/α2 system and the α2 system.

Strathern et al (1981) proposed that the MATα2 protein is a negative regulator of some of the genes that are required for a mating, and that in a/α diploids, MATa1 and MATα2 act together as a negative regulator of genes that are not required in diploids, including MATα1 (Figure 1.2). (For a recent review, see Klar et al, 1984) In a number of cases this control appears to be exerted at the level of transcription - for example the α2 protein acts as a repressor of transcription of the a-specific gene STE6 (Wilson and Herskowitz, 1984) and is believed to act similarly on other genes involved in producing the a mating phenotype, such as STE2, BAR1 and MFa1 (which encodes a factor) (Nakayama et al, 1985; V.Mackay personal communication; Brake et al, 1985)(See Table 1.1). Once an α cell has mated to produce an a/α diploid, the function of α2 becomes

modified. In addition to its function in haploids, it now acts in conjunction with the MATa1 gene product to repress another set of genes. Transcription of these genes is repressed only if both MATa1 and MAT α 2 are expressed. The genes that are under this a1/ α 2 repression are the "haploid specific" genes, whose function is not required in diploids. Such genes include the MAT α 1 gene (Nasmyth et al, 1981a, Klar et al, 1981a) (which activates α -specific genes), the H0 gene (Jensen et al, 1983) (which initiates mating type switching) and STE5 (which is required for mating in both a and α cells) (J.Thorner personal communication and V.Mackay personal communication.) Thus the α 2 gene product represses one set of genes in α cells, but in a/ α diploids it represses an additional set of genes.

Shepherd et al (1984) discovered that both MATa1 and MAT α 2 show protein sequence homology with the " homeo domain". The " homeo domain" is a highly conserved protein sequence found in gene products involved in selecting developmental fates in Drosophila melanogaster and it has also been located in genes of unknown function in other phyla, including vertebrates (McGinnis et al, 1984a,b; Scott and Weiner, 1984; Carrasco et al, 1984; Poole et al, 1985; Wharton et al, 1985; Kuroiwa et al, 1985). Antibodies raised against different gene products that contain the homeo domain recognize nuclear proteins (White and Wilcox, 1984; Beachy et al, 1985; Carroll and Scott, 1985; DiNardo et al, 1985). In the case of MATa1, part of the homeo domain is strikingly similar to the DNA binding domain of bacterial repressors (Sauer et al, 1982), suggesting that MATa1 encodes a repressor protein. There is more direct evidence that the MAT α 2 gene product is a repressor - A.Johnson has recently shown that an α 2- β galactosidase fusion protein binds specifically to the 5' flanking DNA of STE2 (Johnson and Herskowitz, 1985)

In this chapter, I describe work which identifies the DNA sequences which are recognized by the $\alpha 2$ and the $a1/\alpha 2$ repressors first by finding candidates on the basis of DNA sequence comparisons and second by demonstrating that restriction fragments containing the proposed sequences are sufficient to confer the appropriate control on an unrelated promoter. The proposed $\alpha 2$ and $a1/\alpha 2$ control sequences are different but related, and I discuss the means by which the two sets of DNA targets ($\alpha 2$ versus $a1/\alpha 2$) may be distinguished.

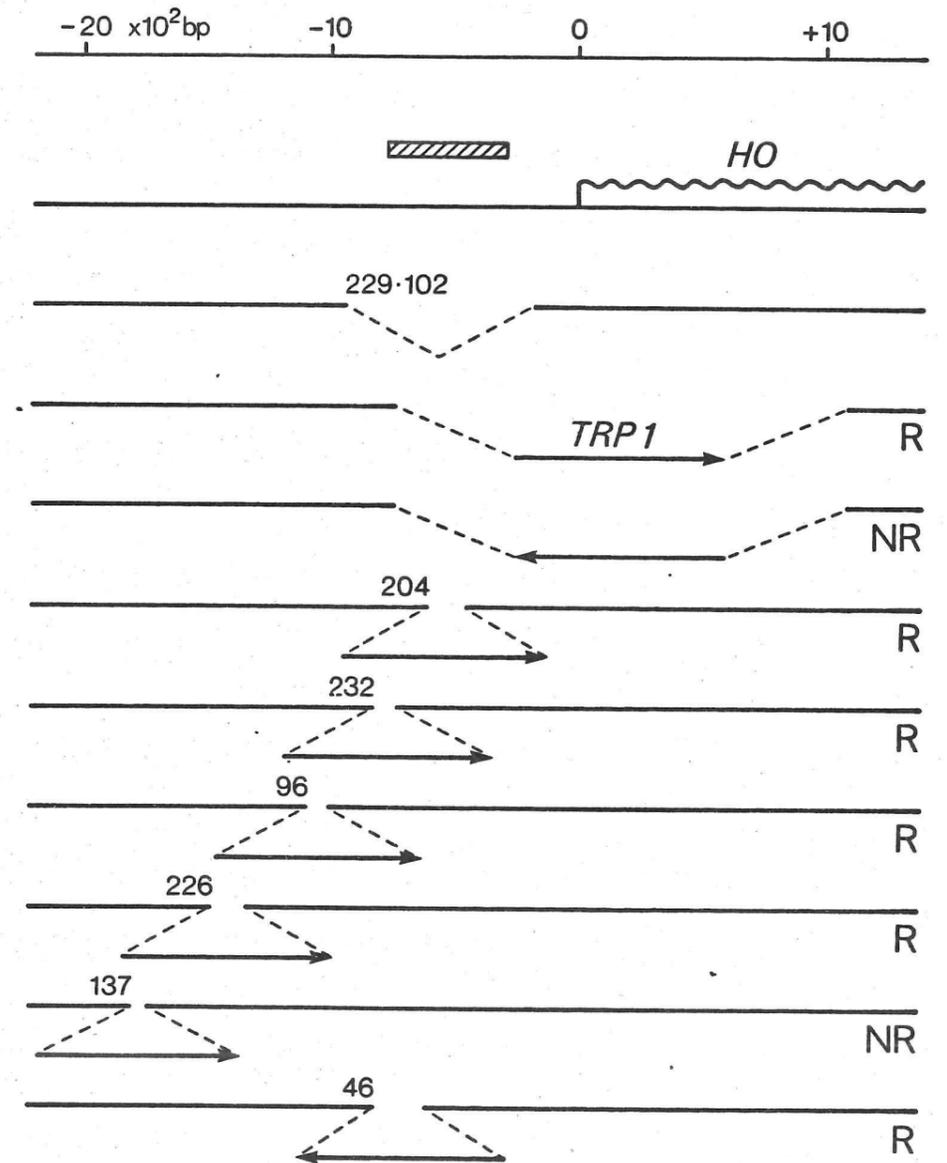
5.1.2 $a1/\alpha 2$ Control Of The $H0$ Gene

One of the haploid-specific genes which is repressed by the action of $a1$ and $\alpha 2$ is the $H0$ gene (Jensen *et al.*, 1983), which is responsible for initiating mating type switching (Kostriken *et al.*, 1983, and F. Heffron personal communication). Two lines of evidence suggest that the sequences which confer $a1/\alpha 2$ repression on $H0$ transcription are present in multiple copies widely distributed within the 5' flanking DNA. The first comes from the deletion analysis by Jensen (1983) of plasmids carrying an $H0-lacZ$ fusion. A deletion removing all but 296bp of the DNA flanking the 5' end of the gene (as measured from the AUG) is fully constitutive, whereas one removing all but 762bp is fully repressed by the a/α state. This suggests that $a1/\alpha 2$ control sequences are 5' to the gene and that they must lie between -296 and -762 from the AUG (Figure 5.1). Surprisingly, a deletion (No. 229.102), which removes all the DNA between -172 and -928 is not only transcribed efficiently, but also is fully repressed by the a/α state when transplanted at the $H0$ locus (Nasmyth, 1985a). This suggests that there must be at least two copies of the putative $a1/\alpha 2$ control element, one between -296 and -762, and another either upstream of -928 or else in the 3' flanking sequences.

FIGURE 5.1

α/α control of the HO gene is redundant

The hatched box denotes the region defined as important by Jensen (1983) and the wavy line represents the HO transcript. The extent of the deletion 229.102 is shown. Insertions of the TRP1 gene into various deletions are also shown; R indicates that, in a particular insertion, the TRP1 gene is repressed in α/α cells, NR indicates that the TRP1 gene is expressed constitutively. The arrow on the TRP1 gene indicates the direction of transcription. NOTE: These constructions were made and analysed by K.Nasmyth, and not the author.



The distribution of the a1/α2 control elements has also been studied by insertions of the yeast TRP1 gene at various positions in the H0 locus. The 850bp EcoRI-BglIII fragment of TRP1 contains all of the coding sequences of the gene encoding PR-anthranilate isomerase but few of its 5' flanking sequences (only out to -102 from the AUG). Due to its incomplete promoter, the level of expression of this fragment is prone to position effects. It has been used, therefore, as a probe for the regulatory properties of the DNA flanking H0 by analysing the Trp⁺ phenotype in a, α or a/α trp⁻ cells containing various insertions of TRP1 in H0.

A variety of TRP1/H0 insertions (See Figure 5.1) have been transferred to the chromosomal H0 locus using the SUP4-o transplacement technique (Rothstein, 1983; Nasmyth, 1985a). All of the constructions shown in Figure 5.1 lead to a fully Trp⁺ phenotype in the parent a strain, K765 (Table 2.1). Many of the insertions, however, lose the ability to confer tryptophan prototrophy when an a/α diploid is formed by mating to strain K822. This specific loss of tryptophan prototrophy has been used as a qualitative measure of a/α repression. For example, the TRP1 gene is fully repressed by the a/α state when inserted (in the same orientation as the H0 coding sequence) as a replacement for the H0 DNA from -718 to +1096, but it is not repressed in a/a or α/α diploids. In contrast, the gene is constitutive (i.e. can confer a Trp⁺ phenotype to a, α, and a/α cells) when inserted in the opposite orientation. The simplest interpretation of this result is that H0 sequences to the left of position -718 can confer a/α repression on the TRP1 gene and, moreover, they will only do so if joined to its 5' end. This result contrasts with the properties of the E element involved in SIR-mediated repression, which will act when 3' to a gene. The TRP1 gene is also regulated when

placed (in the same orientation as the HO coding sequence) in a number of positions within the interval -300 to -1450 (See figure 5.1). On the other hand, it is not repressed (at least qualitatively) in deletion no.137 (-1737 to -1777).

This analysis of the HO gene (which was performed by Kim Nasmyth) taken together with the results of Jensen (1983), suggest that there may be several a/α control elements in the interval between -400 and -1400, and that the 5' end of the TRP1 gene must be near one such element in order to be regulated. Alternatively, there may be fewer elements (though there must be at least two) which are effective over longer distances. It is interesting that an insertion of TRP1 in the opposite orientation (i.e. with its 5' end proximal to the HO coding region) within deletion no. 46(-582 to -791) is also under a/α control. This is consistent with the data of Jensen showing that a/α control elements exist within the -296 to -762 interval and suggests that there may be one such element to the right of position -582.

In summary, the a/α control of the HO gene must be redundant - containing two or more control sites within the region -1777 to -292.

5.2 The Search For A Repeated Sequence

If the a/α control of the HO gene is redundant, and if a single DNA sequence motif is responsible for the control, then such a sequence should be repeated a number of times within the HO promoter, and furthermore the positions at which the motif appears should be consistent with the analyses described above. The HO regulatory region was therefore searched by computer to see if it contained a repeated sequence motif that satisfied the above criteria.

The 5' flanking sequence of the HO gene (length 1950 bp) was

compared with itself, and with its complementary strand, using DIAGON (Staden, 1982). This revealed a number of examples of a possible a1/α2 element, with the sequence (T/C)C(A/G)TGTNN(A/T)NANNTACATCA . If this sequence really is involved in a1/α2 control, then we would expect to find it at other genes that are similarly controlled, such as MATα1. The H0 sequence was therefore DIAGONed against the MATα intergenic region, revealing the same sequence with the MATα promoter region. These examples were aligned to produce a frequency matrix using GETFRQ. (This process produces a matrix in which the rows represent the four nucleotides, T,C,G, and A, while the columns represent the sequential residue positions in the sequence motif. The numbers within the matrix represent the frequencies at which each nucleotide T,C,G or A appears at each position in the sequence motif.) This matrix was used to search for further examples of the sequence motif within the H0 gene and also within the STE5 5' flanking sequence. This revealed a total of ten examples of the motif at the H0 gene, and two examples within 705 bp of STE5 5' flanking DNA.

Thus we have in total 13 examples of the sequence (Table 5.1). 10 of them are found widely dispersed throughout the H0 5' flanking sequence, with a distribution that is consistent with the analyses described in the previous section. At MATα we find one occurrence of this sequence (one "a1/α2 element"), which lies in the region defined as necessary for a1/α2 control (Siliciano and Tatchell, 1984). At STE5, we find two such "a1/α2 elements" in tandem.

The computer search of the H0 gene also revealed another repeated sequence, CACGAAAA. However all occurrences of this sequence are contained within the endpoints of the 229.102 deletion, which still retains a1/α2 control. This CACGAAAA motif is involved in the cell-cycle-dependent transcription of the H0 gene

TABLE 5.1

Aligned Sequences of the $\alpha 1/\alpha 2$ Elements

Sequence	Source	"Score"
CAATGTAGAAAAGTACATCA	MAT α 1	-13.5
GCTTGTTAATTTACACATCA	STE5 -196	-15.0
TCATGTACTTTTCTGCATCA	STE5 -179	-13.0
CCGCGTTAAAACCTACATCA	HO -1752	-14.4
TTATGTTAAAAGTTACATCC	HO -1391	-13.2
GCCTGCGATGAGATACATCA	HO -1328	-17.2
TAGAGTGAAAAAGCACATCG	HO -1208*	-18.5
TCATGTATTCATTCACATCA	HO -736*	-10.6
ACATGTCTTCAACTGCATCA	HO -669*	-14.7
TCGTGTATTTACTTACATCA	HO -576*	-10.0
TCATGTTATTATTTACATCA	HO -411	-8.7
TCATGTCCACATTAACATCA	HO -371	-12.8
GCGTTTAGAACGCTTCATCA	HO -150*	-18.1

CONSENSUS: TC^A_GTGTNN^A_TNANNTACATCA

TABLE 5.1

Aligned Sequences of the $\alpha 1/\alpha 2$ Elements

The sequences of the $\alpha 1/\alpha 2$ elements that we have found in the $MAT\alpha$, $STE5$ and HO genes are shown. For the elements marked *, the orientation of the element has been reversed (taking the complementary strand) to give a closer fit to the consensus sequence. The "score" shown is a measure of each element's fit to the consensus sequence (Staden, 1984). The larger the negative number, the poorer the fit. Thus the best fit to the consensus sequence is the HO -411 element which was used in the sufficiency experiment (Insert B in Figure 5.2). An approximate consensus sequence is shown at the foot of the table. Note that this is an imperfect inverted repeat. Sequence information is from (Astell *et al.*, 1981), (J.Schultz and J.Thorner personal communication), and (M.Squire and M.Smith, personal communication) respectively.

The frequency matrix used to search the EMBL library was derived from these examples, except that the $MAT\alpha 1$ sequence was included three times, to give it greater weight. I have re-evaluated all the occurrences I have found, testing their score against a frequency matrix that lacks this arbitrary weighting. Thus all the "scores" mentioned in this dissertation refer to a frequency matrix derived from the set of examples shown here.

TABLE 5.1

Aligned Sequences of the $\alpha 1/\alpha 2$ Elements

Sequence	Source	"Score"
CAATGTAGAAAAGTACATCA	$MAT\alpha 1$	-13.5
GCTTGTTAATTTACACATCA	$STE5$ -196	-15.0
TCATGTACTTTTCTGCATCA	$STE5$ -179	-13.0
CCGCGTAAAAACCTACATCA	HO -1752	-14.4
TTATGTTAAAAGTTACATCC	HO -1391	-13.2
GCCTGCGATGAGATACATCA	HO -1328	-17.2
TAGAGTGAAAAAGCACATCG	HO -1208*	-18.5
TCATGTATTCATTCACATCA	HO -736*	-10.6
ACATGTCTTCAACTGCATCA	HO -669*	-14.7
TCGTGTATTTACTTACATCA	HO -576*	-10.0
TCATGTTATTATTTACATCA	HO -411	-8.7
TCATGTCCACATTAACATCA	HO -371	-12.8
GCGTTTAGAACGCTTCATCA	HO -150*	-18.1

CONSENSUS: $TC\overset{A}{G}T\overset{A}{N}N\overset{A}{T}NANNTACATCA$

(Nasmyth, 1985b).

5.3 The Repeated Sequence Is Sufficient To Confer a_1/a_2 Control

We have asked whether the sequence (T/C)C(A/G)TGTNN(A/T)NANNTACATCA is able to bring an unrelated promoter under cell type control. To do this we have used the CYC1 promoter fused to the E.coli β -galactosidase gene (Guarente and Ptashne, 1981). The CYC1 gene encodes cytochrome c and is regulated by the availability of heme, and by the carbon source, being partially repressed when the yeast are grown on media containing glucose. The promoter consists of a TATA box and another element about 300 bp upstream, which is required for efficient transcription, called the Upstream Activating Site (UAS). It has been shown that of these elements it is the UAS which is sensitive to the carbon source (Guarente et al, 1984). We have investigated the effect of inserting an a_1/a_2 element into the region between the UAS and the TATA box of the CYC1 promoter. To do this we took two different examples of the a_1/a_2 element from the HO gene (carried on fragments of length 47 and 50 base pairs) and inserted them into a unique XhoI site in the CYC1 promoter, and asked how this insertion affected the activity of the promoter in different cell types. We find that all three of these constructions produce similar levels of β -galactosidase activity in isogenic a , α , and a^- cells (when tested on indicator plates, data not shown).

If we take an α cell carrying the plasmid, we can either mate it to an a cell or an a^- cell to produce two different diploids. One, the a/α , is a normal diploid, and will repress its haploid-specific genes. The other, the a^-/α cell, will not repress its haploid specific genes, since MATa1 is absent. These strains are otherwise

isogenic. When the β galactosidase activity of these strains is measured, it is found that the parent plasmid is not affected significantly by the difference in mating type (Figure 5.2a). The plasmids which contain an $a1/\alpha2$ element, however, show a dramatic reduction in the level of β galactosidase activity in the a/α diploid, as compared with the a^-/α diploid. This implies that a CYC1 promoter containing an $a1/\alpha2$ element is sensitive to a/α control.

There are two ways of viewing this result. It could arise if the insertion of the $a1/\alpha2$ element inactivated the CYC1 promoter, and simultaneously generated a novel a/α -sensitive promoter to replace it. Alternatively, it could be that the $a1/\alpha2$ element is acting in a manner analogous to a prokaryotic operator sequence, bringing the CYC1 promoter under a/α control. In the latter case we would expect the promoter to show the carbon source dependence which is characteristic of the CYC1 promoter. Figure 5.2a shows that the parent plasmid and the two plasmids containing the $a1/\alpha2$ elements all show a similar, approximately three-fold, repression of β galactosidase activity when grown on glucose as carbon source. The insert-containing promoters are therefore still being driven by the CYC1 UAS, and the activity of this UAS is blocked in a/α cells by the presence of an $a1/\alpha2$ element within the promoter.

Figure 5.2 also shows that in α cells the insert-containing plasmids produce transcripts identical to those produced by the parent plasmid, but in a/α cells no transcript is seen. Therefore, the control is operating at the level of transcription.

5.4 The Search of other Genes.

Since the $a1/\alpha2$ element appears sufficient to confer cell-type-specific repression, we would not expect to find good matches to

FIGURE 5.2

The a1/α2 element is sufficient to confer a/α control

Two different examples of the a1/α2 element from the H0 gene were inserted between the UAS and the TATA box of the CYC1 promoter driving the βgalactosidase gene. These plasmids were used to transform an a/α diploid, and the isogenic a⁻/α diploid.

(A) Levels of βgalactosidase (βGal) produced in exponentially growing cells are shown, both for raffinose and for glucose as carbon source. Levels of RNA produced by the plasmids were also measured by the protection of radiolabelled DNA from S1 nuclease digestion.

(B) Sequences of the constructions with the a1/α2 elements underlined. Insert A is H0 -736; Insert B is H0 -411 (Table 5.1).

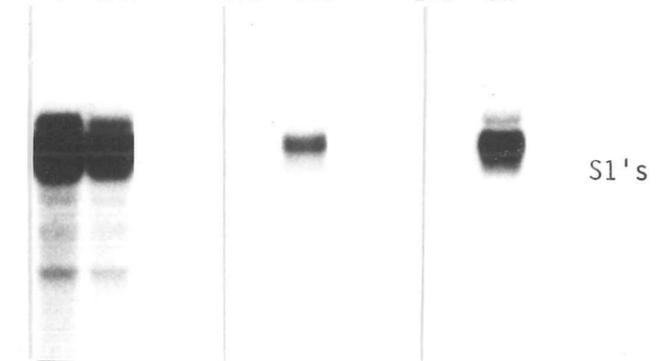
The asterisks denote the limits of the inserted fragments.

The MATα strain used was 822; this was mated with either K699 (MATa) or RS3 (mat::LEU2). See Table 2.1.

A

EFFECT OF THE a1/α2 ELEMENT ON THE CYC1 PROMOTER

PLASMID: MATa1	Parent		Insert A		Insert B		βGal Assays
	+	-	+	-	+	-	
Gluc	56	92	1.9	62	0.6	47	
Raff	250	240	4.3	190	1.5	81	



B

Parent

(← UAS) (TATA →)
 GCGTAGGCAATCCTCGAGGTCGACCTCGAGCAGATCCGCCAGGC
 **** *

Insert A

GGCAATCCTCGAGATCCACGAAAATGATGTGAATGAATACATGAAAGATTCATGAGATCTCGAGCAGATCCGCC
 **** *

Insert B

GGCAATCCTCGAGGTACGAAAATGATCCACGAAAATCATGTTATTATTACATCAACATATCGTCGAGCAGATCCGCC
 **** *

this sequence in genes that are not under a/α control. We have searched the EMBL nucleotide sequence library to see if there are other genes which contain this sequence motif. This was done using the ANALYSEQ program (Staden, 1984). A matrix was constructed which contains the frequencies at which each nucleotide T,C,G or A occurs at each position within the examples of the $a1/\alpha2$ element already identified. This matrix was then used to scan DNA sequences to find regions which show a good fit to the consensus sequence.

Some caution must be exercised in interpreting the results of such a search, since the "goodness of fit" observed refers to the consensus sequence which I have constructed and does not necessarily reflect exactly what is important for a/α repression in vivo.

I have searched 61kb of yeast DNA sequence in the EMBL nucleotide sequence library and have found only one significant match within the 5' flanking region of any gene. The $MF\alpha1$ gene (Singh et al, 1983) has the sequence CCATGTGAAAAGATGCATCT (which would score -15.5 in table 1) at -230 bp from the AUG. $MF\alpha1$ encodes the α cell's mating pheromone and is an α -specific gene, whose transcription is dependent on the $MAT\alpha1$ gene product (Strathern et al, 1981, see Figure 1.2). $MF\alpha1$ is not expressed in a/α diploids partly because $MAT\alpha1$ is absent. There must, however, be a second a/α dependent control since $MF\alpha1$ is still not expressed in a/α diploids when $MAT\alpha1$ is expressed constitutively (Siliciano and Tatchell, 1984; Ammerer et al, 1985). Our finding that $MF\alpha1$ contains possible $a1/\alpha2$ control element suggests that this second control may be exerted directly by the $a1/\alpha2$ repressor, but does not rule out the possibility that $a1/\alpha2$ may also repress the expression of a transcriptional activator other than $MAT\alpha1$, and that this activator is required for the expression of various genes involved in mating.

We also find a number of matches that occur within the open reading frames of genes which are not known to be regulated by mating type. In order to test whether such elements might have any effect on transcription, I have inserted into the CYC1-lacZ fusion a restriction fragment known to cause repression when inserted into the promoter.

Insert A (Figure 5.2) was inserted into the open reading frame of the CYC1-lacZ fusion, producing the construction A-Bam (Figure 5.3), and the transcription from this plasmid was measured in yeast strains K700 and Y65. The construction is expressed at the same level in both strains (Figure 5.3). Thus no detectable effect of an a1/α2 element within an open reading frame was observed. This implies that the examples of the a1/α2 element found within open reading frames are not, in themselves, going to affect the transcription of the gene. Therefore the examples that have been found in all probability represent random occurrences, being neither selected for, nor selected against, in the evolution of the open reading frame concerned. If this is the case, then these random matches give us an empirical measure of the significance of the matches shown in Table 5.1. The search of 61 kb of DNA sequence revealed 10 examples of putative a1/α2 elements that would score -16.2 or better in Table 5.1, suggesting that such good fits to the consensus sequence should only occur at random once every 6 kb.

5.5 Effect of an Element, when Placed outside of the Promoter

In the experiments described here, as in those of Brent and Ptashne (1984), the repressing element has been inserted between the UAS and the TATA box, although not overlapping either promoter element. What effect does the element have when placed upstream of the UAS: that is, upstream of all sequences required for transcription?

FIGURE 5.3

Effect of inserting the $a1/a2$ element in different places.

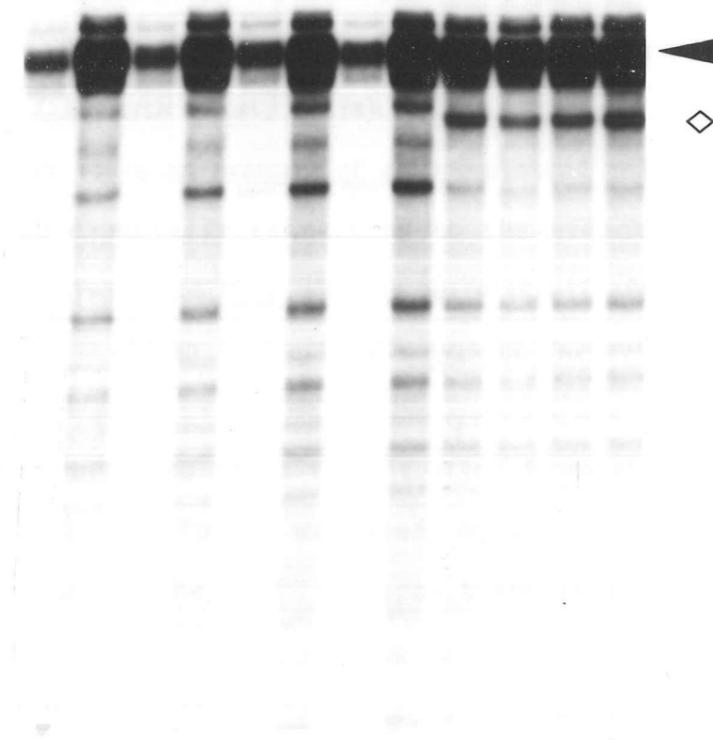
(A) Insert A (Figure 5.2) was inserted into the SmaI site (which is upstream of the UAS) and into the BamHI site (which is within the β -galactosidase transcription unit) of the plasmid pLG-312, and the levels of RNA produced in isogenic α and a/α cells were measured. Within each group of four lanes, the first and third show protection by RNA from a/α cells, while the second and fourth show protection by RNA from α cells. The arrow-head indicates the expected transcription start site, while the diamond indicates the BamHI site. Since the DNA probe is derived from the parent plasmid, the RNA produced from the A-Bam construction should contain a short insert at this point that is not homologous to the DNA probe. The faint signal indicated by the diamond presumably results from partial S1 digestion of the probe DNA where the non-homologous RNA is looped out of the hybrid.

The strains used were K700 (MAT α HMRa) and its derivative Y65 (MAT α HMRaE Δ , which expresses a1 (See Table 2.1.) RNA samples were made from cells in exponential growth on glucose as carbon source.

(B) Sequences of the constructions used. The asterisks denote the limits of the inserted sequences. Note that both the insertions into the SmaI site contain small deletions of vector sequence. See Figure 7.2 for a diagram depicting the different constructions.

a

A'-Sma | A-Sma | A-Bam



b

Parent
GCTCCCGGGTAATAACTGATAT
SmaI

Insert A'-Sma
GCTCCCGATCTCATGAATCTTTCATGTATTCATTCACATCATTTTCGTGGATCCTGATAT

Insert A-Sma
GCTCCCGATCCACGAAAATGATGTGAATGAATACATGAAAGATTCATGAGATCGGTAATAACTGATAT

Parent
TGACCGGATCCGGAGC
BamHI

Insert A-Bam
TGACCGGATCCACGAAAATGATGTGAATGAATACATGAAAGATTCATGAGATCCGGAGC

One line of evidence suggests that it does not repress transcription completely. The MAT α promoter consists of a divergent UAS with an a1/ α 2 element to the right (See Figure 7.1 and Siliciano and Tatchell, 1984). In diploids the MAT α 1 gene is repressed more than 100 fold, whereas MAT α 2 transcription is repressed approximately five-fold. (Nasmyth *et al*, 1981a). This is, however, a rather special case in that it is an example of autoregulation and therefore the MAT α UAS may be especially refractory to the effect of an a1/ α 2 element. Alternatively the MAT α UAS may be in itself affected by a1/ α 2-mediated repression via an indirect pathway, not involving the a1/ α 2 element at MAT α .

In order to mimic the situation found at MAT α with another UAS, the CYC1-lacZ fusion was used. At the junction between the CYC1 sequence and the URA3 sequence there is a SmaI site, which can be used to insert an a1/ α 2 element upstream of the CYC1 UAS. Insert A was inserted into this SmaI site in both orientations to produce two constructions, A-Sma and A'-Sma (Figure 5.3). They were used to transform yeast strains Y65 and K700 and the levels of RNA produced by the CYC1-lacZ fusion were measured. Figure 5.3 shows that both of the plasmids carrying the a1/ α 2 element are transcribed at the same level as the parent in strain K700, but that in strain Y65 transcription from the a1/ α 2 element-containing plasmids is partially repressed. The results of this experiment and another similar one were quantitated by cutting the protected fragments out of the gel, and measuring their radioactivity in a scintillation counter. This revealed that the plasmids with an a1/ α 2 element in the SmaI site were repressed 7 - 11 fold as compared with the approximately 200 fold repression seen with an element in the XhoI site. Thus the a1/ α 2 element does function when placed upstream of the UAS, but not to full effect. It appears that,

to achieve effective repression, the element must be inside the region considered as the promoter, but that it need not overlap any essential promoter element, in contrast to the situation in procaryotes, where operator sequences appear to overlap promoter elements.

5.6 α_2 -mediated Repression Recognizes A Different Sequence

We have shown that the sequence motif (T/C)C(A/G)TGTNN(A/T)NANNTACATCA is sufficient to confer α_1/α_2 -mediated repression. We know from the biology of yeast mating type that genes carrying this sequence (the haploid specific genes) are not repressed by α_2 alone (Figure 1.2). α_2 mediated repression must involve a different, though possibly related, DNA sequence. To address this question, we have analysed the DNA sequences of the 5' flanking regions of three genes which are repressed by α_2 alone, looking for sequences specific to these genes which are related to (but different from) the α_1/α_2 consensus. Using the DIAGON program of Staden, we have found that MFa1, BAR1 and STE2 all have a 30 bp sequence in common. Although in this instance we have only three examples of the sequence, it is clear from the extent of conservation that this sequence has some function.

To test the function of this sequence, we have inserted a restriction fragment carrying it into the CYC1 promoter. In this way, we have shown that a 97 bp fragment containing the putative α_2 element from the STE2 gene can confer α_2 control on the CYC1 promoter. The fragment was inserted between the UAS and the TATA box of the CYC1 promoter, and the levels of β -galactosidase in isogenic α_1^- and $\alpha_1^- \alpha_2^-$ strains was then measured; the only genotypic difference between these strains is that the former produces α_2 but the latter does not. If this sequence confers α_2 -mediated repression, we would

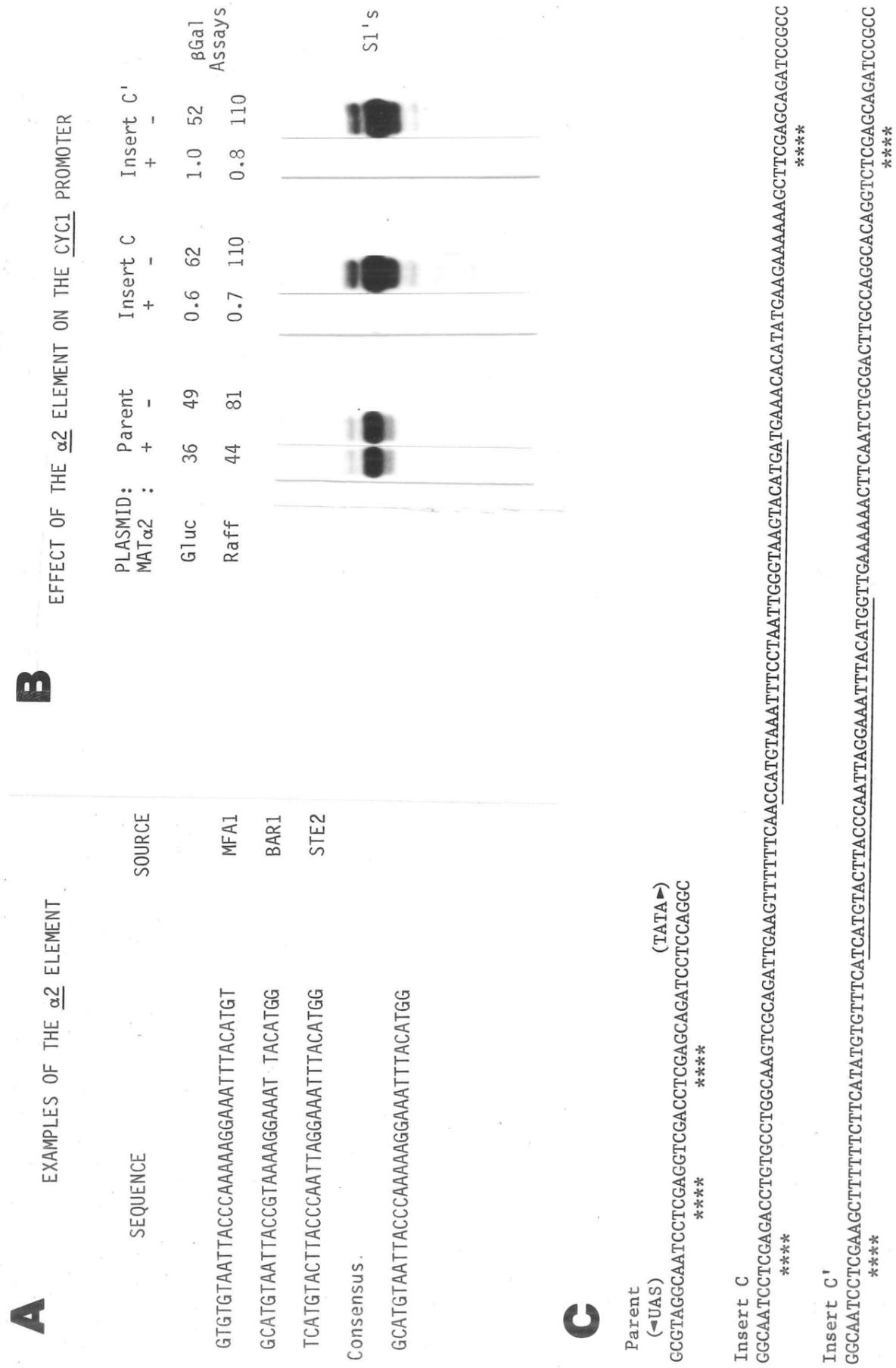
FIGURE 5.4 The Sequence of the α_2 element and its effect on the CYC1 promoter

(A) The sequences of the α_2 elements that have been found at the MFa1, BAR1, and STE2 genes are shown. Note that the consensus sequence is an approximate inverted repeat. Sequence information is from (A. Brake personal communication, V. Mackay, personal communication; and this work, respectively)

(B) The effect of the α_2 element on the CYC1 promoter. A 97 bp fragment of the STE2 gene carrying an α_2 element was inserted in both orientations into the CYC1 promoter. Here I show the levels of β galactosidase produced by these constructs in isogenic α , α_1^- and $\alpha_1^- \alpha_2^-$ cells, both for raffinose and for glucose as carbon source. Levels of RNA produced were measured by the protection of radiolabelled DNA from S1 nuclease digestion.

The strains used were RS1:(MAT α) and its derivatives RS3:(MAT α_1^- MAT α_2^-) and M48:(MAT α_2^+ mat α_1 ..LEU2). See Table 2.1.

(C) Sequences of the constructions with the α_2 element underlined. Asterisks denote the limits of the inserted fragments.



expect the plasmid to direct β -galactosidase production in the $\alpha 1^- \alpha 2^-$ cell, but not in the $\alpha 1^-$ cell. Figure 5.4 shows, as predicted, that the 97 bp fragment regulates β galactosidase activity in α cells or in $\alpha 1^-$ cells, but not in $\alpha 1^- \alpha 2^-$ cells. Correctly initiated transcript is found in $\alpha 1^-$ cells, but no transcript is found in $\alpha 1^- \alpha 2^-$ cells and plasmid copy number is invariant, implying that the control is at the transcriptional level. Furthermore, the constructions are still under the same carbon source dependence as the parent plasmid, demonstrating that they are being driven by the CYC1 promoter.

Thus, the 97 bp fragment is able to bring the CYC1 promoter under $\alpha 2$ control. Because of the size of this fragment, it is not clear from these data whether the consensus sequence alone is sufficient for repression, or whether some other sequence on the fragment is also involved. It seems probable that the consensus sequence described here is responsible since it is the most dramatic homology between the MFA1, BAR1 and STE2 sequences, and because it has a sequence which is related to that of the $\alpha 1/\alpha 2$ element. To a lesser extent, the experiments with restriction fragments containing the putative $\alpha 1/\alpha 2$ element that are described above do not demonstrate rigorously that the DNA sequence elements, identified by sequence comparisons, are sufficient to confer cell-type-specific transcription. It is formally possible that both the restriction fragments that were inserted into the CYC1 promoter carry other sequences, which are required for control of transcription. This possibility can be eliminated by inserting into the CYC1 promoter a segment of DNA which contains only those sequences we believe are required. This can be done most simply by synthesizing the requisite DNA fragments in vitro, using the technique of in vitro oligonucleotide synthesis. The synthetic oligonucleotide can then be

inserted into the CYC1 promoter and its effect on transcription measured. Such an experiment is shown in Figure 5.5. In a cell expressing both the a1 and the α2 gene products, the transcription from the CYC1-lacZ fusion is repressed, demonstrating that the sequence 5'-TTCATGTTATTATTACATCAT-3' is sufficient to confer cell-type-specific repression on an unrelated promoter. The analogous experiment with the α2 element has been performed by Johnson and Herskowitz (1985) and shows that the sequence

5'-CATGTAATTACCTAATAGGGAAATTTACACGC-3' is sufficient for α2-mediated repression.

Johnson and Herskowitz also inserted their α2 element into the SmaI site just upstream of the CYC1 UAS and found 10 to 20 fold repression in α2⁺ cells, demonstrating that these two different elements have similar properties, that is, both the a1/α2 and the α2 elements exert partial repression when placed upstream of the CYC1 UAS.

FIGURE 5.5

Effect of a synthetic oligonucleotide on the CYC1 promoter.

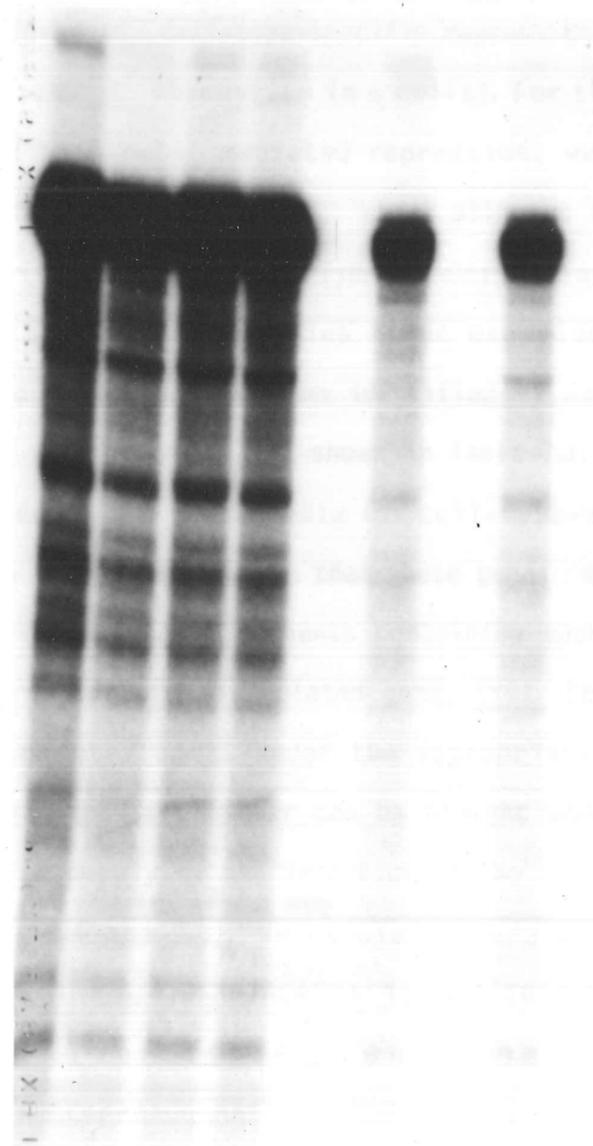
A synthetic oligonucleotide containing the sequence of the a1/a2 element found at HO -411 (Table 5.1) was inserted between the UAS and the TATA box of the CYC1 promoter and RNA levels were measured in α and a/α cells. The strains used were K700 (MAT α HMR α) and its derivative Y65 (MAT α HMR α E Δ , which expresses a1 (See Table 2.1).)

(B) Sequences of the constructions used. The asterisks denote the limits of the inserted sequence.

a

EFFECT OF A SYNTHETIC OLIGONUCLEOTIDE ON THE CYC1 PROMOTER.

PLASMID: Parent Insert D
a1: + - + - + - + -



b

Parent:
GGCTAATCCTCGAGGTCGACCTCGAGCAGA

Insert D:
GGCTAATCCTCGAGGTCGATTCATGTTATTATTTACATCATTTCGACCTCGAGCAGA
**** ****

5.7 Discussion

5.7.1 Summary of Results

We have identified, using the technique of DNA sequence comparisons, two separate sequence motifs that appear to confer two different forms of cell-type-specific repression (α_1/α_2 repression in diploids, and α_2 repression in α cells). For the sequence which is recognized by α_1/α_2 -mediated repression, we have about a dozen separate instances of the motif, which allow us to devise a consensus sequence (T/C)C(A/G)TGTNN(A/T)NANNTACATCA (Table 5.1). For the α_2 sequence, this work identifies three examples (Figure 5.4) and a further two examples have been identified by Johnson and Herskowitz (1985); all five examples are shown in Table 7.1. To demonstrate that these sequences are responsible for cell-type-specific expression, rather than some other feature that these genes might have in common, we have inserted small fragments containing each of these 'elements' into the promoter of an unrelated gene, CYC1. In these constructions the CYC1 promoter comes under the appropriate cell-type control, demonstrating that a promoter can be brought under a particular type of regulation merely by the insertion of the appropriate element into it (Figures 5.4 and 5.5). ^{+ Brent & Ptashne 1984} We conclude therefore that the sequence (T/C)C(A/G)TGTNN(A/T)NANNTACATCA is sufficient to confer α_1/α_2 -mediated repression, and that the sequence GCATGTAATTACCCAAAAAGGAAATTTACATGG (or a similar sequence) is probably sufficient to confer α_2 -mediated repression. In an analogous experiment, A. Johnson and I. Herskowitz (1985) have shown that the sequence 5'-CATGTAATTACCTAATAGGGAAATTTACACGC-3' is sufficient for α_2 -mediated repression.

We find α_1/α_2 elements spread throughout 1400 bp of DNA upstream of the HO AUG. The distribution of these elements is able to account

for our observation that a number of deletions of this region all retain a1/α2 control, and that when the TRP1 gene is inserted into a variety of positions within this region it comes under a/α control. Similarly the position of the a1/α2 element within the MAT_α gene is consistent with the deletion analysis of this gene (Siliciano and Tatchell, 1984), which defined the lefthand end of the region that is necessary for a/α control. The first deletion which abolishes a/α control has in fact removed 8 bp of the a1/α2 element defined here.

These two sequences, therefore, define two "eukaryotic operator" sequences, in that they are short DNA sequences which repress adjoining genes when present in cis. The a1/α2 element however differs from a prokaryotic operator in that it is not necessarily found adjacent to promoter elements (the UAS or the TATA box) and is unlikely to function by directly excluding RNA polymerase from binding to the promoter. In a typical prokaryotic promoter, the sequences that are essential for efficient transcription lie within 45 bp of the transcription start site (Pribnow, 1979), and the operators to which repressors bind are usually situated within this interval. Eukaryotic promoters, by contrast, extend over much greater intervals, with a TATA box that directs correct initiation and an upstream activation site (or enhancer) often hundreds of base pairs away (Banerji et al, 1981; Benoist and Chambon, 1981). We find that the a1/α2 element can repress transcription when inserted between the UAS and the TATA box, even though it is 60 bp away from either promoter element. The element does not work when inserted into the transcription unit itself, but when inserted upstream of the UAS, then an intermediate result is observed. Instead of the approximately 200 fold repression seen when the element

is between the UAS and the TATA box, we see a 7 to 11 fold effect. This effect is also seen *in vivo*, where an a1/α2 element upstream of a UAS exerts a partial effect at MATα. The MATα intergenic region comprises a divergent UAS, with an a1/α2 element to the right (Siliciano and Tatchell, 1984). In diploids, a1/α2 causes the complete repression of the α1 gene, and causes only partial repression of the α2 gene (Nasmyth *et al*, 1981a). It seems likely that the order of the elements (α2 - UAS - a/α element - α1) is responsible for the different levels to which α2 and α1 are repressed in the diploid.

5.7.2 Possible Mechanisms of Repression

Given these data, there are a number of possible mechanisms by which the element might act. One possibility is that it acts like a prokaryotic operator, binding a repressor protein, which by steric hindrance directly excludes the binding of an activator to promoter elements. It seems unlikely that this is the only mechanism by which the element acts, since complete repression can be achieved by an element that is at least 60 bp away from any essential sequences in the promoter. The steric hindrance could however be indirect, in that binding of a repressor to the a1/α2 element disturbs the positioning of other proteins, thereby blocking essential promoter sequences. A third type of model involves the signal transmission from the UAS to the TATA; although we do not understand how a UAS works, we must assume that some signal is transmitted from the UAS to the transcription start site. The a1/α2 element could block this transmission. For example the signal that is transmitted may be a transcription factor which interacts with the UAS and then slides along the DNA to the TATA box, repression could be achieved by

blocking the progress of the factor. Since the element is able to exert some effect when upstream of the UAS, blocking signal translocation cannot be the sole mechanism by which repression is achieved. In the constructions with inserts upstream, it is possible that steric hindrance of the UAS is occurring.

It is possible that a detailed series of experiments on the effect of different positionings of the a1/α2 element might reveal some information about the mechanism of action of a1/α2-mediated repression or about the mechanism of transcriptional activation itself. (For example, the work of Hoschild and Ptashne (1986) on lambda repressor reveals important information on the way in which cooperative binding occurs.) I have not however made a systematic study of the effect of distance between the a1/α2 element and other DNA elements such as the UAS or the TATA box. Recent work by Siliciano and Tatchell (1986) has shown that the a1/α2 element does not function when placed 375 bp upstream of the CYC1 UAS.

5.7.3 The a1/α2 and α2 elements have related, but distinct sequences

We know from the biology of yeast mating type that the a1/α2 element cannot be recognised in any productive way by the α2 protein alone. α2 will recognize the α2 element, and can only recognize an a1/α2 element if the a1 gene product is present in the same cell (See Figure 1.2). How is this differential specificity achieved by the proteins and DNA sequences involved? Examination of the DNA sequences of the two different "operators" reveals that both are inverted repeats, suggesting that both bind protein dimers (or tetramers etc.). Furthermore there is the interesting observation that the two "operators" have similar sequences, in that they carry the same sequence motif (ATGT...ACAT) in inverted

orientation (See Figure 5.6a). What then are the differences between the sequences of the two operators which might be responsible for their different effects? The first and most marked difference is the spacing between the inverted repeats: the common ATGT...ACAT motif is separated by 11bp in the a1/α2 element, but by about 24bp in the α2 element. This observation may be important to understanding how the two "operators" are distinguished in vivo, since it suggests that the two elements are recognized by protein complexes with similar DNA sequence specificity but different steric properties. Secondly the sequence motifs found in inverted orientation may not be identical: the α2 element is symmetrical, carrying the sequence TACATG at both ends, whereas the a1/α2 element carries the sequence TACATCA at one end (Figure 5.6a).

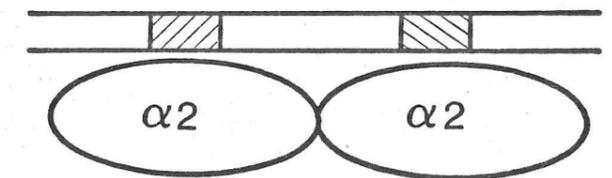
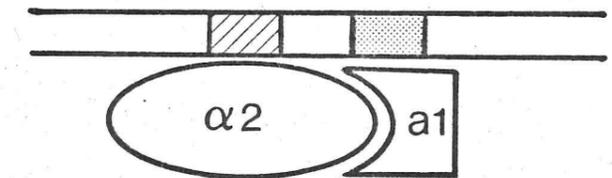
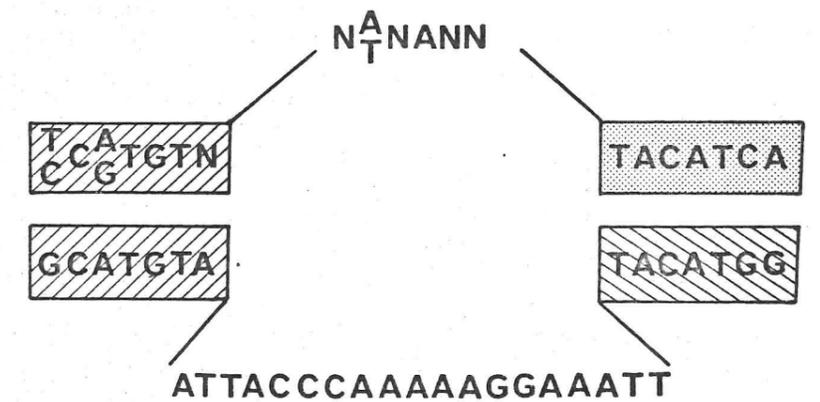
5.7.4 Possible Mechanism of a1 Gene Product Function

What does the presence of the a1 gene product in the cell do to extend the realm of action of α2 to include the a1/α2 element? The a1 gene product could, for example, bind to the α2 repressor complex such that the latter's quaternary structure was changed, thereby allowing it to recognize the shorter inverted repeat at the a1/α2 element. Alternatively the a1 and α2 gene products could form a heterodimer, in which each protein binds to one half of the a1/α2 element. The distinction between these models is: does a1 bind DNA directly? Protein sequence data suggest that it does. The carboxy terminal domains of a1 and α2 are homologous (Shepherd et al, 1984), and we know an α2-βgalactosidase fusion protein shows sequence-specific DNA binding (Johnson and Herskowitz, 1985). Furthermore, the homologous region includes the helix-turn-helix motif found in prokaryotic DNA binding proteins (Sauer et al, 1982). In the light of these

FIGURE 5.6

A) Comparison of the $\alpha 1/\alpha 2$ and $\alpha 2$ elements. The $\alpha 1/\alpha 2$ element (top) is shown aligned with the $\alpha 2$ element (below), with the common sequence motifs boxed. The TACATG motif is hatched, and the TACATCA motif is stippled.

B) The combinatorial aspect of mating type control. The $\alpha 2$ element (bottom) consists of two $\alpha 2$ half-sites correctly spaced to accommodate the $\alpha 2$ dimer. The $\alpha 1/\alpha 2$ element (above), in contrast, consists of an $\alpha 1$ and an $\alpha 2$ half site spaced correctly for the "heterodimer". It may be that only the spacing of the half-sites is important, or alternatively the difference between their sequences may also be important. The critical aspect is that the $\alpha 2$ half-site in the $\alpha 1/\alpha 2$ element cannot be productively recognized by the $\alpha 2$ dimer since the other half of the element will not allow both halves of the dimer to bind cooperatively.



homologies, it seems probable that both $\alpha 2$ and $\alpha 1$ are sequence-specific DNA binding proteins, and they might be expected to bind to similar sequences. We therefore prefer the model in which $\alpha 1/\alpha 2$ exists on the DNA as a heterodimer, with each protein binding to one half of the $\alpha 1/\alpha 2$ element (to one 'half-site')(Figure 5.6b). Thus the $\alpha 2$ dimer (or tetramer, etc.) will bind to the $\alpha 2$ element with its two half-sites 24bp apart, but cannot bind productively to an $\alpha 2$ half-site which is adjacent to an $\alpha 1$ half-site.

According to this model, there may be at least two reasons why $\alpha 2$ cannot recognize the $\alpha 1/\alpha 2$ site. First, the 10 bp distance between the ATGT...ACAT sequences may allow the binding of an $\alpha 2$ subunit to one half and the cooperative binding of an $\alpha 1$ subunit to the other, but not allow the cooperative binding of two $\alpha 2$ subunits. Second, it is possible that the TACATCA motif found in the right hand half of the $\alpha 1/\alpha 2$ element binds $\alpha 1$ more tightly than it does $\alpha 2$, whereas the sequence TACATG binds $\alpha 2$ (Figure 5.6).

The model depicted in Figure 5.6 for $\alpha 1/\alpha 2$ versus $\alpha 2$ control proposes that the specificity of an operator can be determined by three major variables: 1) the presence of short DNA sequences to which particular factors bind specifically; 2) the presence of two or more such sequences in close proximity such that cooperative binding may occur; and 3) cooperative binding may be critically dependent on the spacing between the sequences. By extension, we can consider a system in which there are a number of proteins which recognize a number of half-sites, but which will only bind productively to such sites if two proteins can bind to their half-sites co-operatively. For this to occur, the proteins must interact with each other, and the half-sites must be correctly spaced to allow such an interaction. The advantages of such a system are the increased specificity thus gained,

and the immense flexibility of the system. The system is flexible in evolutionary terms, in that these small half-sites may arise by mutation and can be shuffled to create operator sequences with different specificities. The system is also flexible in that it allows combinatorial interaction of these regulatory proteins, since regulation is only achieved at a composite site "X-Y" if both "X" and "Y" proteins are present in the same cell (or only at site "X---Y" if both "X" and "Z" are present, etc.) In such a model a set of DNA-binding repressors may recognize the same DNA sequence (half-site), but form different protein-protein interactions. Hence, proteins which recognize the same sequence could perform very different functions, repressing quite unrelated sets of genes as a result of the different protein-protein interactions they can make.

In this light the observations on the homeo domain are significant. This protein sequence is found in the engrailed, Antennapedia, fushi tarazu and Ultrabithorax gene products (Poole et al, 1985; McGinnis et al, 1984a; Scott and Weiner, 1984), as well as in a number of genes of unknown function in higher organisms, including vertebrates (Carrasco et al, 1984; McGinnis et al, 1984b). MAT α 1 and MAT α 2 both show protein sequence homology to the homeo domain, and the homeo domain also appears to conserve the prokaryotic helix-turn-helix motif and is therefore believed to be a DNA binding domain (Laughon and Scott, 1984); this hypothesis is consistent with the observation that antibodies raised against different gene products that contain the homeo domain recognize nuclear proteins (White and Wilcox, 1984; Beachy et al, 1985; Carroll and Scott, 1985; DiNardo et al, 1985). The four genes from Drosophila have quite distinct functions, involved in different aspects of the

determination of cell fate. It therefore seems paradoxical that some of these proteins are identical in the region which (by analogy with the prokaryotic proteins) would be mainly responsible for the sequence specificity of DNA binding (Laughon and Scott, 1984). It seems likely that the different functions of these different proteins are encoded by the non-homologous regions. These regions could either specify a second DNA-binding specificity, or, analogously to a1 and α2, they could specify a protein-protein interaction. Indeed, the homeo domain may be just one of several classes of sequence-specific DNA binding domains. In the case of a1 and α2 these domains are associated with polypeptide sequences which confer negative control, but there is no reason why they could not alternatively be associated with protein domains which confer positive control.

The idea that homeo proteins may interact in a combinatorial manner with each other and with other proteins is rather attractive, since (as with a1 and α2) different combinations of these selector genes in different cells would activate different cell-type specific genes. This would allow the combinatorial genetic addresses (Struhl, 1982; Lawrence and Morata, 1983) of different cells to be "read" directly by the promoters of the different cell-type-specific genes.

CHAPTER SIX

ISOLATION OF MUTATIONS IN MAT α 2

6.1 Introduction

In this chapter I will describe the isolation and characterization of mutations in the MAT α 2 gene. These mutations have a curious phenotype, in that they prevent a1/ α 2-mediated repression from occurring, but have little effect on α 2-mediated repression. The mutants behave as if the α 2 gene product is unable to "recognise" or interact with the MATa1 gene product, but repression by MAT α 2 alone does still occur.

Mutations at MAT α 2 have been isolated in a number of ways: screening for sterility (MacKay and Manney, 1974a,b), in vitro construction of Xho-linker mutations (Tatchell et al, 1981), and site-directed mutagenesis (Porter and Smith, 1986). All three methods produced mutations which are defective in both functions of MAT α 2, mating and sporulation; this is the phenotype which results from a total loss of function at MAT α 2. However, all three methods also produced rare mutations which are defective in mating, but are still sporulation-competent (MAT α 2-4, α X109, and clones 56 and 96 respectively.) Xho-linker mutation α X109 is of particular interest, since it lies outside of the transcribed region of the MAT α 2 gene, and therefore the protein product produced by α X109 should be identical to the protein made in MAT α 2⁺. The Xho-linker deletes the α 2 TATA box, probably leading to reduced transcription levels. The phenotype of this mutation suggests that merely reducing the level of α 2 gene product can lead to loss of mating ability, without affecting sporulation.

Although both of the in vitro mutageneses of MAT α 2 produced mutations with such a phenotype, the converse effect was not seen,

that is, mutations affecting only diploid function were not observed. The isolation of such mutations is described here.

6.2 Isolation of Mutations

These mutations were isolated by serendipity. The screen in which they were isolated was designed with the intention of isolating suppressors of a sir3^{ts} mutation, and worked as follows. The strain M30 was constructed, which has the genotype HMLa MAT α HMRa sir3^{ts} ho::TRP1 trp1-1. As a result of the sir^{ts} mutation, strain M30 is an ordinary α cell at 25°C, but at 34°C it expresses a and α information and becomes sterile. The ho::TRP1 construction in M30 is an insertion of the TRP1 transcription unit into the H0 gene (from -718 to +1096) that has been transplacated into the H0 locus. In this construction TRP1 is subject to a1/a2-mediated repression (See Section 5.1.2). Thus at 25°C, the copy of TRP1 at the H0 locus is expressed, but when the temperature is raised and the silent copies become derepressed, then the TRP1 transcript is switched off. Strain M30 is, therefore, capable of growth on plates lacking tryptophan at 25°C, but not at 34°C.

The first stage in my attempt to isolate suppressors of the sir3^{ts} allele was to select for growth at 34°C on plates lacking tryptophan. There are a number of ways in which this Trp⁺ phenotype could arise. For instance, mating type switching still occurs in strains lacking the H0 gene, albeit at a greatly reduced frequency. Hence we would expect the genotype HMLa MATa HMRa to arise at a frequency of about 10^{-6} (Hawthorne, 1963a), due to rare switching at MAT. This strain lacks α information, and therefore a1/a2 -mediated repression cannot occur.

mat α 2⁻ mutations are expected at a high frequency, but a1⁻ mutations are not, since two copies of the a1 gene (HMLa1 and HMRa1) have to be mutated for this to occur. The Trp⁺ phenotype could also

occur by a gene conversion event, or other DNA rearrangement that 'liberates' the TRP1 transcription unit from a1/α2 control, for instance, if the copy of TRP1 at the H0 locus were used to heal the the chromosomal trp1 allele, then a Trp^+ phenotype would arise.

None of these events, however, can overcome the mating defect of the parent strain: that is, none of these mutants will mate as α at the restrictive temperature. To eliminate the events listed above, the colonies which grew on plates lacking tryptophan at 34°C were screened for their ability to mate as α at that temperature. The only mutations that I expected to generate both the Trp^+ phenotype and the α -mating phenotype at 34°C were suppressors of sir3^{ts}. Approximately 2×10^8 cells were plated out on the selective plates at 34°C , and gave rise to 477 Trp^+ colonies. Of these, 110 showed some α mating ability.

I thought that the 110 mutants that were Trp^+ α maters were suppressors of sir3^{ts}: I wished to investigate suppressors that were dominant, so the mutant strains were mated to the strain M28 to test for the dominance of "suppression." M28 has the genotype HMLa MATΔ HMRa sir3^{ts}, so recessive suppressors would revert to having the α/α phenotype at 34°C in this strain. A dominant suppressor of sir3^{ts}, on the other hand, would still be Sir^+ in this diploid, and would therefore have α phenotype. Recessive suppressors of sir⁻ and sir^{ts} mutations have been isolated at gene knock out frequencies by other workers, and were therefore expected here.

The mutant/M28 diploids were tested for α mating ability and for ability to grow on plates lacking tryptophan. Most mutants were unable to do either, while at least six independent events gave rise to Trp^+ diploids that would not mate. When the mating of the haploid mutants from which these diploids were derived was retested, they all showed severe defects in α mating.

Only four distinct examples were found of strains which produced α maters at 34°C when mated with strain M28. Two of these examples came from the same culture, and therefore may not represent entirely independent events, but their phenotypes were sufficiently different for them to be analysed separately.

6.3 Analysis of the Mutants

Since a mutation in MAT α 2 was not expected, the genetic analysis which leads to that conclusion is rather confusing.

In order to determine whether the mutations isolated were second-site suppressors of the sir3^{ts} mutation, or if they were intragenic suppressors (revertants or pseudo-revertants), the mutants were crossed to an ordinary a strain K163 (HML α MATa HMRa SIR⁺).

If the mutation were an unlinked suppressor of sir, then the cross:

a	α	a	<u>sir3^{ts}</u>	<u>SSR^D</u>
α	a	a	<u>SIR⁺</u>	<u>ssr⁺</u>

would be expected to give

mating phenotype	α^{ts}	α^+	a^{ts}	a^+
relative frequency	2	6	1	7

linkage between SSR and SIR will decrease the numbers of temperature sensitive maters produced; in the extreme case, where a reversion has occurred, then no temperature-sensitive maters will be segregated. However, two thirds of the ts maters produced should be α 's. The reason for this asymmetry is the presence of the HMLa allele segregating in the cross; most yeast strains are HML α HMRa and are therefore non-maters if they are Sir⁻, but strains carrying HMLa HMRa and MATa will mate as a , regardless of their Sir status. (The linkage

between the mating type loci can be ignored.)

6.3.1 Analysis of the mutant 5-8

One of the four mutations isolated, 5-8, did not behave like a mutation in MAT α 2, and I will describe its behaviour first to demonstrate the sort of results that were originally expected. 11 tetrads were dissected from the diploid 5-8/K163 (a tetrad consists of the four haploid spores that result from a single meiosis, and therefore within each tetrad copies of each Mendelian allele should segregate 2:2.) Mating type segregated 2:2 in all tetrads, and all segregants mated at 34⁰C. Thus no Sir^{ts} segregants were observed, implying that the suppressor of sir3^{ts} is tightly linked to SIR3, and implying that the strain is probably a revertant. However, when this same strain, 5-8, was crossed to another "ordinary a cell", AB10(HML α MATa HMRa SIR⁺), then a completely different result was obtained. 8 tetrads were dissected, mating type segregated 2:2, but individual tetrads contained 1, 2 or 3 maters at 34⁰C. The four possible mating phenotypes were distributed as follows:

5-8/AB10

mating phenotype	α^{ts}	α^+	a ^{ts}	a ⁺
frequency	10	5	1	13

The strain 5-8 is behaving in a manner similar to the parent strain M30, which gave:

M30/K163

mating phenotype	α^{ts}	α^+	a ^{ts}	a ⁺
frequency	8	6	3	11

Theory predicts.

8 8 4 12 for a cross a α a sir^{ts}/ α a a SIR

The data from the 5-8/K135 cross imply that the mutation has disappeared. It is not clear why this mutant is behaving differently in what should be identical crosses, and the mutation has not been pursued further.

6.3.2 Analysis of the mutants 3-6, 5-9, and 8-2.

Some difficulty was encountered in obtaining efficient sporulation from the diploids 3-6/K163, 5-9/K163, and 8-2/K163. High levels of asci were eventually achieved; they were dissected and analysed, with a result unlike those described above. Mating type segregated 2:2, but every tetrad contained 2 temperature-sensitive maters, with no difference between the behaviour of the a's and the α 's.

	α^{ts}	α^+	a^{ts}	a^+
3-6	9	11	11	9
5-9	7	15	15	7
8-2	17	17	17	17
TOTAL	33	43	43	33

A number of explanations for the segregation pattern observed were considered, but none fit the data. There was no way of explaining the absence of additional a^+ maters, which should arise from the genotype HMLa MATa HMRa sir3^{ts}. Southern analysis confirmed that both M30 and the mutants isolated contained HMLa.

One possibility considered was that the SIR3 allele had undergone gene conversion to sir3^{ts}. To investigate this possibility, segregants from the first cross were outcrossed a second time to see if the apparently Sir^+ strains did all carry a sir3^{ts} allele.

The spores from one tetrad from each mutant 3-6, 5-9, and 8-2 were mated to SIR⁺ strains of the opposite mating type and sporulation was attempted. In each case the MATa segregants yielded diploids which sporulated, whereas the MAT α segregants showed a sporulation defect reminiscent of the original mutant strains. Since the a and α strains used to outcross the F1 were isogenic, this suggested that the sporulation defect was linked to MAT. This in turn suggested that the original mutation might be in MAT α , and might allow the MAT α locus to function normally in haploids, but fail to respond to the expression of a1 information. Thus cells carrying this hypothetical mutation in MAT α will mate as α even if a1 information is being expressed, and diploids carrying the mutation will be sporulation-defective.

The screening procedure asked for a strain of the genotype

HMLa MAT α HMRa sir3^{ts}

HMLa MAT Δ HMRa sir3^{ts}

to mate as α , and to fail to perform a1/a2-mediated repression, the assumption being that such a phenotype could only be produced by switching off the expression of the silent loci. If, however a mutation caused the strain to become refractory to the presence of a1 gene product, then the same phenotype would arise. This model makes the specific prediction that the diploids formed from such a strain will still mate as α : for example

HMLa MAT α ^{*} HMRa sir3^{ts}

HML α MATa HMRa SIR3⁺

should be an α mater.

This prediction was tested, and Table 6.1 shows that the diploids 5-9/K163 and 8-2/K163 mate strongly, whereas 3-6/K163 shows a

TABLE 6.1

Mating Behaviour of the Mutants

Strain	Mating as α at			Mating as a at		
	25°C	30°C	33°C	25°C	30°C	33°C
3-6	--/+	--/+	--/+	-	--/+	-
5-8	+/-	+/-	+/-	-	--/+	-
5-9	+++	+++	+++	-	--/+	-
8-2	+++	+++	+++	-	--/+	-
M30 (parent)	+++	--/+	-	-	--/+	-
3-6/K163	+/-	+/-	--/+	+/-	+/-	+/-
5-8/K163	--/+	--/+	--/+	--/+	--/+	--/+
5-9/K163	+++	+++	+++	-	++	-
8-2/K163	+++	+++	+++	-	++	-
K163	-	-	-	+++	+++	+++

considerable mating defect.

These diploids should not sporulate properly. In the original outcrossing of the mutations, difficulty was encountered in persuading these strains to sporulate, but eventually high levels of asci were produced. How could this be achieved? If the expression of HML α were able to overcome the mutation at MAT α , then the strain shown immediately above will sporulate only if it undergoes homozygosis to sir3^{ts}, thereby allowing HML α expression. Now, when a strain of genotype

HML α MAT α ^{*} HMR α sir3^{ts}

HML α MAT α HMR α sir3^{ts} is sporulated,

and haploid progeny are analysed, four types of spore will arise.

<u>HML</u>	<u>MAT</u>	phenotype
a	a	a ⁺
a	α [*]	α ⁺
α	a	a ^{ts}
α	α [*]	α ^{ts} because <u>MATα</u> [*] is hypostatic to <u>HMLα</u> , as proposed.

This explains the segregation pattern observed. The only diploids that could give rise to spores were those that were homozygous for sir3^{ts}. In the progeny, the strains carrying HML α were maters at the restrictive temperature, irrespective of whether they were MAT α or MAT α , while the strains carrying HML α were all temperature-sensitive maters. Outcrosses confirmed that the apparently Sir⁺ progeny were in fact HML α sir3^{ts}, as expected.

This explanation rests on the assumption that HML α ⁺ is able to overcome the mutation at MAT α , and that therefore homozygosis has

occurred at sir3^{ts}. We can test this by starting off with a strain that is homozygous at sir3^{ts}, and observing the segregation pattern that results. The three mutants 3-6, 5-9, and 8-2 were mated to strain K123 : HML_α MAT_α HMR_α sir3^{ts} to produce diploids that already have the genotype proposed above :

HML_α MAT_α^{*} HMR_α sir3^{ts}

HML_α MAT_α HMR_α sir3^{ts}

Tetrads from these diploids were dissected and analysed. The results were identical to those obtained with the SIR⁺ strain K163. All tetrads produced 2 spores that were temperature-sensitive for mating, and the distribution of phenotypes was as follows:

	<u>α</u> ^{ts}	<u>α</u> ⁺	<u>a</u> ^{ts}	<u>a</u> ⁺
3-6	7	7	7	7
5-9	7	9	9	7
8-2	13	9	9	13
TOTAL	27	25	25	27

Thus the same distribution of temperature sensitive maters is produced when any of the mutants is crossed to either K123 or K163, as predicted by the proposed model.

6.3.3 Conclusion

It appears that the original mutant strains carried a lesion which allows α mating even in the presence of a1 expression. From the segregation pattern observed, one can deduce that the mutation is hypostatic to HML_α expression both in haploids and in diploids. The 2:2 segregation also implies that the mutation is fairly closely linked to MAT_α for the following reason. If MAT_α recombined away from the lesion, then it would produce a spore that was Mat_α⁺, half such spores would carry HML_α and would therefore produce a tetrad with 3

temperature-sensitive maters, as opposed to two. No such tetrads were found.

The observation that the effect of the mutation can be overcome by HML α expression suggests that it is a loss of function mutation. This observation, along with the dispensability of MAT α 1 for sporulation, further imply that the lesion is within MAT α 2.

6.4 Cloning and Sequencing the Mutations

Because of the novel and curious nature of the mutations, it was decided to clone and sequence the mutant alleles that gave a strong phenotype. Mutation 3-6 was not cloned; its defect in mating ability puts it in the same category as Xho-linker mutations such as α X75 (Tatchell *et al*, 1981).

The MAT α alleles from 5-9 and 8-2 were cloned into M13 sequencing vectors in the following manner. DNA was prepared from the yeast strains, and digested with HindIII, which produces a 4.2 kb restriction fragment. The genomic restriction fragments were separated by electrophoresis, and DNA was recovered from the region of the gel corresponding to 4.1 to 4.3 kb, approximately. These DNA fragments were ligated into HindIII cut M13mp10 and plaques were screened by plaque hybridisation.

Recombinant M13 clones carrying the MAT α fragment could then be sequenced directly. This was done by using, as primers in the dideoxy-sequencing reaction, synthetic oligonucleotides that were complementary to sequences within the MAT α 2 gene. For this method to work, however, the MAT α DNA must be inserted into the M13 vector in the correct orientation, such that the M13 template strand is complementary to the synthetic oligonucleotide primer, rather than identical to it in sequence.

Recombinant M13 carrying the MAT α fragment were screened by restriction digestion to find inserts that were in the correct orientation. These were isolated for the mutant alleles 5-9 and 8-2, but all the recombinants isolated from the parent strain M30 carried inserts in the wrong orientation. In this case, therefore, the cloned DNA was excised and re-inserted into M13 in the correct orientation. Single-stranded DNA was then prepared and the MAT α 2 sequence determined by dideoxy-sequencing, using two oligonucleotide primers complementary to MAT α .

When the sequence of the MAT α DNA isolated from M30 is compared with the sequence of Astell *et al* (1981), two changes are observed (Figure 6.1). These changes represent polymorphisms. The base change at the 55th codon has been observed previously (Takano *et al*, 1984), and the change at the 5th codon is silent.

When the MAT α 2 sequences from 5-9 and 8-2 were compared with M30, they were found to differ at a single nucleotide (Figure 6.1). In both 5-9 and 8-2 nucleotide number 1405 had mutated from G to T (on the sense strand). This causes a change in coding capacity from cysteine to phenylalanine at codon number 32.

6.5 The Cloned DNA is Sufficient to Confer the Phenotype

Thus the two mutant alleles that I have sequenced carry the identical base pair change, and can be considered as a single allele, which I will refer to as MAT α 2*. This allele has been generated on two independent occasions. In order to demonstrate that the cloned MAT α * allele is sufficient to confer the mutant phenotype, the cloned DNA was used to replace the chromosomal MAT DNA in the strain M28 (HMLa mat::LEU2 HMRa ho leu2 sir3^{ts}). Yeast transformants in which the MAT α DNA had replaced the chromosomal mat::LEU2 construction were

FIGURE 6.1

Comparison of the cloned MAT sequences with the sequence of Astell *et al*, 1981.

The coding regions of *MAT α 2* DNA from strain M30 and from the mutants 5-9 and 8-2 were compared with the sequence originally published by Astell *et al* (1981). All changes that were identified are shown here, along with their effect on the predicted protein sequence (if any).

FIGURE 6.1

Comparison of cloned MAT sequences with the sequence of Astell *et al*, 1981.

```

                                1485.
(M) N K I P I K D L L N P Q I T D E
GCAAGAAAAAATGAATAAAATACCCATTAAAGACCTTTTAAATCCACAAATCACAGATGA
M30                                ATA
5-9                                ATA
8-2                                ATA
                                    I

                                1405.
F K S S I L D I N K K L F S I C C N L P
GTTTAAATCCAGCATACTAGACATAAATAAAAAGCTCTTTTCTATTTGCTGTAATTTACC
M30                                TGC
5-9                                TTC
8-2                                TTC
                                    F

                                1336.
K L P E S V T T E E E V E L R D I L V F
TAAGTTACCAGAGAGTGTAACAACAGAAGAAGTGAATTAAGGGATATATTAGTATT
M30                                GGA
5-9                                GGA
8-2                                GGA
                                    G

L S R A N K N R K I S D E E K K L L Q T
CTTATCTAGGGCCAACAAAACCGTAAGATTAGTGATGAGGAGAAGAAGTTGTTGCAAAC

T S Q L T T T I T V L L K E M R S I E N
AACATCTCAACTCACTACTACCATTACTGTATTACTCAAAGAAATGCGCAGCATAGAAAA

D R S N Y Q L T Q K N K S A D G L V F N
CGATAGAAGTAATTATCAACTTACACAGAAAAATAAATCGGCGGATGGGTTGGTATTTAA

V V T Q D M I N K S T K P Y R G H R F T
TGTGGTAACTCAAGATATGATAAAACAAAAGTACTAAACCTTACAGAGGACACCGGTTTAC

K E N V R I L E S W F A K N I E N P Y L
AAAAGAAAATGTCCGAATACTAGAAAAGTTGGTTTGCAAAGAACATCGAGAACCCATATCT

D T K G L E N L M K N T S L S R I Q I K
AGATACCAAGGGCCTAGAGAATCTAATGAAGAATACCAGTTTATCTCGCATTCAAATCAA

N W V S N R R R K E K T I T I A P E L A
AAACTGGGTTTTCGAATAGAAGAAGAAAAGAAAAACAATAACAATCGCTCCAGAATTAGC

D L L S G E P L A K K K E *
GGACCTCTTGAGCGGTGAGCCTCTGGCAAAGAAGAAAGAATGAGCCCGAAAAACAATAT

```

identified by virtue of their ability to mate as α , and by their inability to grow on media lacking leucine. Both MAT α ⁺ DNA (from M30) and MAT α ^{*} (from 5-9) were used.

When the mating ability of these transformants was tested at 34°C, rather than 25°C, the transformants carrying MAT α ⁺ DNA were no longer able to mate, whereas the transformants carrying MAT α ^{*} DNA were still able to mate, implying that the cloned MAT α DNA fragment is sufficient to confer the mutant phenotype.

Finally, in order to obtain a quantitative measure of the effectiveness of the MAT α ^{*} mutation at α 2-mediated repression, and of its defect in a 1/ α 2-mediated repression, the levels of two cell-type-specific transcripts were measured: STE2 and H0. The STE2 promoter contains an α 2 element (Chapter 5), and RNA production is subject to α 2 mediated repression (data not shown), while H0 contains about ten a 1/ α 2 elements within its promoter, and is subject to a 1/ α 2-mediated repression (Jensen *et al*, 1983). Thus in wild type strains STE2 is expressed in a cells, but not in α cells, whereas H0 is expressed in both a and α cells, but not in a/α diploids. From the phenotype of the MAT α ^{*} mutation one would predict that strains carrying this mutation would fail to repress H0 effectively, but would still be capable of repressing STE2.

For both M30 and 5-9 MAT α DNA two independent transformants of strain M28 were taken, and mated to strain K699 (HML α MAT α HMR α ho SIR⁺). RNA was prepared from all these strains and the levels of STE2 and H0 RNA were measured by protection of radiolabelled DNA from S1 nuclease. Figure 6.2 shows that H0 is transcribed in all the haploids, and in the a/α ^{*} diploid - thus α ^{*} fails to repress H0. The first six lanes in Figure 6.2 demonstrate that α ^{*} is able to repress STE2 very well, as predicted, although some transcript is detectable.

FIGURE 6.2

Effects of the mutations on STE2 and HO transcript levels.

Cloned MAT α DNA from either M30 or mutant 5-9 was used to replace the chromosomal MAT DNA of the strain M28 (mat::LEU2). Transformants were taken and mated to strain K699 (MAT α), and RNA was prepared from both the haploid transformants and the diploids produced. The levels of STE2 and HO RNA were determined. Radiolabelled DNA complementary to Histone H2B RNA was mixed with the other DNA probes to act as an internal control.

Lane 1) MAT α cell (Strain K699)

Lane 2) MAT Δ cell (Strain M28)

Lanes 3 and 4) two independent transformants using MAT α DNA from the mutant 5-9

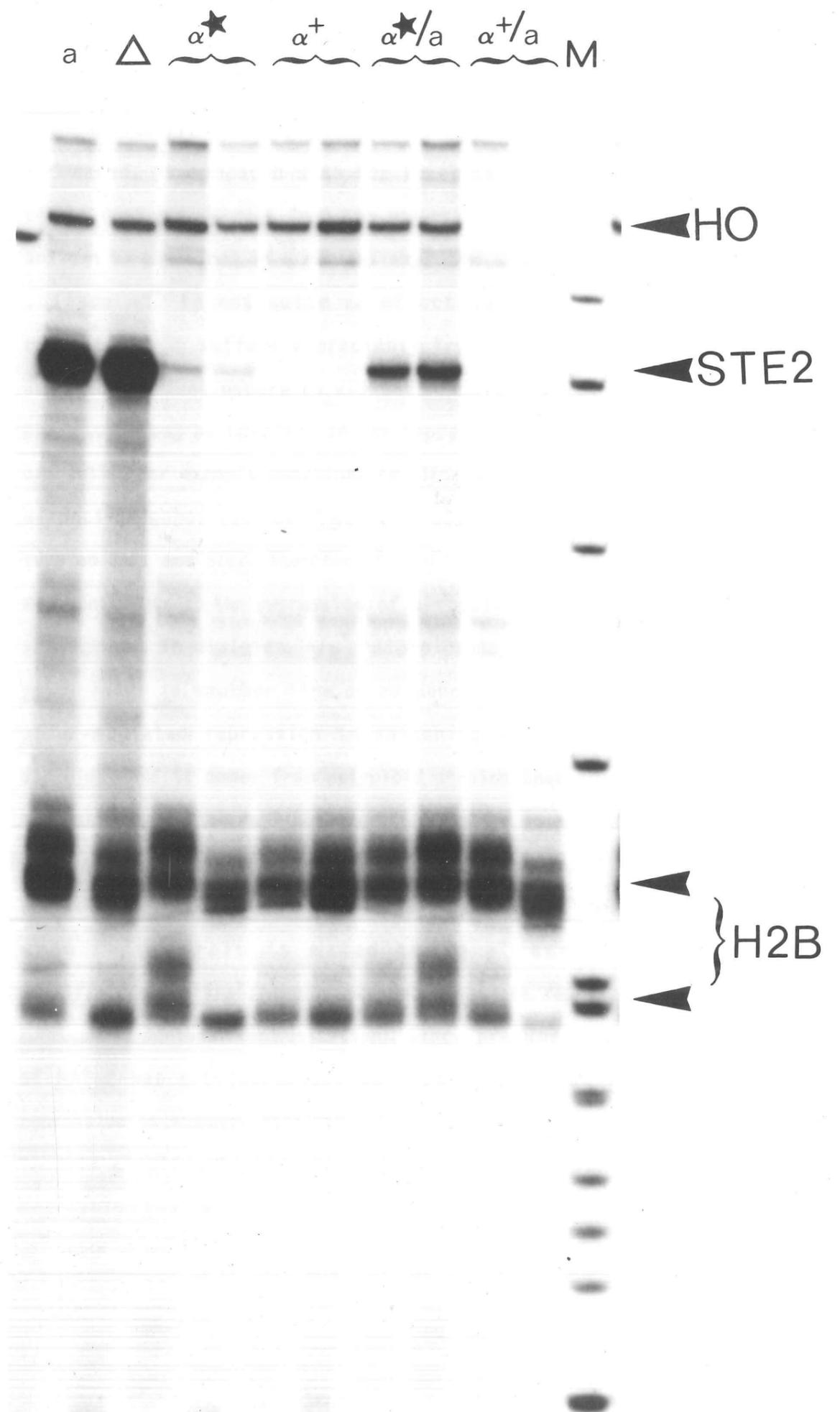
Lanes 5 and 6) two independent transformants using MAT α DNA from the parent M30

Lanes 7 and 8) K699 mated to the strains used in lanes 3 and 4, respectively

Lanes 9 and 10) K699 mated to the strains used in lanes 5 and 6, respectively

Lane 11) pBR x MspI, Sizes are 623, 527, 404, 309, 242, 238, 217, 201, 190, 180, 160(x3).

Cells were growing exponentially when harvested for RNA preparation. 40 μ g of total RNA were used per lane.



Curiously, in the $\underline{a}/\underline{\alpha}^*$ diploids, STE2 expression is considerably increased, as compared with the level seen in $\underline{\alpha 2}^*$ haploids. This effect does not detract from the essential nature of the $\underline{\alpha 2}^*$ mutation and can be explained in a various ways. One simple explanation is that, since $\underline{\alpha 2}^*$ is not quite so effective as $\underline{\alpha 2}^+$ at $\underline{\alpha 2}$ -mediated repression, it suffers appreciably from the doubling of cell and genome size that occurs in diploids. Additionally, $\underline{a1}/\underline{\alpha 2}$ -mediated repression may be involved in the repression of a-specific genes in $\underline{a}/\underline{\alpha}$ cells. For example mutations in STE5 lead to reduced expression of α -specific genes such as MF α 1 and STE3, and also a-specific genes, such as BAR1 and STE2. Therefore the $\underline{a1}/\underline{\alpha 2}$ -mediated repression of STE5 may contribute to the repression of a-specific and α -specific genes that occurs in diploids. $\underline{a}/\underline{\alpha}^*$ diploids do not, we assume, repress STE5. There is another line of evidence in support of the role of $\underline{a1}/\underline{\alpha 2}$ -mediated repression in switching off a-specific genes in diploids, and it comes from polyploid strains that are constructed by successively mating "a-like fakers" (alf's) to an $\underline{\alpha}$ strain. Strains of the genotype $\underline{\alpha}^+/\underline{alf}/\underline{alf}/\underline{alf}/\underline{alf}$ express a-specific genes, presumably due to some dilution or titration of $\underline{\alpha 2}^+$ gene product, but when this strain is mated to an \underline{a}^+ strain to produce $\underline{\alpha}^+/\underline{alf}/\underline{alf}/\underline{alf}/\underline{alf}/\underline{a}^+$, the a-specific genes are repressed once more (K.Tatchell, personal communication). Since presumably the expression of MAT $\alpha 2$ has not increased, we must assume that $\underline{a1}/\underline{\alpha 2}$ -mediated repression is somehow assisting in the repression of a-specific genes in this $\underline{\alpha}^+/\underline{alf}/\underline{alf}/\underline{alf}/\underline{alf}/\underline{a}^+$ strain. For these reasons, the STE2 expression that we observe in $\underline{a}/\underline{\alpha}^*$ diploids does not imply that the $\underline{\alpha 2}^*$ gene product is titrated by the presence of the a1 gene product.

6.6 Discussion

This chapter describes the isolation and characterisation of mutations in the MAT α 2 gene. These mutant alleles of MAT α 2 are unable to perform a1/ α 2-mediated repression but are still able to perform α 2-mediated repression, thus separating the two functions of the MAT α 2 gene product.

The genetic screen in which the mutations were isolated was originally designed to isolate dominant suppressors of a sir3^{ts} allele. The screen would however also isolate mutant strains in which a1/ α 2-mediated repression had failed, but α 2-mediated repression still occurred, as long as such mutations were dominant. Such mutations were rare, and three out of four of the mutations isolated were within the MAT α 2 gene, while the fourth mutation was not successfully characterised.

Of the three mutations at MAT α 2, one showed a considerable mating defect, placing it in the same category as other previously isolated MAT α 2 alleles that fail to support sporulation, but also show severe mating defects (such as α X75 in Tatchell *et al*, 1981.) The other two mutant strains showed strong α mating and unaffected α Factor production (data not shown), and these two mutant alleles were cloned and sequenced. This revealed that they both contained the identical nucleotide change, with the 32nd codon changing from TGC to TTC (cysteine to phenylalanine).

Further characterisation of this mutation confirmed that it exerted its effect at the level of tran^scription of the target genes. Thus MAT α 2^{*} was unable to repress H0 transcription in an a1/ α 2^{*} diploid, but it was able to repress the a-specific gene, STE2. This experiment also supported the hypothesis that a1/ α 2-mediated repression may have an important role in the repression of a-specific

genes in diploids. The observation made was that MAT α 2^{*} seems capable of repressing the α -specific gene STE2 in haploids, but in diploids it is not so effective. No such loss of efficiency is seen for the wild type allele. The reason for this difference may be MAT α 2^{*}'s inability to perform a1/ α 2-mediated repression. According to this model, α -specific genes are repressed in diploids via two mechanisms, firstly α 2-mediated repression acting directly on their promoters, and secondly by the a1/ α 2-mediated repression of an activator, such as STE5.

The MAT α 2^{*} allele appears quite unable to perform a1/ α 2 -mediated repression adequately - for example MAT α 2^{*} /MAT α diploids fail to sporulate at all. When I attempted to sporulate such a strain, sporulation only occurred when a gene conversion event at sir3^{ts} allowed the expression of α 2⁺ information from HML α .

Thus we have a mutation that confers a "dominant α mating phenotype" (that is, dominant over MAT α 1) but is hypostatic with respect to HML α ⁺ expression. A mutation with the same phenotype has been isolated by S.Harashima and Y.Oshima. They screened a HML α MAT α HMRA sir⁻ strain for α mateds and isolated a mutation at HML α . The first such allele to be sequenced has mutated at the same nucleotide as the MAT α 2^{*} alleles that I have sequenced: in Harashima's mutant the 32nd codon has changed from cysteine to tyrosine. In total then, three independent mutations in the α 2 gene that give this phenotype have been sequenced, and all lie within the same amino acid. Although three mutant alleles is too small a sample for the significance of this coincidence to be clear, it seems probable that the number of different mutations that can cause this phenotype is very small.

One intriguing possibility raised by the fact that the mutated amino acid is a cysteine, is that the interaction between the MAT α 1

gene product and MAT α 2 involves an disulphide bridge.

DISCUSSION

This thesis describes work which investigates the control of gene expression in the yeast Saccharomyces cerevisiae. I have looked at two different systems of control - the SIR system and the a1/α2 system. Both SIR and a1/α2 are involved in determining the mating behaviour in Saccharomyces cerevisiae, and both act by repressing, rather than activating, transcription. These two systems of negative control are however different, having distinct properties which suggest that they may act via different mechanisms.

Acting together, the products of the four genes SIR1,2,3 and 4 repress the expression of the silent mating type loci HML and HMR. Deletion analysis of plasmid-borne copies of these loci indicated that each locus contained two DNA elements involved in repression. One element was essential for repression ("E") while the other was important ("I"). Deletion of E led to complete derepression, whereas deletion of I led to only partial derepression (Abraham et al, 1984; Feldman et al, 1984). These elements lie between 900 and 1700 bp from the affected promoters (Abraham et al, 1984; Feldman et al, 1984), and in fact the HMRE sequence has been shown to repress the HMRa1 promoter at a distance of 2600 bp (Brand et al, 1985). SIR repression appears, therefore, to be unlike other previously characterised systems of repression in that the DNA sequences which are required in cis are found to lie a long way from the affected promoter. In contrast, deletion analysis of the MATα1 promoter implied that the a1/α2-mediated repression of this gene involved sequences within the promoter region (Siliciano and Tatchell, 1982)

Deletion analysis of HML and HMR revealed another curious feature of SIR-mediated repression: the cis-acting DNA elements, HMRE, HMRI, HMLE, and HMLI, all appeared to be tightly linked to ARS elements. Since ARS elements are believed to be specific origins of DNA replication, this suggested that DNA replication might be involved in SIR-mediated repression.

I have therefore investigated the possible role of DNA replication in the SIR-mediated repression of the silent loci. This was done using temperature-sensitive alleles of two of the SIR genes, SIR3 and SIR4. Chapter 3 describes preliminary experiments to choose a convenient assay for $Sir^{+/-}$ status. The measurement of RNA levels appeared to be a suitable method for monitoring the status of the silent loci, and furthermore revealed that the a1 transcript is spliced.

In Chapter 4, I describe experiments in which the level of silent locus transcription is monitored in a sir^{ts} strain after a shift in temperature, and the effect of inhibitors of cell cycle progress is tested. In one experiment, I asked whether blocking cell cycle progress could affect the derepression of the silent loci that occurs when the temperature is raised in a sir3^{ts} strain. The result shown in Figure 4.7a implies that SIR-mediated repression can be lost in cells which are blocked in early G_1 and are, therefore, not executing DNA replication. Hence DNA replication does not appear to be required for the derepression of the silent loci under these conditions. When the reciprocal shift (from the restrictive temperature to the permissive temperature) is performed, a different result is obtained. Cells blocked either in early G_1 or in late G_1 fail to establish repression, whereas cells blocked later in the cell cycle in late G_2 are able to repress their silent loci (Figures 4.3

and 4.6).

These results are consistent with the hypothesis that DNA replication is a critical step in establishing the repressed state, but I cannot eliminate the involvement of some other event within this period of the cell cycle. Recent studies by Brand (1986) have further defined the association between HMRE and ARS activity. Brand has shown that HMRE can be divided into three DNA elements, and the deletion of any two of these elements leads to the expression of HMR: two of these elements are ARS sequences. Taken together, these results strongly imply a role for DNA replication in the establishment of SIR-mediated repression. Possible mechanisms are considered in Section 4.6. Other examples, in which it was suspected that an ARS element might be involved in the control of transcription, have not stood up to more detailed analysis. For instance, it was for some time believed that an ARS sequence 3' to the Histone H2B gene was required for the cell-cycle dependent activation of H2B (Osley and Hereford, 1982). More recent results imply that the ARS is not necessary, and in fact the regulatory sequences are located within the 5' promoter region of H2B (Osley *et al.*, 1986). Similarly the ARS element 3' to the H0 gene (Kearsey, 1984) does not appear to be involved in transcriptional control (Nasmyth, 1985a).

In Section 4.5, I describe an experiment which asks whether a1/α2 repression can be established in G1-arrested cells. The results imply that the MATα1 gene can be repressed by a1/α2 in cells that are not undergoing DNA replication (Figure 4.7b). Chapter 5 goes on to describe work identifying the DNA sequence elements which are involved in a1/α2-mediated repression. Previous work on the H0 gene by Jensen *et al.* (1983) and by Nasmyth (outlined in Section 5.1.2 and Figure 5.1) implied that the DNA sequences responsible for a1/α2 repression must

be present a number of times within the HO promoter. I therefore performed a computer search, looking for a DNA sequence motif that was repeated within the HO promoter at positions consistent with the previous data, and that was also present within the MAT α 1 promoter (which is also repressed by a1/ α 2). This revealed the sequence TC(A/G)TGTNN(A/T)NANNTACATCA (where N denotes any nucleotide). (See Table 5.1.) In order to confirm that this sequence was responsible for a1/ α 2 control, rather than some other function, examples of this sequence were inserted into the promoter of a gene not involved in mating type, CYC1. First small restriction fragments containing this sequence motif were inserted into the CYC1 promoter (Figure 5.2), and subsequently a synthetic oligonucleotide comprising only the conserved sequence was used (Figure 5.5). In both cases the CYC1 promoter was brought under a1/ α 2 control in the constructions generated.

This work shows that DNA sequence comparisons can be used to identify the DNA sequences that are involved in a particular type of control, as long as there are a number of regions of DNA, each of which is known to contain an example of the sequence. DNA sequence comparisons have identified a number of sequence motifs shared between certain of the genes encoding ribosomal proteins (Teem *et al.*, 1984), although the functional significance of most of these motifs is not yet clear.

DNA sequence comparisons can confirm the results of other analyses; for example the sets of genes that are co-ordinately regulated by GCN4 (Hinnebusch and Fink, 1983b; Hinnebusch *et al.*, 1985) or by GAL4 (Giniger *et al.*, 1985; Bram and Kornberg, 1985) each share a common sequence motif that is responsible for their co-ordinate regulation.

FIGURE 7.1

Distribution of $\alpha 1/\alpha 2$ and $\alpha 2$ elements.

The positions of the elements (denoted by open boxes) are shown relative to the presumed structural gene starts. For the HO and MAT α genes, the UAS's and TATA boxes are also shown (Siliciano and Tatchell, 1984 Nasmyth, 1985a), as are the transcription start sites.

The $\alpha 1/\alpha 2$ elements at RME were found by Mitchell and Herskowitz, and the $\alpha 2$ elements at MFa2 and STE6 were identified by Johnson and Herskowitz (1985).

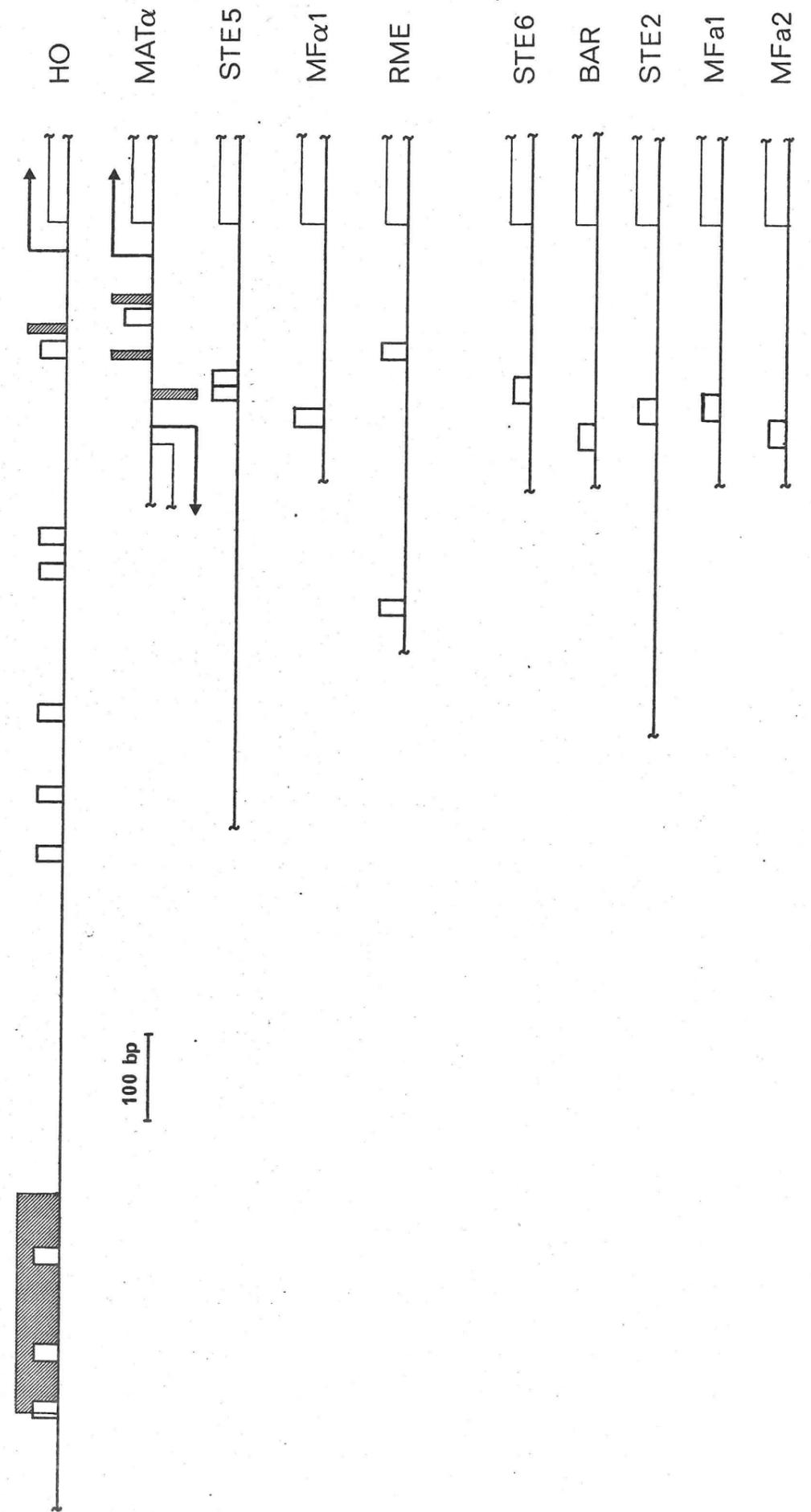


TABLE 7.1

Examples of the α_2 Element

Source	Sequence
BAR1	CATGTAATTACCGAAAAAGGAAAT TACATGG
STE2	CATGTACTTACCCAATTAGGAAATTTACATGG
MFa1	TGTGTAATTACCCAAAAAGGAAATTTACATGT
MFa2	CATGTATTTACCTATTCGGGAAATTTACATGA
STE6	CATGTAATTACCTAATAGGGAAATTTACACGC
Consensus	CATGTAATTACCNAATAAGGAAATTTACATGN

The first three sequences were identified in this work, the last two are from Johnson and Herskowitz (1985).

In Section 5.6 I identified the DNA sequence element involved in $\alpha 2$ -mediated repression using DNA sequence comparisons. The DNA sequences of the 5' flanking regions of three genes known to be under $\alpha 2$ -mediated repression were compared, and a sequence motif with the following consensus sequence was found.

GCATGTAATTACCCAAAAAGGAAATTTACATGG.

A restriction fragment containing an example of this sequence was able to bring the CYC1 promoter under $\alpha 2$ -mediated repression (Figure 5.4). Johnson and Herskowitz (1985) have shown that a similar sequence is indeed sufficient to confer $\alpha 2$ -mediated repression, and also that an $\alpha 2$ - β galactosidase fusion protein will bind to their sequence *in vitro*. In Section 5.7 and in Figure 5.7, the sequences of the $\alpha 1/\alpha 2$ and the $\alpha 2$ elements are compared, and possible models for their action are considered. Figure 7.1 shows the positions of various $\alpha 1/\alpha 2$ and $\alpha 2$ elements relative to the coding regions of the controlled genes, while Tables 5.1 and 7.1 list the sequences of $\alpha 1/\alpha 2$ elements and $\alpha 2$ elements respectively.

In Section 5.5, I describe an experiment in which a restriction fragment containing the $\alpha 1/\alpha 2$ element is inserted just upstream of the CYC1 UAS; in such a construction $\alpha 1/\alpha 2$ -mediated repression exerts only 7 - 11 fold repression (Figure 5.3), as compared with the approximately 200 fold repression observed when the same element is inserted within the promoter (Figure 7.2). A similar effect is seen in the MAT α gene, where a single UAS directs the transcription of the two divergent transcripts $\alpha 1$ and $\alpha 2$ (Siliciano and Tatchell, 1984). The $\alpha 1/\alpha 2$ element at MAT α is located to the right of the UAS, the $\alpha 1$ side (Figure 7.1). In a/α cells, MAT $\alpha 1$ transcription is completely repressed, while $\alpha 2$ transcription shows an approximately 5 fold reduction (Nasmyth *et al* 1981a). It appears probable that the

FIGURE 7.2

Different Insertions of $\alpha 1/\alpha 2$ elements into the CYC1 promoter.

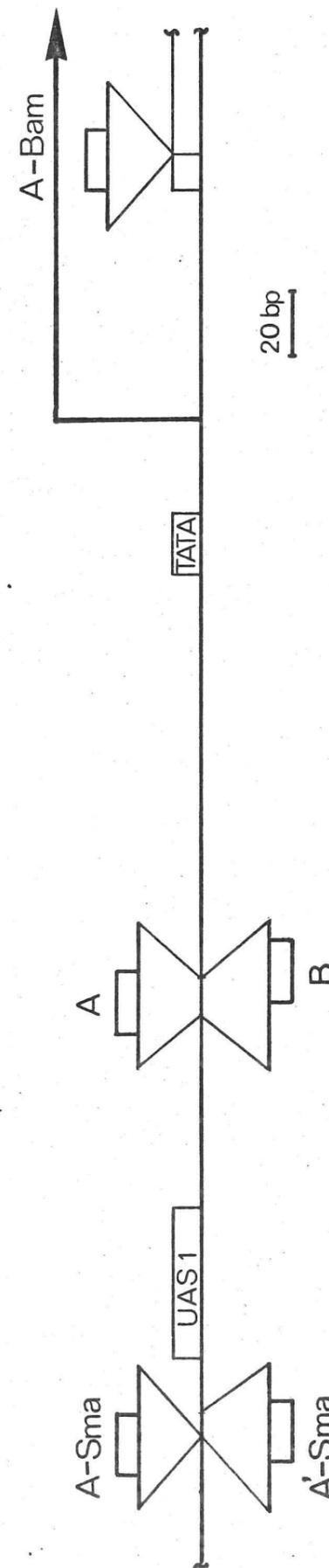
The different insertions of restriction fragments containing $\alpha 1/\alpha 2$ elements into the CYC1 promoter are shown.

A and B: Fragments A and B inserted into the XhoI/SalI/XhoI linker between UAS and the TATA box (Figure 5.2.)

A-Sma and A'-Sma: Fragment A inserted in both orientations into the SmaI site upstream of the UAS (Figure 5.3.)

A-Bam: Fragment A inserted into the BamHI site within the open reading frame (Figure 5.3.)

In each case the open box represents the position of the $\alpha 1/\alpha 2$ element within the inserted restriction fragment. The arrow denotes the transcript. The TATA box and UAS1 are also shown; UAS1 is the sequence responsible for activation when glucose is used as carbon source.



relative positions of the UAS and the a1/α2 element are responsible for this difference. These data imply that the a1/α2 element can only function to full effect when located inside the sequences necessary for transcription. The α2 element shows a similar position dependence (Johnson and Herskowitz, 1985). Siliciano and Tatchell (1986) have furthermore shown that an a1/α2 element has no effect when placed 375 bp upstream of the CYC1 UAS.

In summary, SIR repression and a1/α2 repression represent two different forms of negative control - a1/α2 repression acts only over short distances, and does not require DNA replication, whereas SIR-mediated repression, which acts over considerable distances, seems to involve DNA replication. The characteristics of a1/α2-mediated repression are reminiscent of procaryotic operators, but SIR-mediated repression is unlike any previously characterised system of repression of transcription.

Finally Chapter 6 describes the isolation of mutations in the MATα2 gene. These mutations have the curious phenotype that they are able to perform α2-mediated repression, but fail to perform a1/α2-mediated repression. Two independent isolates have been sequenced, and found to contain an identical point mutation (Figure 6.1). In itself, this "MATα2^{*}" mutation tells us little about a1/α2 and α2-mediated repression, except that such mutations can exist. It may however be a useful tool in further studies of a1/α2, such as how α2-mediated repression and a1/α2-mediated repression recognize their different DNA targets (discussed in Section 5.7). The phenotype of the mutation implies that an interaction with some a-specific factor necessary for a1/α2 control has been perturbed. Perhaps the most likely candidate for the a-specific factor is the MATa1 gene product itself. If this is true, then it may be possible to isolate an extragenic suppressor of

the MAT α 2^{*} mutation by mutagenesis of plasmid-borne MAT α 1.

Other approaches to studying the differential specificity would include investigating the importance of the spacing between the inverted repeats in the two sequence motifs, and developing an in vitro assay for each activity, such that the effects of purified factors (including the α 1 and the α 2 gene products) could be studied biochemically. It would then be possible to correlate the in vivo phenotypes of various mutations (such as MAT α 2^{*}) with their effects in vitro.

- Abraham, J., Nasmyth, K.A., Strathern, J.N., Klar, A.J.S., and Hicks, J.B. (1984) *J. Mol. Biol.* 176, 307-331.
- Abraham, J., Feldman, J., Nasmyth, K.A., Strathern, J.N., Klar, A.J.S., Broach, J.R., and Hicks, J.B. (1982) *Cold Spring Harbor Symp. Quant. Biol.* 47, 989-997.
- Alberts, B., Worcel, A., and Weintraub, H. (1977) in *The Organization and Expression of the Eukaryotic Genome* (eds. E.M. Bradbury, K. Javaherian) 165-191. (New York: Academic Press).
- Ammerer, G., Sprague, G.F., and Bender, A., (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 5855-5859.
- Astell, C.R., Ahlstrom-Jonasson, L., Smith, M., Tatchell, K., Nasmyth, K.A., and Hall, B.D. (1981) *Cell* 27, 15-23.
- Banerji, T., Rusconi, S., and Schaffner, W. (1981) *Cell* 27, 299-308.
- Bankier, A.T., and Barrell, B.G. (1983) In *Techniques in Life Sciences* B508, Elsevier Scientific Publications Ltd. pp 1-34.
- Beach, D., Piper, M., and Shall, S., (1980) *Nature* 284, 185-187.
- Beachy, P.A., Helfand, S.L., and Hogness, D.S., (1985) *Nature* 313, 545-551.
- Beggs, J.D. (1978) *Nature* 275, 104-109.
- Benoist, C. and Chambon, P. (1981) *Nature* 290: 304-310.
- Berget, S.M., Moore, C., and Sharp, P.A., (1977) *Proc. Natl. Acad. Sci. USA* 74, 3171.
- Berk, A.J., and Sharp, P.A., (1978) *Proc. Natl. Acad. Sci. U.S.A.* 75, 1274-1278.
- Betz, R., and Dunz, W., (1979) *Eur. J. Biochem.* 95 469.

- Birnboim, H.C., and Daly, J., (1979) *Nucl. Acids Res.* 7, 1513.
- Brake, A., Brenner, C., Najarian, R., Laybourn, P., and Merryweather, J., (1985) In Transport and Secretion of Proteins Cold Spring Harbor Laboratory, New York.
- Bram, R.J. and Kornberg, R.D. (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 43-7.
- Brand, A.H., Breeden, L., Abraham, J., Sternglanz, R., and Nasmyth, K., (1985) *Cell* 41, 41-48.
- Brand, A.H., (1986) Ph.D. Thesis, Univ. of Cambridge.
- Breeden, L., and Nasmyth, K.A., (1985) *Cold Spring Harbor Symp. Quant. Biol.* 50, 643.
- Bromley, S., Hereford, L., and Rosbash, M., (1982) *Mol. Cell. Biol.* 2, 1205-1211.
- Bucking-Throm, E., Dunzle, W., Hartwell, L.H., and Manney, T.R., (1973) *Exp. Cell Res.* 76 99-110.
- Carrasco, A.E., McGinnis, W., Gehring, W.J., and de Robertis, E.M. (1984) *Cell* 37, 409.
- Carroll, S.B., and Scott, M.P., (1985) *Cell* 43, 47-57.
- Celniker, S.E., and Campbell, J.L., (1982) *Cell* 31, 201-213.
- Chan, C.S.M., and Tye, B-K., (1980) *Proc. Natl. Acad. Sci. USA* 77, 6329.
- Chan, R., (1977) *J. Bacteriol.* 130 766.
- Chen, W., and Struhl, K., (1985) *EMBO Jou.* 4, 3273-3280.
- Chow, L.T., Gelinis, R.E., Broker, T.R., and Roberts, R.J., (1977) *Cell* 12, 1-8.

- Contreras,R., Gheysen,D., Knowland,J., van de Voode,A., and Fiers,W.
(1982) Nature 300, 500-505.
- DiNardo,S., Kuner,J.M., Theis,J., and O'Farrell,P.H., (1985) Cell 43, 59.
- Duckworth,M.L., Gait,M.J., Goelet,P., Fang Hong,G., Singh,M., and
Titmas,R.C., (1981) Nucleic Acids Res. 9, 1691-1706.
- Dunzte,W., MacKay,V.L., and Manney,T.R., (1970) Science 168 1472.
- Dunzte,W., Stotzler,D., Bucking-Throm,E., and Kalbitzer,S., (1973) Eur. J.
Biochem. 35 357.
- de Crombrugge,B., Mudryj,M., DiLauro,R., and Gottesman,M. (1979) Cell 18,
1145-1151.
- Elliott,T., and Geiduschek,E.P. (1984) Cell 36, 211-219.
- Fields,S., and Herskowitz,I., (1985) Cell 42, 923-930.
- Fried,H.M., and Warner,J.R., (1982) Nucleic Acids Res. 10, 3133-3148.
- Gait,M.J. (1984) Ed. Oligonucleotide Synthesis IRL PRESS, Oxford.
- Geiduschek,E.P., Elliott,T., and Kassavetis, G.A. (1980) in Bacteriophage T4
(eds Mathews,C.K., Kutter,E.M., Mosig,G., and Berget,P.B.) 189-192
(Washington D.C. Amer. Soc. for Microbiol.)
- Gilbert,W., and Muller-Hill,B. (1970) in The lactose operon (ed.
J.R.Beckwith and D.Zipser) 93-109 (Cold Spring Harbor Laboratory
Cold Spring Harbor, New York.)
- Giniger,E., Varnum,S.M., and Ptashne,M., (1985) Cell 40, 767-774.
- Guarente,L., and Ptashne,M. (1981) Proc. Natl. Acad. Sci. U.S.A.
78,2199-2203.

- Guarente,L. Yocum,R.R., and Gifford,P. (1982) Proc Natl. Acad. Sci. U.S.A. 79, 7410-7414.
- Guarente.L., and Mason,T., (1983) Cell 32, 1279-86.
- Guarente,L., and Hoar,E., (1984) Proc. Natl. Acad. Sci. USA 81 7860-4.
- Guarente,L., Lalonde,B., Gifford,P., and Alani,E., (1984) Cell 36 503-11.
- Haber,J.E., and George,J.P. (1979) Genetics 93, 13-35.
- Hahn,S. Hoar,E.T. and Guarente,L. (1985) Proc. Natl. Acad. Sci. USA 82 8562-6.
- Hanahan,D., (1983) J. Mol. Biol. 166, 557.
- Hartwell,L.H., (1976) J.Mol. Biol. 104, 803-817.
- Hartwell,L.H., (1980) J. Cell. Biol. 85 811-822.
- Hawthorne,D.C., (1963a) Abstr.Proc. 11th Int.Congr.Genet. 7, 34-35.
- Hawthorne,D.C., (1963b) Genetics 48, 1727-29.
- Hentschel,C., Irminger,J.-C., Bucher,P., and Birnstiel,M.L., (1980) Nature 285, 147-151.
- Hereford,L., and Hartwell,L.H., (1974) J. Mol. Biol. 84,445-461.
- Hereford,L., Bromley,S., and Osley,M.A. (1982) Cell 30, 305-310.
- Herman,A.I., (1971) Antonie van Leeuwenhoek J. Microbiol. Serol. 34 379.
- Hicks,J.B., (1975) Ph.D. Thesis, University of Oregon, Eugene.
- Hicks,J.B., and Herskowitz,I., (1976a) Nature 260 246.
- Hicks,J.B., and Herskowitz,I., (1976b) Genetics 83 245-58.

- Hicks, J.B. and Herskowitz, I. (1977) *Genetics* 85 373-93.
- Hicks, J.B., Strathern, J.N., and Klar, A.J.S., (1979) *Nature* 282 478-83.
- Hicks, J.B., Strathern, J.N. and Herskowitz, I. (1977) In DNA Insertion elements, Plasmids and Episomes, Ed. A. Bukhari, J. Shapiro, S. Adhya, pp 457-62. New York Cold Spring Harbor Lab.
- Hinnebusch, A.G., and Fink, G.R. (1983a) *Proc. Natl. Acad. Sci. U.S.A.* 80, 5374-5378.
- Hinnebusch, A.G., and Fink, G.R., (1983b) *J. Biol. Chem.* 258, 5238-47.
- Hinnebusch, A.G., Lucchini, G., and Fink, G.R., (1985) *Proc. Natl. Acad. Sci. U.S.A.* 82, 498-502.
- Holmes, D.S., and Quigley, M., (1981) *Anal. Biochem.* 114, 193.
- Hopper, A.K., and Hall, B.D., (1975) *Genetics* 80 41-59.
- Ivy, J., Klar, J.J.S., and Hicks, J.B., (1986) *Mol. Cell. Biol.* 6, 668.
- Jacob, F. and Monod, J., (1961) *J. Mol. Biol.* 3, 318-356.
- Jensen, R., (1983) Ph.D. Thesis, University of Oregon.
- Jensen, R., Sprague, G.F., and Herskowitz, I., (1983) *Proc. Natl. Acad. Sci. U.S.A.* 80, 3035-3039.
- Jensen, R., and Herskowitz, I. (1984) *Cold Spring Harbor Symp. Quant. Biol.* 49, 97.
- Johnson, A.D., Poteete, A.R. Lauer, G. Sauer, R.T. Ackers, G.K. and Ptashne, M. (1981) *Nature* 294, 217-223.
- Johnson, A.D., and Herskowitz, I., (1985) *Cell* 42, 237-247.

- Johnston, S.A. and Hopper, J.E. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 6971-6975.
- Kassir, Y., and Simchen, G., (1976) Genetics 82 187-206.
- Kearsey, S., (1984) Cell 37, 299-307.
- Keller, J.M., and Alwine, J.C., (1984) Cell 36, 381-389.
- Kilmartin, J.V., (1981) Bioch. 20, 3629-3633.
- Kitchingman, G., Lai, S., and Westphal, H. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 4392-5.
- Klar, A.J.S., Fogel, S., and Macleod, K. (1979a) Genetics 93, 37-50.
- Klar, A.J.S., Fogel, S., Radin, D.N., (1979b) Genetics 92 759.
- Klar, A.J.S., (1980) Genetics 95, 631-648.
- Klar, A.J.S., Strathern, J.N., Broach, J.B., and Hicks, J.B., (1981a) Nature 289 239.
- Klar, A.J.S. Strathern, J.N. and Hicks, J.B. (1981b) Cell 25, 517-24.
- Klar, A.J.S., Hicks, J.B., and Strathern, J.N., (1982) Cell 28, 551.
- Klar, A.J.S., Strathern, J.N. and Hicks, J.B. (1984) In Microbial Development Eds. R.Losick and L.Shapiro. Cold Spring Harbor Laboratory. 151-195.
- Klug, A., and Lutter, L.C., (1981) Nucl. Acids Res. 9, 4267-83.
- Kostriken, R., Strathern, J.N., Klar, A.J.S., Hicks, J.B., and Heffron, F. (1983) Cell 35, 167-174.
- Kuroiwa, A. Kloter, U. Baumgartner, P. and Gehring, W.J., (1985) EMBO Jou. 4, 3757-3764.

- Langford,C., and Gallwitz,D., (1983) Cell 33, 519-527.
- Laughon,A., and Scott,M.P., (1984) Nature 310, 25-30.
- Laughon,A. and Gesteland,R.F. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 6827-6831.
- Lawrence,P.A., and Morata,G. (1983) Cell 35, 595-601.
- Levi,J.D., (1956) Nature 177 753.
- Lindgren,C.C., and Lindgren,G., (1943) Proc. Natl. Acad. Sci. USA 29 306.
- Losson,R., and Lacroute,F. (1983) Cell 32, 371-377.
- Mackay,V.L., and Manney,T.R., (1974a) Genetics 76, 255-271.
- Mackay,V.L., and Manney,T.R., (1974b) Genetics 76, 273-288.
- Mackay,V.L., (1983) Methods in Enzymol. 101, 325.
- Maniatis,T., Fritsch,E., and Sambrook,J., (1982) Molecular Cloning a Laboratory Manual published by Cold Spring Harbor Press, NY.
- McGinnis,W., Garber,R.L., Wirz,J., Kuroiwa,A., and Gehring,W.J. (1984b) Cell 37, 403-408.
- McGinnis,W., Levine,M.S., Hafen,E., Kuroiwa,A., and Gehring,W.J. (1984a) Nature 308, 428-433.
- Messing,J., (1979) Recombinant DNA Technical Bulletin 2, 43-48.
- Messing,J., and Vieira,J., (1982) Gene 19, 269-276.
- Miller,A.M., (1984) EMBOJ, 3, 1061-1065.
- Miller,A.M., Sternglanz,R., and Nasmyth,K.A. (1984) Cold Spring Harbor Symp. Quant Biol. 49, 105-113.

- Miller, J.H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory Press, New York 352-355.
- Miller, J.H., and Reznikoff, W.W. (1978) Eds. The Operon Cold Spring Harbor Laboratory Press, New York.
- Mitchell, A.P., and Herskowitz, I., (1986) Nature 319, 738-742.
- Mortimer, R.K., and Hawthorne, D.C., (1969) Yeast Genetics. In The Yeasts Ed. A.H. Rose and J.S. Harrison 1, 385. New York Academic.
- Nagawa, F., and Fink, G.R., (1985) Proc. Natl. Acad. Sci. USA 82 8557-61.
- Nakayama, N., Migajima, A., and Arai, K., (1985). EMBO Jou. 4, 2643-2648.
- Nasmyth, K.A. and Tatchell, K. (1980) Cell 19 753.
- Nasmyth, K.A. Tatchell, K. Hall, B.D., Astell, C.R. and Smith, M. (1981a) Nature 289, 244-250.
- Nasmyth, K.A. Tatchell, K., Hall, B.D. Astell, C.R. and Smith, M., (1981b) Cold Spring Harbor Symp. Quant. Biol. 45 961-81.
- Nasmyth, K.A., (1982a) Annu. Rev. Genet. 16, 439-500.
- Nasmyth K.A., (1982b) Cell 30, 567-78.
- Nasmyth, K.A., (1983) Nature 302, 670-676.
- Nasmyth, K.A., (1985a) Cell 42, 213-222.
- Nasmyth, K.A., (1985b) Cell 42, 225-235.
- Newlon, C.S., and Burke, W., (1980) in Mechanistic studies of DNA replication and Genetic Recombination ICN-UCLA Symposia on Cellular and Molecular Biology, eds Alberts, B., and Fox, C.F., (Academic New York) Vol 19, p309ff.

- Orr-Weaver, T.L., Szostak J.W. and Rothstein, R.J. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 6354-6358.
- Oshima, Y., and Takano, I., (1971) Genetics 67 327-35.
- Osley, M.A., and Hereford, L., (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 7689-93.
- Osley, M.A., Gould, J., Kim, S., Kane, M., and Hereford, L., (1986) Cell 45, 537-44.
- Pikielny, C.W., Teem, J.L., and Rosbash, M., (1983) Cell 34, 395-403.
- Poole, S.J., Kauvar, L.M., Drees, B., and Kornberg, T., (1985) Cell 40, 37-43.
- Porter, S.D., and Smith, M., (1986) Nature 320, 766-8.
- Pribnow, D. (1979) In Biological regulation and Development Vol. 1, 219-277
- Quinlan, R.A., Pogson, C.I., and Gull, K. (1980) J. Cell. Sci. 46, 341-352.
- Reed, S.I., (1980) Genetics 95, 561-577.
- Rine, J.D., (1979) PhD thesis. Univ. Oreg., Eugene.
- Rine, J., Strathern, J.N., Hicks, J.B., and Herskowitz, I. (1979) Genetics 93, 877-901.
- Rosbash, M., Harris, P.K.W., Woolford, J.L. and Teem, J.L., (1981) Cell 24, 679-686.
- Rothstein, R. (1983) Methods in enzymology, 101, 202.
- Salstrom, J.S., and Szybalski, W. (1978) J. Mol. Biol. 124, 195-21.
- Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc Natl. Acad. Sci. USA 74, 5463-5467.
- Sarokin, L., and Carlson, M., (1984) Mol. Cell. Biol. 4 2750-7.

- Sauer, R.T. Yocum, R.R. Doolittle, R.F. Lewis, M. and Pabo, C.O. (1982) Nature 298, 447-451.
- Scott, M.P. and Weiner, A.J. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 4115-4119. (1984)
- Shatkin, A.J., (1976) Cell 9, 645.
- Shepherd, J.C.W. McGinnis, W. Carrasco, A.E. de Robertis, E.M. and Gehring, W.J. (1984) Nature 310, 70-71.
- Shore, D., Squire, M. and Nasmyth, K.A., (1984) EMBO Jou. 3, 2817.
- Siliciano, P.G., and Tatchell, K., (1984) Cell 37 969-978.
- Siliciano, P.G. and Tatchell, K. (1986) Proc. Natl. Acad. Sci. U.S.A. 83, 2320-2324.
- Singh, A. Chen, E.Y. Lugovoy, J.M.L. Chang, C.N. Hitzeman, R.A. and Seeburg, P.H. (1983) Nucleic Acids Res. 11, 4049-4063.
- Southern, E.M., (1975) J. Mol. Biol. 98, 503.
- Sprague, G.F., Jensen, R. and Herskowitz, I., (1983) Cell 32 409.
- Staden, R., (1982) Nucleic Acids Res. 10, 2951-2961.
- Staden, R. (1984) Nucleic Acids Res. 12, 505-519.
- Stinchcomb, D.T. Struhl, K. and Davies, R.W. (1979) Nature 282, 39-43.
- Stotzler, D. and Dunzter, W. (1976) Eur. J. Biochem. 65 257.
- Strathern, J.N. Blair, L.C. and Herskowitz, I. (1979) Proc. Natl. Acad. Sci U.S.A. 76 3425-29.
- Strathern, J.N. Spatola, E. McGill, C. and Hicks, J.B. (1980) Proc. Natl. Acad. Sci. USA 77 3839.

- Strathern, J.N. Hicks, J.B. and Herskowitz, I. (1981) *J. Mol. Biol.* 147, 357-372.
- Strathern, J.N. Klar, A.J.S. Hicks, J.B. Abraham, J.A. Ivy J.M. Nasmyth, K.A., and McGill, C., (1982) *Cell* 31, 183-92.
- Struhl, G., (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79, 7380-7384.
- Struhl, K., Stinchcomb, D.T., Scherer, S., and Davies, R.W., (1979) *Proc. Natl. Acad. Sci. U.S.A.* 76, 1035.
- Struhl, K. (1985) *Proc. Natl. Acad. Sci. USA* 82 8419-24.
- Takano, I., and Oshima, Y., (1967) *Genetics* 57 875.
- Takano, I., and Oshima, Y. (1970a) *Genetics* 65 421.
- Takano, I., and Oshima, Y., (1970b) *Genetics* 64 229.
- Tatchell, K., Nasmyth, K.A., Hall, B.D., Astell, C.R., and Smith, M., (1981) *Cell* 27 25-35.
- Teem, J.L. Rodriguez J.R. Tung, L. and Rosbash, M. (1983) *Mol. Gen. Genet.* 192, 101-103.
- Teem, J.L. Abovich, N. Kaufer, N.F. Schwindinger, W.F. Warner, J.R. Levy A Woolford, J., Leer, R.J., van Raamsdonk-Duin, M.M.C., Mager, W.H., Planta, R.J. Schultz L. Friesen, J.D. Fried, H. and Rosbash, M., (1984) *Nucl. Acids Res.* 12, 8295-8312.
- Thomas, G.P., and Mathews, M.B. (1980) *Cell* 22, 523-533.
- Thorner, J., (1981) In Molecular Biology of the yeast *Saccharomyces cerevisiae* Ed. J. Strathern, E-W Jones, J.R Broach. pp143-180. New York. Cold Spring Harbor Lab.
- Tkacz, J.S., and MacKay, V.L., (1979) *J. Cell. Biol.* 80 326.

Wharton, K.A., Yedvobnick, B., Finnerty, V.G. and Artavanis-Tsakonas, S. (1985) *Cell* 40, 55-62.

White, R.A., and Wilcox, M., (1984) *Cell* 39, 163-171.

Wilkinson, L.E., and Pringle, J.R. (1974) *Exp. Cell Res.* 89 175.

Wilson, K.L., and Herskowitz, I., (1984) *Mol. Cell. Biol.* 4 2420-27.

Brent, R., and Ptashne, M., (1984) *Nature* 312, 612-5.

Brown, D.D., (1984) *Cell* 37, 359-65.

Feldman, J.B., Hicks, J.B., and Broach, J.R., (1984) *J.Mol.Biol.* 178, 815-34.

"Detection is, or ought to be, an exact science, and should be treated in the same cold and unemotional manner. You have attempted to tinge it with romanticism, which produces much the same effect as if you worked a love story or an elopement into the fifth proposition of Euclid."

"But the romance was there." I remonstrated, "I could not tamper with the facts."

"Some facts should be suppressed....."

A. Conan Doyle, in *The Sign of Four*.